

MODEL VALIDATION: INDUSTRY & REGULATORY PERSPECTIVES

Last month, we reviewed the notion of model risk. As the complexity and consequences of financial risk evolve, risk models must necessarily progress. The Financial Times (1) now publishes a series called “Mastering Risk: Welcome to this week’s one in a million event”, which deftly illustrates the escalating difficulties of risk management. This ever-increasing risk has resulted in a related ever-increasing reliance on computer-based models to manage risk, which has created this new form of risk, model risk.

In order to assist U.S. national banks in managing, mitigating, and perhaps model risk, the OCC issued BULLETIN 2000-16: RISK MODELING, DESCRIPTION: MODEL VALIDATION (2). This month, we review several model validation issues. In order to answer the Who, What, When, Where, Why, and How of model validation we again refer to the “best practices” approach outlined in RISK STANDARDS FOR INSTITUTIONAL INVESTMENT MANAGERS AND INSTITUTIONAL INVESTORS (3) as well as the OCC’s Bulletin.

Who should perform the validation?

According to *Risk Standard 19: Independent review of methodologies, models, and systems*, management “should appoint an independent third party to review the soundness of valuation methodologies, models, and related systems. If this is not possible, an internal group should perform such assessments, with appropriate checks and balances”.

In many regional banks, and virtually all community banks, the level of expertise required for proficient model assessment typically resides solely within the specific modeling area. This tends to be as true in A/LM as it is in credit scoring or marketing. The OCC specifically notes that “model validation is often outside the scope of audit responsibilities”. In all cases, however, Internal Audit still has a valuable role to play in ensuring data integrity, certifying policy

compliance, reviewing documentation, and ensuring management oversight.

Creditable and independent review is thus typically outsourced. As Captain Renault says in CASABLANCA (4), let’s “round up the usual suspects” (in alphabetical order):

- *Accounting firms*
Many of the larger accounting firms have a sizeable risk management-consulting department dedicated to the financial services industry. They can provide comprehensive life-cycle assistance from model selection and implementation through audit and validation. In addition, they can offer an integrated approach to related risk management and audit issues. Regional and larger banks occasionally follow this approach. The drawback for many regional and community banks is that the cost of this approach is frequently greater than the entire annual A/L departmental budget. In smaller institutions, this is coupled with the concern that the consulting firm may be developing their staff’s expertise on the bank’s time and dime.
- *Broker/Dealers*
Several broker/dealers offer outsourced A/L modeling and validation on a fee-based or relationship basis. Their broad-based market experience is an advantage. Possible concerns include objectivity and relevant experience. These concerns can be allayed by the selection of credentialed professionals (such as CFAs or CPAs) whose certification includes a formalized ethics program, coupled with verification of prior hands-on banking A/L management and modeling experience.
- *Consulting firms specializing in risk management or A/LM*
They can also provide comprehensive modeling assistance from selection and implementation through audit and validation. The largest of banks use enterprise risk management consultants, whereas community and regional banks more typically utilize the services of specialized A/L consultants. The benefits of this approach include

an independent, unbiased perspective, coupled with a broad-based view of a wide cross-section of depository institutions. Familiarity with industry “best practices” and regulatory developments are yet additional advantages of using specialized consultants. Cost and other concerns are similar to those noted above.

- *Software vendors*

Many of the leading A/L models offer consulting and validation services based on the strength of their model-specific expertise. Whereas the other outsourced solutions also offer valuable insight into the appropriateness of the A/L model used by a bank, this approach usually does not offer that objective point of view. Much as banks outgrow core and other systems, many banks outgrow their original A/L model. This form of model risk is increasingly seen as the industry consolidates and small banks become large banks. This is exacerbated due to increasing risk complexity at all sizes of banks, due, in part, to the rapid evolution of option-laden assets and liabilities and their proliferation on bank balance sheets.

What should the validation include?

Risk Standard 16: Assessing model risk specifies model factors that are relevant to bank A/L model validation:

- *Data integrity.*
- *Cash flow definitions.* Are interest rate contingent cash flows scenario-dependent and option-adjusted?
- *Algorithms, formulae, or other mathematical models.* Does the yield curve model use linear interpolation or a more sophisticated cubic spline methodology? Are projected yield curves based on a single factor model such as Black-Derman-Toy or multifactor model such as Brace-Gatarek-Musiela? Are ARM prepayments based on a simple constant prepayment rate or a more robust model such as Espiel or Andrew Davidson?
- *Going concern or liquidation approach.* If the liquidation approach is used, are time, liquidity

and spread assumptions defined? Do they differ from going concern assumptions?

- *Model specific parameters.* These include, but are not limited to:
 - Yield curves. For example, are implied forward curves used to derive values? If flat rate or economic forecast scenarios are used, calculated values may be difficult to reconcile with actual market values and may not be suitable for external disclosure (see SEC Market Risk Disclosure FAQ’s available at www.sec.gov)
 - Rates and spreads. Are discount rates for value calculations market-based, rather than institution-specific? Are wholesale market spreads updated frequently? If institution-specific spreads change in different rate environments, say for Core Deposits, are these spreads changed in the modeling process manually or via formula? If via formula, are equation parameters updated and back-tested?
 - Scenario selection and stress test variables. Are reasonable and realistic interest rate changes modeled based on historic or implied rate volatilities? Are non-parallel yield curve shifts considered? For example, simplistic shock scenarios utilizing the entire yield curve are easy to model, but rarely, if ever, occur in the U.S. The use of non-realistic rate shock scenarios in the 2000s is akin to only using a single scenario Gap report in the 1990s (or playing golf with only one club). Are multiple balance sheet scenarios considered? Are important assumptions evaluated via sensitivity analysis? Are stress tests severe enough? One community bank, for example, stress tests its model and outcomes by simulating a 600 basis point Prime rate move because the C.E.O. remembers that such a move is possible as it happened in the early 1980s. This pragmatic view is consistent with the article on “Welcome to this week’s one in

a million event” in the Financial Times noted earlier.

- Probability intervals and correlation assumptions. Are probability and confidence intervals wide enough or too wide (see above for Prime rate note and the following “normal” definitions)? Are regression and/or correlation estimates varied to ascertain model robustness and sensitivities?
- Time horizon. Is only one year of earnings simulated? Frequently mismatch and basis risk develop over longer time horizons. As above, is the overarching approach that of liquidation or going concern?

The validation should also include a review of reporting and report terminology. If reports are developed “outside the model” due to model limitations, process and validation documentation should note the degree of supplementary effort, including any manual input (a frequent source of model risk, as the author can attest). The use of an outside perspective on internal reporting is constructive due to a broader experience set, which can result in the design and use of more effective and/or easier to understand reports.

The final “what” concerns terminology; clarification of terminology is seen as the province of semantic professors, but it is also the domain of risk reporting professionals. The term “normal” to a Value at Risk analyst may refer to a value profile that is normally distributed. To this professional, normal, at a 95% confidence interval, may mean that a bank exceeds policy limits on trading book losses once every twenty days. Alternatively, it may mean a bank that goes into default one out of every twenty times (or years). To a C.E.O., “normal” may mean ordinary, usual, or customary. The notion of exceeding trading policy limits certainly doesn’t meet this definition of “normal”. For a C.E.O. that intends to have a twenty-year tenure, the notion of default during that time frame certainly isn’t “normal”. Thus, the clarification of both technical and managerial terminology and

expectations may be a beneficial byproduct of the model validation process.

When, or how often, should the validation take place?

The OCC, in *bulletin 2000-16*, suggests that ongoing validation is an integral element of sound validation policy. We do not believe that they mean that every part of every model needs to be continuously reviewed; rather we believe that validation should be activity and risk appropriate, in terms of both timing and process. Bulletin 2000-16 delineates three primary components of models:

1. The information input component, which includes data and assumptions. Most banks validate the data input process continuously – that is, do model balances equal the General Ledger, external reporting, and subsystem balances? Assumptions, while equally critical, tend to be validated on a periodic basis, if ever. For both data and assumptions, documentation is paramount.
2. The processing component, which includes financial theory and computer models. In this instance, validation should focus on post-implementation, pre-production testing. As new business lines or products are added, pre-production validation should be used to ascertain whether the model in use is still sufficient. This frequently involves ongoing and/or periodic benchmarking to alternative models or to historical results.
3. The reporting component, especially the management-reporting component. Is management aware of model limitations (see following “why” comment) or the context in which reported results are relevant (for example, unrealistic parallel instantaneous yield curve shocks, with static rates thereafter)? Do non-technical managers and Board members readily understand the limitations and results? Do ALCO reports have to be boring (hopefully not)?

As noted last month, *Risk Standard 15: Back testing* suggests that risk and return forecasts be back tested at least quarterly. While back-testing is seen as the bane of many a forecaster, it is an effective means to improved modeling. Model calibration, or fine-tuning, can emanate from comparison of actual results to model estimates. For example, if a bank uses generic consensus prepayment estimates for its' residential loan portfolio (or securitized equivalents) and its' review demonstrates that prepayments are always 8%-10% lower, it would behoove the bank to scale future prepayment forecasts. In a similar vein, some banks scale their marketing department and/or line manager estimates of new or budgeted volumes. We further review "back testing" in the "how" section below.

Where

Regulators and auditors typically combine on site review with off site analysis and reporting. It is prudent for model validation professionals to follow a similar regime.

Why

As the use of computer models as a basis for decision-making increases, so does the potential for losses due to model risk. The consequences of model risk can be dramatic, as Meridien Research and Capital Market Risk Advisors estimated model-risk related losses at \$5.5 billion during 1999 (2, *ibid*). Validation reduces the risk of fraud, which is frankly more of a concern for trading and valuation models than for A/L models, although the same executive typically oversees both functions.

Validation also helps reduce the risk of GIGO, or Garbage In Garbage Out. According to the OCC, the use of a vendor-provided model or outsourced reporting does not absolve a bank from the task of model validation. Nonetheless, the largest outsourced provider of A/L reports in the U.S. (producing reports for several thousand of the roughly 10,000 banks during 2000) uses a default prepayment rate of 6% CPR in the base case for loans (for 1998-2000 per available on-line documentation). One wonders how many of the banks had the outsourced provider

override this default and/or informed their ALCOs and Boards of this assumption and the impact of its use. Certainly, few of these banks had their residential portfolio prepay at 6% during 1998 or early 2001. Certainly, the use of this default must have materially misstated the impact of their estimated base case and scenario-dependent rate shock value profile in at least some of the banks.

How

Auditors and examiners tend to utilize checklists in their work papers. As an example, consider a grid approach to model validation based on the OCC bulletin 2000-16 (see table 1). This ABC Grid, or Assessment, Benchmarking, and Comparative Procedures approach is also effective at the micro, or individual model line level, as well as at the macro, or institution level. That is, as you validate or review individual accounts, the "what" should change, but not the "how".

The "what" is influenced by the relative importance of market and other factors to a given model account. When considering residential loans, or their securitized equivalents, one of the more significant "whats", or assumptions, is the scenario-dependent prepayment estimate. FNMA 7.5% MBS, corresponding to 1999 residential loan production originated at approximately 8.12%, prepaid at an 8% annual rate during the fourth quarter of 2000, but is projected to increase to a 24% annual prepayment rate by the fourth quarter of 2001 (5). This threefold increase in estimated prepayments can have a material impact on a given banks' liquidity, earnings, and value measures, thus the documentation, review, and validation of this assumption is not trivial. The validation process should reveal whether this marketplace change is relevant for similar loans (or pools) in a given institution.

Another "how" that is frequently mentioned, but less frequently undertaken is "back testing". Many institutions are already using an effective comparative tool for back testing, the familiar rate/volume report. While more advanced approaches such as attribution

analysis has supplanted the use of this tool in some organizations, management may more readily understand enhancements and extensions to a familiar format. The rate variance component can be decomposed into several factors based on model “key” or “driver” rates. In addition, where the rate variance is based on yield curve shifts, this variance can be decomposed into key rate durations or “risk point measures” (6).

Thus, the task of evaluating estimated value changes can be enhanced by the use of key rate durations, especially when coupled with the use of calculated portfolio durations. The use of correlations of various key rate and portfolio durations can further reduce the key duration set, thus abridging the analytical task. This approach can result in the use of as few as three “key rates” or “risk points” to explain most of the rate variance. Alternatively, one could calculate level, slope, and curvature duration variances (7), but a variance analysis using this technique may be problematic, especially when introducing it in an ALCO or Board setting.

A macro-level “back test” on the total and components of the net interest margin can be readily completed using a related technique, that of reducing “key” or “driver” rates to a manageable few. Twenty to thirty quarters of data regressed against this handful of key rates, appropriately lagged, is frequently enough to produce statistically significant correlation and regression beta measures. This approach may be complicated by structural shifts in an organizations’ balance sheet, in which case a more micro approach is required.

One more who: Who is the validation for?

The management of risk impacts all organizational stakeholders; thus all stakeholders benefit from the validation of financial models, including the designers and users of models. It is worth noting that a prime beneficiary of model validation is executive management, as it may provide the “peace of mind” regarding financial models that they may have lacked

to date. If the only benefit is to certify that the model is materially correct, it may result in greater reliance of the model and the modeling team, which can be a mixed blessing, as many readers of Bank A/LM can readily attest.

FRED POORMAN JR., CFA

Endnotes

1. available at ft.com
2. available at cmra.com
3. available at occ.gov
4. see an excellent film
5. see Bloomberg Financial
6. see Ho, T., “Key Rate Durations: Measures of Interest Rate Risks”, *The Journal of Fixed Income*, September 1992, Vol. 2. No. 2; see also Dattatreya, R.& Fabozzi, F., “The Risk Point Method for Measuring and Controlling Yield Curve Risk” *Financial Analysts Journal*, July-August 1995, Vol. 51, No. 4; see also Golub, B. & Tilman L., “Measuring Yield Curve Risk Using Principal Components Analysis, Value At Risk, and Key Rate Durations”, *The Journal Of Portfolio Management*, Summer 1997, Vol. 23. No. 4.
7. see Willner, R. “A New Tool for Portfolio Managers: Level, Slope, and Curvature Durations”, *The Journal of Fixed Income*, June 1996, Vol. 6, No. 1;

ABC VALIDATION PROCEDURES Model Components	ASSESSMENT PROCEDURES: INDEPENDENT REVIEW OF LOGICAL AND CONCEPTUAL SOUNDNESS	BENCHMARKING PROCEDURES: EVALUATE RESULTS VERSUS THE MARKET AND OTHER MODELS	COMPARATIVE PROCEDURES: COMPARE MODEL ESTIMATES TO ACTUAL RESULTS (BACKTESTING)
Input Component: Data and Assumptions	Data: Is input automated or manual? Do balances tie to the ledger? Assumptions: Are the sources for yield curve, spreads, and prepayment estimates documented and reasonable? What is the source?	Assumptions: How do the scenario-dependent assumptions and estimates compare to historical and/or marketplace pricing estimates? Are non-parallel yield curve moves and basis risk considered? Is additional sensitivity analysis periodically performed?	Assumptions: Are assumptions and other sub-model projections compared to actual results? Is this monitored on a regular basis? Is there a consistent directional bias over time?
Processing Component: Financial Theory and Computer Models	Financial Theory: Are the models consistent with: * ALCO's business approach * Market and/or academic models? Computer Model: Has the underlying theory been migrated into software component?	Computer Model: Does the model price securities consistent with the consensus or marketplace pricing estimates in all scenarios? Are values benchmarked to another model (e.g. Bloomberg, Yield Book, BondEdge)?	Computer Model: Have projected and actual interest income been compared via a rate/volume or other analysis? Have scenario-dependent total return estimates been realized?
Reporting Component: Integration with Management Information Systems (MIS)	Integration with MIS: Is the report readily comprehensible to ALCO? Is ALCO aware of material model limitations? Is the level of reporting reasonable and sufficient? Is the account stratification consistent with internal reporting and/or marketplace convention?	Integration with MIS: Is the report's content and form consistent with industry convention or regulatory and/or internal reporting? Is additional sensitivity analysis periodically performed? Are results compared to policy limits and external benchmarks?	Integration with MIS: Is ALCO provided with periodic actual vs. projected reports? Is the modeling and reporting set rich enough to allow for line-, portfolio-, and institution-level back testing? Are projections and results benchmarked to marketplace and peers?