



RMA's Model Validation and Governance Survey *Points to Best-Practices Benchmark*

•• The survey of **45 global institutions** also distinguishes the areas that set leading practitioners apart from other institutions. Given the recent market turmoil, it also underscores the need for financial institutions to scrutinize more carefully and back-test the models that they use to validate their underlying assumptions.

RMA is grateful to the members of its 2008-09 Council on Market Risk for undertaking this initiative:

- **Bryan Osmar**, senior vice president, Market & Trading Credit Risk, Royal Bank of Canada, chair of RMA's Market Risk Council
 - **Aleem Gillani**, chief market risk officer, SunTrust Banks
 - **Joseph Masri**, vice president & head of Investment Risk Management, Canada Pension Plan Investment Board
 - **Murray McIntosh**, senior vice president, CIBC
 - **Paul Sassieni**, senior vice president, Enterprise Risk Management, State Street Corporation
 - **Sharon Schick**, vice president/Credit Risk Management & Advisory, Goldman Sachs & Co.
-

RMA staff members contributing to the study were Fran Garritt, Curtis Knight, Stephen Revucky, Loretta Spingler, and Kim Gordon. The writing of the final report was undertaken by RMA.

RMA would also like to thank David Fehrenbacher, First Horizon National Corporation, Yong Kim and Russell Cottam, Key Corp, for their help in crafting and editing the survey, vetting the final online version, and providing clarification on issues during the course of the survey. Without their help, this survey would not have been possible. The executive summary for this survey was written by James G. Rickards, managing director, Omnis, Inc. and Francis Garritt, RMA Market Risk manager. In addition to Garritt, other RMA staff members contributing to the study were Curtis Knight, Stephen Revucky, Loretta Spingler, and Kim Gordon.

IN A SURVEY conducted last spring, RMA sought to determine the practices and standards used to manage model risk of financial institutions in the United States, Canada, Europe, Asia, and Australia. The survey responses, in addition to providing benchmark data, offered valuable insights into common industry practices and the degree to which certain validation techniques have achieved general acceptance and standardization.

Specifically, the Model Validation and Governance survey yielded information about what impact, if any, the recent credit crisis and market disruptions have had on respondents' approaches to managing model risk—specifically, their methods of model validation. Based on their responses, institutions might want to consider whether independent validation is enough when both the end user and the independent validator use the same paradigms of degree distribution and stochastic methodology. Inviting experts with alternative paradigms—some of which are discussed here—to contribute to the validation process might prove useful in identifying shortcomings in existing models.

Objectives

The report was divided into the following sections:

- Organizational structure.
- General policy and governance of model validation.
- Independent validation activities and practices.
- Scope of validation.
- Vendor/third-party models.
- Stress testing, sensitivity analysis, and scenario analysis.

Risk management has become more of a key player in policy issues and is playing a stronger role in governance and oversight.

Organizational Structure

More than 82% of respondents reported that their organizations have a unit dedicated solely to the model validation process, independent from the model builders and business units/users. Generally, this independent unit resides within the risk management function. Having an independent unit was the primary characteristic that distinguished a leading practitioner from other institutions.

Clearly, the recent credit crisis and market disruptions have affected institutions' model validation and model risk management practices, forcing some to create specific model risk management groups. Steps taken have included updating models with new data as soon as possible and accelerating validation cycles. Respondents believe the crisis has further emphasized the need to understand and quantify risks outside of traditional limits, including the impact of liquidity on model computations, especially those involving credit derivatives. Overall, institutions are putting a heightened focus on risk management.

In the future, if an institution is not already doing so, it should conduct an annual review of internally developed rating models, as well as more prudently backtest probability of default (PD) calculations along with quarterly monitoring of default rates. For corporates, an even more conservative approach is to validate the pricing models for the eligibility of collateral with no observable market prices, but there needs to be closer contact with front-office risk managers. Some institutions have increased their model validation staff, which in turn has increased its involvement in quantitative regulatory issues.

General Policy and Governance

Understandably, the financial crisis has increased the focus on governance. The survey confirmed that industry best practice (as reported by 84% of respondents) is to have a written, enterprise-wide policy for managing and validating model risk. In fact, respondents said having such a policy is another attribute that distinguished a leading practitioner from other institutions. For those institutions with a written, enterprise-wide policy, it is primarily owned by the chief risk officer or the model risk management/validation committee.

Responses were equally split between those institutions with a model risk management/validation committee and

those without one. Seven of eight leading practitioners have a model risk management/validation committee, while most institutions with fewer than \$100 billion in assets do not. Respondents said establishing such a committee should be a priority, especially since it can be invaluable in cases of disagreement between model builders/users and model validators. When there is clear authority, generally the model risk management/validation committee is authorized to arbitrate disputes. This practice is predominant among leading practitioners and is generally the case across all asset size classes.

According to respondents, risk management has become more of a key player in policy issues and is playing a stronger role in governance and oversight. The crisis has accelerated institutions' efforts to communicate and explain guidelines as well as the institutional and individual group requirements necessary to meet those guidelines and risk thresholds. In most institutions, model control policy, guidelines, and governance have been rewritten, and firms are tightening standards and controls around models. Most institutions are developing clearer governance structures for economic capital models, and many are being forced into reorganization in order to bring control of different valuation streams (models, adjustments, price testing, and trade reviews) under a consistent governance framework.

Independent Validation Activities and Practices

Respondents expressed concern about the nature and extent of specific validation practices and how independent they are within the firm's risk structure. Institutions that plan to independently validate their models generally expect to do so within the next two years. Institutions that plan to revalidate models classified as "high impact" (models that would adversely affect earnings) generally expect to do so within the next two years as well. Leading practitioners generally revalidate in a period of less than two years, with the responses evenly divided between "one year or less" and "one to two years." Some respondents said their institutions do not use model impact classification, but a number of them are developing such a classification.

In Canada, Basel II models are required for annual validation by the Office of the Superintendent of Financial Institutions (OSFI). For institutions that do not revalidate these models and for high-impact models outside Basel II regulation, the best practice is clear: An institution needs to revalidate these models within the next two years and have a policy of revalidation in place.

Most respondents require a model to be independently evaluated or validated prior to implementation, with 75% of leading practitioners and 50–65% across all asset classes doing so. However, institutions that do not require independent model validation have numerous reasons for not doing so, including:

- Every trade on an unapproved model must be approved

by senior management, such as the risk-area committee and enterprise risk management (ERM) committee.

- The head of model validation must approve exposure limits (nominal amount) on top of time limits.
- Necessary steps must be taken to ensure that risk is mitigated if a model is put into production prior to an independent validation.
- Validation is triggered by valuation risk assessment.

When model validation is outsourced, the external consultants usually report their findings to risk management, which is headed by the chief risk officer. Some consultants may report to ERM or the model validation group, including the head of the model validation unit. In addition, the chief financial officer and the chief risk officer also may be heavily involved.

The survey also examined existing models and asked about the need for approvals when new or material modifications must be made. Three out of four respondents require formal approval prior to using new models or making material modifications to existing models. For the nearly 20% of institutions with no clear rule, material modifications to existing models should be formally approved prior to use, according to best practices.

The survey also revealed that model risk management and validation activities have changed in response to recent market conditions. There is increased emphasis on reassessment of model validation policy and formalization of a framework, as well as on assessment of ongoing monitoring performance and monitoring/sensitivity evaluation activities. The Lehman Brothers default, in particular, has highlighted the need for formal model validation processes. In addition, respondents now spend more time looking at model assumptions to consider if they are realistic, taking into account market knowledge to determine a model's appropriateness and whether or not it needs reviewing as a result of market dislocations. There also is a need to revisit previously approved models and stress test them using new scenarios and historical simulations.

Scope of Validation

The credit crisis and ensuing market turmoil have intensified the scrutiny on validations, affecting the range of procedures and standards required. Conceptual soundness (leading practitioners, 100%) and documentation (leading practitioners, 75%) stand out as the two most often included among the top five activities in the scope of independent model validations.

Most respondents agreed that there is no established time limit for how long a deficient model can be used before it is replaced. Institutions with a specific limit generally set it in the zero to six-month time frame. Leading practitioners are evenly divided between "no established limit" and "a deficient model cannot be used." When a deficient model

is used, however, it needs to be subject to reserves and the ability to quantify its deficiencies. Also, models with minor shortcomings can be operated typically for no more than 90 to 180 days. For more pronounced deficiencies, the model issues should be addressed immediately.

The crisis has increased the scope of institutions' validations in both frequency and coverage. For some respondents that actively validate credit models, PD models, internal risk ratings, and loss given default (LGD) models have been extensively reviewed and validated. For market risk models, most credit products, especially structured credit products, are now reviewed intensively for both valuation and risk. Under current accounting rules, credit valuation adjustment (CVA) is required for both counterparty credit and the institutions' own credit, which has significantly affected the scope of model validation.

Respondents said that, going forward, there should be increased discussion and more frequent meetings with model owners and a greater focus on monitoring (including pre-implementation monitoring) to ensure that models remain "fit for purpose." In addition, model validation schedules should be fluid to allow for more timely response in implementing new models.

Of course, model validation typically involves examining source codes, equations, and algorithms designed to produce outputs that adhere to a paradigmatic understanding of degree distributions of risk events and covariance of risk factors. These events and factors themselves are based on a priori assumptions about the dynamics of capital market activity. A more rigorous validation process would involve testing the validity of the paradigm itself and its ability to accurately mirror the actual degree distribution of price time-series and the implications of that distribution for an understanding of market dynamics.

For example, most practitioners understand that the efficient market hypothesis (EMH), as well as the concomitant normal (Gaussian) distribution of market events and the "random walk" nature of those events, is a paradigm that has come under enormous criticism in recent years. In particular, the phenomenon of "fat tails," or extreme events

The credit crisis and ensuing market turmoil have intensified the scrutiny on validations, affecting the range of procedures and standards required.

Significant effort should be spent on developing links between market risks and credit risks as well as outlining extreme stress scenarios, commonly referred to as “black swans.”

that lie outside the normal distribution, has been well vetted, and most models provide some specialized handling of these events in addition to the stochastic methodology applied to normal distributions. Despite these refinements, however, few have rejected EMH and random walk entirely. This is partly due to the utility of EMH in most cases and the lack of well-understood alternatives.

Recently, a number of alternative paradigms have appeared. These paradigms do not rely on EMH but rather assume its opposite and build models that are more robust with empirical evidence on market price patterns. Among these approaches are:

- **Behavioral economics.** This field relies on insights into human behavior derived from social science and psychology—in particular, the “irrational” nature of people’s decision making when faced with economic choices. Insights include risk aversion, herding, the presence or absence of cognitive diversity, and network effects. While not summarized in a general theory and not always amenable to quantitative modeling, the insights of behavioral economics are powerful and should be considered in weighing reliance on EMH-style models, which do not make allowance for subjective influences captured in this approach.
- **Imperfect-knowledge economics.** This discipline, commonly known as IKE, attempts to deal with uncertainty inherent in capital-market risk taking by using a combination of Bayesian networks (Baynets), link analysis, causal inference, and probabilistic hypotheses to fill in unknowns using the known. This method is heavily dependent on the proper construction of paths and the proper weighing of probabilities in each hypothesis cell or evidence cell. Used correctly, it can guide decision making without applying the straitjacket of a normal degree distribution not supported by empirical evidence.
- **Econophysics.** This branch of financial economics uses

insights gained from physics to model capital markets’ behavior. These insights include nonlinearity in dynamic critical-state systems, which exhibit an unpredictably deterministic, nonlinear relationship between inputs and outputs (the “butterfly effect”), and scale invariance that agrees with actual time series of capital market prices. More importantly, this field leads to a distribution characterized by a “power curve” rather than the normal “bell curve,” with implications for scaling metrics in the management of systemic risk.

While all of these approaches and others not mentioned here require more research to normalize metrics and build general theories, they are efficacious and robust alternatives to EMH and random walk analysis. Their consideration in a validation process might lead to both reduced reliance on all existing models and increased emphasis on countercyclical risk management.

Vendor/Third-party Models

Institutions’ attitudes are evenly split when asked about managing risk associated with vendor/third-party models. There does not seem to be a clear consensus with regard to vendors’ models being more or less risky than internally developed models. This is true across all asset sizes and among leading practitioners.

One-third of respondents do not require vendors to provide documentation of their validation methods and results, while two-thirds either do require it or are considering requiring it. Half of leading practitioners require or plan to require vendors to provide documentation of their validation methods and results. The finding that one-third of all respondents don’t require documentation from vendors seems to run contrary to the industry practice moving in that direction. Institutions need to have this information about the models they’re using so they can make correct decisions and feel confident that the models are correctly capturing the risks they are undertaking.

In some instances, the recent crisis has placed significant stress on relations between institutions that rely on third-party vendors (typically the smaller institutions) and their vendors. Some respondents said vendors are often slow to respond to market dynamics and react defensively when pushed to provide documentation on their model validation methods. The lack of information may not have been an issue prior to the crisis, but that has changed in the past eight to 12 months as institutions have required and requested more information. In addition, institutions should perform due diligence on all vendors, especially new ones, and closely evaluate their viability in order to avoid being caught in a dangerous position if the vendor goes bankrupt. For that matter, new vendors or models are always a big concern for an institution, simply because the models’ validity has not been tested long enough.

Stress Testing, Sensitivity Analysis, and Scenario Analysis

Stress testing, sensitivity analysis, and scenario analysis do not necessarily need to be carried out by the staff responsible for model validation and governance. Given the recent Basel II consultative document, *Principles for Sound Stress Testing and Practices and Supervision*, however, it would make sense to be familiar with current practices.

When asked how frequently they run the following stress and scenario testing—single-factor stress, multifactor stress, and historical scenarios—institutions say they are moving from monthly and weekly testing to daily testing. Hypothetical scenarios currently are run monthly, although some institutions are not doing testing at all. In addition, approximately two-thirds of all institutions, including leading practitioners, do not perform stress and scenario testing at all using brute-force Monte Carlo simulation.

The credit crisis has shown that many risks are not just correlated but interconnected. The survey queried institutions on whether their scenario analysis allows for the interaction of market risk with other risks—for example, credit risk or counterparty risk. Although fairly evenly divided, more respondents than not allow for the interaction of market risk with other risks. Scenario analysis allowing for the interaction of market risk with other risks is incorporated more at larger institutions than smaller ones. Data availability and integration of market risks and credit risks are the two main limitations on stress and scenario testing capability.

As with model validation, there should be specific guidelines for scenario generation to ensure that the scenario is properly vetted. Also, the appropriateness and soundness of the assumptions should properly reflect the current economic environment. Surprisingly, less than the majority of institutions have specific guidelines for scenario generation.

Some 70% of leading practitioners distinguish between stress testing and sensitivity testing. Sensitivity testing is useful when the parameter estimation is not reliable and/or some parameter values are indicative of drastic shifts in market factors. On the other hand, scenario analysis is used when considering situations outside of the current market environment or rare events that happened in the past.

In addition to ad hoc testing based on model and risk type, including what is being stressed, the parameter values most frequently tested in sensitivity analysis are short-term changes in interest rates and probability of default. Underlying factors driving stress testing are PD for credit risk models and market instruments (e.g., rates and volatilities) for market risk.

The credit crisis has affected institutions' stress testing, triggering an overall review of stress-testing programs. Significant effort should be spent on developing links between market risks and credit risks as well as outlining extreme stress scenarios, commonly referred to as "black swans." That is, institutions should focus on bigger stress events outside of normal historical levels—for example, taking 10-

20-, and even 50-year event types into account much more than in the past. There also needs to be greater emphasis on integrating market and credit risks and quantifying the impact on CVA as well as mark to market.

The crisis has also led institutions to perform a complete review of scenarios and to add more recent historical data, now that there is the "new" 1987 stress test. As some participants remarked, hypothetical scenarios have become historical scenarios. Going forward, institutions should include stress tests that affect earnings, capital, and liquidity. A firm's amount of liquidity and how that liquidity in the market (i.e., credit, equities, and fixed income) will affect individual institutions must be stress tested. An institution should also consider including secondary factors such as bond basis and cross-currency basis in its stress-testing set of scenarios.

More importantly, the insights from econophysics (discussed above) suggest that the extent of the worst possible catastrophe in a stress test is not a function of history or a linear extrapolation of known events, but rather an exponential function of systemic scale. This implies that as the system itself is scaled upward—for example, in the growth of the notional value of derivatives—the "worst case" is neither mitigated by hedging nor related to such an increase in a linear way, but rather grows as a 100, 1,000, or 10,000 times function of notional value. This makes a powerful case for "descaling" techniques such as clearinghouses, which provide netting, and for canceling old trades rather than sterilizing them with "pair-offs."

Conclusion

The survey succeeded in creating an industry benchmark of best practices with a varying degree of emphasis across different firms. The results provide a best practice in model validation with regard to organizational structure, general policy and governance of model validation, independent validation activities and practices, scope of validation, and vendor/third-party models.

In addition, given the recent market turmoil and the Basel II Accord, the models that financial institutions rely on need to be scrutinized more carefully and back-tested more appropriately to validate the underlying assumptions. This will allow institutions to develop a more prudent approach to generating scenarios and running simulations, resulting in a better stress-testing practice.

The survey also proves that more education is needed about what model validation is and why it is important. It reveals a glimpse of where the industry stands on the issue and allows individual firms to assess their positions relative to their peers. ❖



Copies of the full survey are available for purchase on www.rmahq.org. Click through to the Market Risk section.