

Keys to Success in Model Risk Management for CCAR & DFAST

By Sisi Liang and Joseph L. Breeden

Introduction

Among the many causes of the US Mortgage Crisis, should we add to the list that the crisis occurred because banks did not have forecasting models? Not exactly. All lenders had loss forecasting models of some sort. Presumably, effective loss forecasting should have resulted in lenders withholding loss reserves and capital sufficient to lessen the crisis.

Federal regulators appear to have reached similar conclusions, as two sets of guidelines have been evolving since. Until recent years, stress testing (forecasting under economic scenarios) had only been performed half-heartedly. But this has now become a central modeling activity for top lenders in the form of CCAR (Comprehensive Capital Analysis and Review¹) and DFAST (Dodd-Frank Act Stress Tests²).

On a parallel track, the Model Risk Management guidelines^{3,4,5,6} have sought to assure that there is a model governance program over the population of relevant models and any covered model is fit for use at inception and continues to be fit for use. Models that fail to capture serious risks may lead to misguided decisions by management, over-exposure, and dramatic financial losses.

The mortgage crisis of 2008, for example, would have been much less severe if the models accurately captured the risks being assumed. Timely and effective management of model risk, therefore, is key to preventing the crisis from repeating itself. But this may be easier said than done. Numerous challenges are faced by every relevant group in the banks to execute this successfully. The regulatory environment is continuously evolving, and therefore the standard of “best practice” is a moving target. New classes of models, which are the stress test models, are required for most lenders, who are struggling to obtain the talent and resources to keep up with model development.

In addition, new requirements are coming into play for both model validation and internal audit. In fact, there are specific regulatory expectations for those institutions that apply the “advanced approaches,” for internal audit to review the capital management program and provide its assessment to the Board of Directors¹.

¹ Board of Governors of the Federal Reserve System. Basel Coordination Committee Bulletin 13-6 Guidance for Internal Audit under the Advanced Approaches Rule. May, 2013.

“Under section 22(j)(5) of the Federal Reserve Board's advanced approaches risk-based capital rule (rule), a bank or a bank holding company must have an internal audit function independent of business-line management

CCAR and DFAST rules are maturing; therefore Model Risk Management must evolve accordingly. At the core of all of these challenges, is the need to enable every relevant part of the bank to work together effectively and competently in addressing the regulatory requirements while meeting the business needs.

This may be the most fundamental challenge and the least discussed, and yet without solving this issue, no amount of regulations, enforcement, or excessively generous budget will yield the result we truly want – better models, better practice, and a healthier lending environment. One reason for this challenge to be the least discussed is that it is among the most subtle and difficult problems to uncover and tackle. It takes brutal honesty, commitment and well-grounded minds to make an impact in solving this problem.

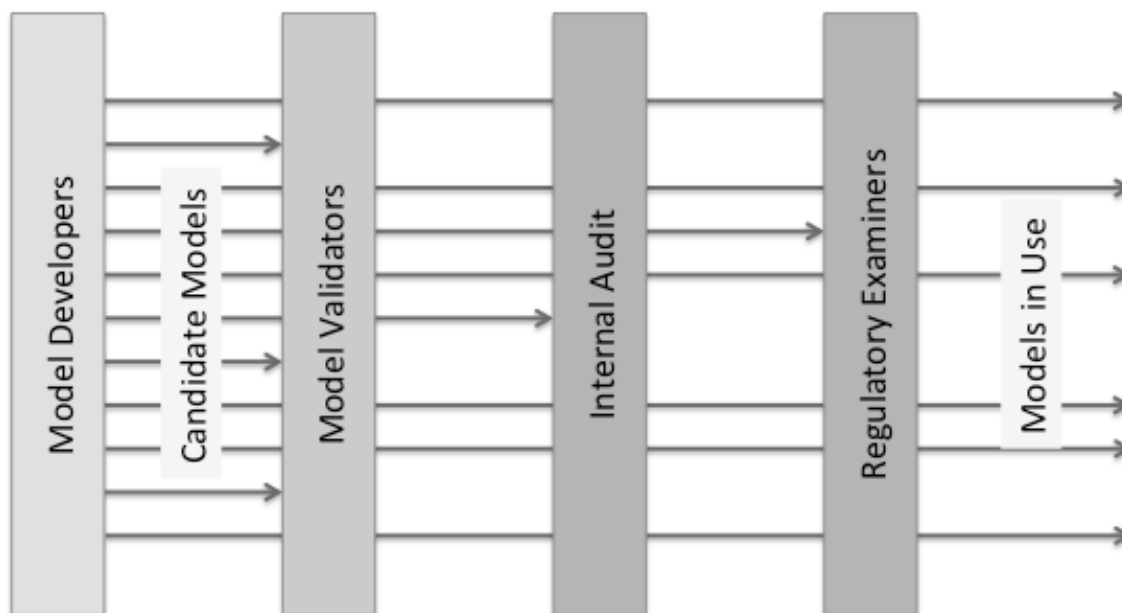
This article focuses on the best practices in Model Risk Management for CCAR and DFAST resulting from interviews with multiple Top 20 lenders. The observations and recommendations presented here are intended to capture the challenges faced by institutions as they strive to comply with the regulators' requirements, while implementing Model Risk Management in a way that makes practical sense to their own unique institutions.

Key Players in Model Risk Management

The goal of Model Risk Management is to prevent bad models from making it into production and to monitor good models for indications of deterioration. Good business decisions cannot derive from bad models. To prevent this, lenders use multiple lines of defense and levels of governance to filter out the bad models: Model Developers, Model Validators, Internal Audit, and often a Model Risk Committee. Regulatory Examiners form the last line of defense, reviewing the work of all previous lines.

The graph below conceptualizes the filtering process whereby not all models survive all the levels of review. Most model rejections will occur with model validators, but on rare occasions audit or examiners may reject a model.

that at least annually assesses the effectiveness of the controls supporting the [bank]'s advanced systems and reports its findings to the [bank]'s board of directors (or a committee thereof)."



The model developers create, validate, and document the candidate models. In some situations, they are also involved when the model goes into production to monitor performance.

Model validation has a multi-faceted set of responsibilities. These responsibilities vary slightly from one institution to another, but along the OCC guidelines, they generally include the evaluation of conceptual soundness, on-going monitoring of model implementation and performance, and outcome analysis in comparison to the actual outcomes. The biggest challenge for the model validation seems to be deciding how much is enough: how many tests are sufficient to prove effectiveness? What level of in-sample or back-test accuracy is acceptable?

A model validation's list usually includes the following questions:

Data: Is it segmented best for the model chosen and in compliance with examiner suggestions?

Conceptual Soundness: Does the model capture all key dynamics and work with the available length and breadth of data?

Estimation Method: Is it robust and unbiased?

Overrides: Are these used sparingly and well-explained with business justifications?

Limitations and Weaknesses: Are the limitations and weaknesses understood and well documented? The worst thing to present to examiners is not a bad model, but a bad model promoted as a good one.

Performance: Does the model pass stability tests, out-of-time backtests, and sensitivity tests?

Source Code Control: Are code access and change controls in place?

Monitoring: Are model estimation stability and performance against actuals monitored regularly?

Internal audit as the third line of defense plays a key role in assessing the effectiveness of the bank's model risk management framework. In recent years, the regulators have been criticizing the effectiveness of internal audit, which pushes the auditors to provide more rigorous challenges to both the models themselves and the model validators' work. Unfortunately, there is ambiguity as to exactly how or in what specific areas should internal audit be involved, and what kind of talent would be required to perform this task. Particularly, this hybrid group of auditors need to both be able to understand (enough of) the technical aspects of model development, and perform an effective audit at the same time. Experiences show that the people performing this special audit tend to lean heavily on one side or the other. It has been a challenge to find the sweet spot in terms of what to audit, and it is an even bigger challenge to find the right people.

In addition to the above groups, banks usually set up a "Model Risk Committee" to approve the use of models. The people on the Committee are senior level personnel drawn from various groups of the bank such as treasury and credit. This mix of professionals brings a healthy amount of debate but banking organizations should be sensitive to the balance between model users and those in control functions, particularly when disagreements arise. Conflicts of interest may arise when approving or rejecting a model. Therefore appropriate escalation and resolution processes should be in place with particular weight given to the control functions over the lines of business. Additionally, there tends to be a lack of expertise in the members of the Model Risk Committee to rigorously opine on the many issues with respect to a model.

Separately, a group called Model Risk Assessment will assign a level of importance to each model, effectively identifying those portfolios requiring detailed treatment. In addition, this or another group will need to assess the final "Risk Not in Model" to be included as an additional capital buffer. We suggest that Risks Not in Model be owned by model validation, while internal audit should be responsible for the assessment of governance, control and sustainability of this procedure.

Risk Not in Model is not the same as management overlays, which are management adjustments to the model output. Model approvals by the Model Risk Committee may include a requirement for overlays to address acknowledged weaknesses. Overlays are a standard part of the process, but should be used sparingly. An excessive number of overlays in one model suggest a failed model. Large numbers of overlays across a set of models suggest a failed development process. Overlay amounts are created by the model developers or businesses, where the risks should be best known. The overlay estimates should then follow the same model validation review and audit process review as the original model.

Lastly, banks often assemble a separate committee to interpret the regulators' feedback, which includes the MRIA's (Matters Requiring Immediate Attention) and MRA's (Matters Requiring Attention). One significant risk is that this committee may not be sufficiently technically savvy to translate regulatory feedback into specific model-related actions. The risk, too often seen in practice, is that incoherent, incomplete, or misinterpreted action plans may be carried into the next submission cycle. Conversely, a committee providing these interpretations cannot be comprised purely of quants, since business issues must factor in as well.

Model Risk Management on a Deadline

What all participants realize is that fixed deadlines from the regulators make CCAR unique. The natural cycle of model development and enhancement is a continual, on-going process. Injecting deadlines in the midst of this process creates artificial milestones and break points. This, if not managed with strategy and discipline, may encourage the banks to forego long-term benefits in exchange for short-term gains.

An effective approach involves an extensive, iterative feedback loop among the model developers, model validation and internal audit to ensure that the right issues are taken care of at the right time. The latter two groups cannot wait until the CCAR deadline to provide feedback or fail a model. A single end-point review does not support developing the best possible models or satisfying the regulators. Instead, model validation and internal audit should provide constructive feedback early in the process (within reason) so that the model developers would have time to make improvements.

In particular, model validation has the license to be more proactive in the process, and if they are disciplined in giving model developers enough time to make changes, model validation is in the position to make the biggest impact. In the same way, audit would want to provide feedback earlier in the process, at multiple touchpoints. In current practice across the banks, there are two major categories of feedback that audit provides to the developers and the validators. These include: true technical model validation, and assessment on governance, sustainability, control and data. The first aspect, technical model validation, has been new to audit in recent years, after the regulators have criticized the lack of rigor of internal audit at banks. As a result, the audit departments stretched their capability, and moved from the traditional checklist review function into one that somewhat mirrors model validation to the best of their ability, separate from the validators. Adding this second line of (partial) model validation highlights the importance of high-performing models in our current banking and regulatory environment. In addition, the second facet of responsibility in audit is increasingly crucial to a bank's success. Some banks do not yet have an effective governance framework in place, and struggle to build one when confronted by failing stress-test results. Internal audit plays a primary role in ensuring the framework's sustainability, effective control

and good data quality. In the same way as model validation, audit should provide feedback and assessment on these aspects to both the developers and the validators relatively early in the process to ensure an active feedback loop and effective challenging.

The dynamic among the model developers, model validation, internal audit and Model Risk Committee reveals how well (or badly) the model risk management framework is functioning within a bank. These groups are likely to need on-going, lengthy interactions with each other in order to get the best possible models, but this may lead to unwanted hand-holding over time. One group of validators explained that when they first instituted an iterative validation process, it degenerated into daily phone calls from developers on minor details to determine what the validators wanted to see in the model. Such interactions blur the distinction between developer and validator.

We also find wide variation in the communication of findings between validator and developer. In some institutions, validators will only point out those items that are insufficient. In other organizations, validators have provided detailed recommendations. In either case, the validators need to be careful that their criticisms or concerns reflect clear model weaknesses and not expressions of personal preference. The presumption should be that the developer is the expert in creating the model (which we generally see to be the case) and the validator is present to check their work, rather than override the work or become the de facto model designer.

A more disciplined approach to the developer – validator interactions would be to set healthy boundaries among these groups with a focus on identified weaknesses, allowing the developers to lead the design of the solutions. Model validation, for example, can make a list of items that they would agree to answer for the model developers at predetermined check-points. These items may include: data quality, model design, and model estimation. The validators can and should refrain from providing advice that is out of its scope.

Similar questions arise in the interactions between internal audit and developers or validators. Although a wait-to-the-end approach is insufficient with CCAR, continuous interactions can lead to a breakdown in process where there are no fixed interactions or deadlines. Best practice appears to be having a fixed number of reviews by audit. Interim audit feedback can facilitate quick remediation, but the final review will simply be attempting to document all known shortcomings rather than to expect all exceptions to be remediated before the deadline.

Since the technical teams at internal audit are often newly created, banks need clear guidelines on the role, scope, and timing of the interactions. In practice, internal audit is less likely to have the same technical expertise as the developers or validators. Thus internal audit, as suggested by the regulators, should not bear the responsibility of providing full validation of the models. Rather, internal audit

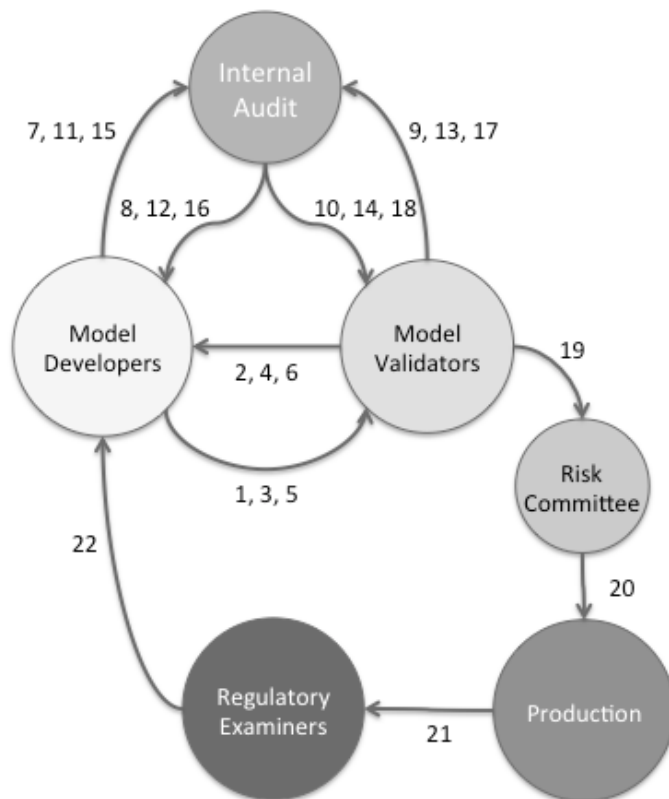
should focus most of the energy and resources on assessing the aspects of sustainability, control, data, and governance in the model risk management framework.

When reviewing model development, audit will follow the same checklist given above for model validation, but with an eye more to the appropriateness of the processes and decisions made, rather than repeating the validation. When reviewing the work of model validation, a corresponding checklist is appropriate:

1. Have the validators established clear guidelines for validating the models?
2. Have the validators reviewed the models for all of the items in their checklist?
3. Are the requested validations sufficient and effective?
4. Have the validators confirmed that the models satisfy all MRAs and MRAs?
5. Have the validators followed its own criteria for accepting or rejecting the models?

The Life of a Model

The interactions of the various groups can be summarized by tracing the life of a model through the process. The following diagram visualizes this flow.



Model developers and model validation should have a predetermined number of touch points, such as Data Review (Steps 1 & 2), Conceptual Soundness Review (Steps 3 & 4), and Performance Review (Steps 5 & 6). Internal audit should also have a predetermined number of iterations of review with the developers (Steps 7, 8, 11, 12, 15, & 16) and with the validators (9, 10, 13, 14, 17, & 18).

Once developers and validators are ready, the model goes to the model risk committee for review (Step 19). Upon approval, it goes to Production (Step 20), and finally to review by examiners from regulatory agencies (Step 21).

After models get into production, it returns to developers for performance monitoring (Step 22) in preparation for the next round of revision or redevelopment.

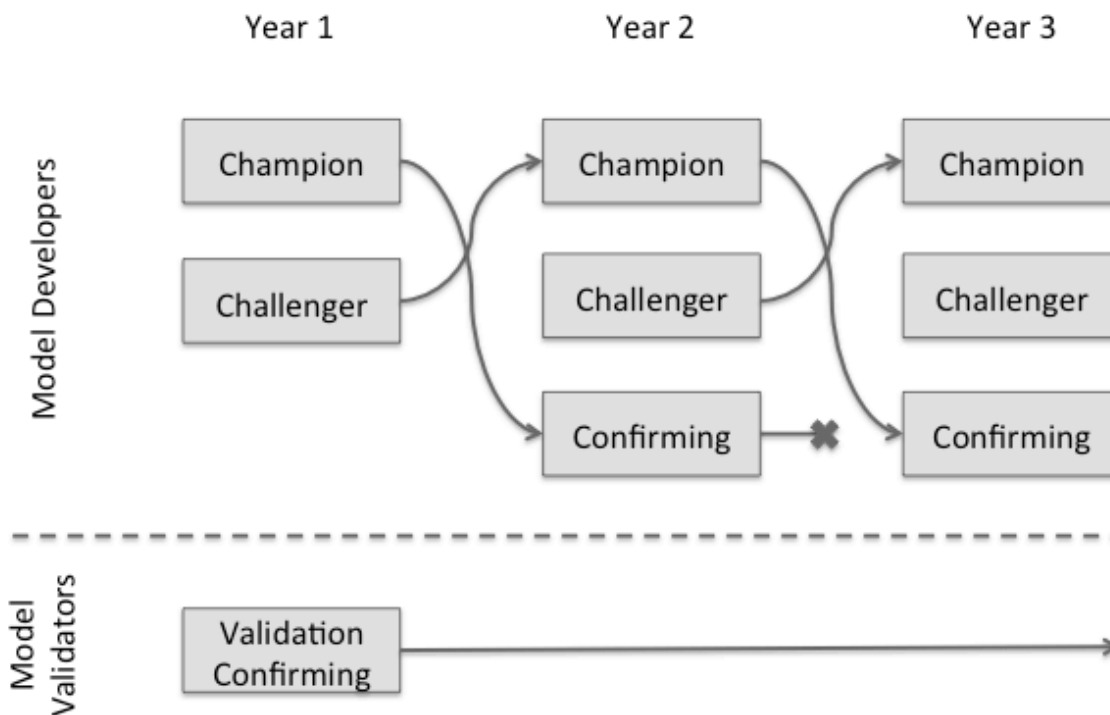
The process of deploying a new model has become much more arduous than in the past, but the new rigor is commensurate with the importance of the result. The challenge is not to lose sight of the importance of efficiency. A key part of the reviews should be challenging developers, validators, and auditors not to waste time on tests or reviews that provide few insights. More testing is not better testing.

A Plethora of Models

When CCAR was first introduced, long-range stress testing models represented a new class of models for most institutions. For the first few years, creating a single set of models was challenging enough. Part of that challenge is how to know if the models are effective. Unlike scoring models, we do not possess hold out samples sufficient to prove the effectiveness of stress test models. If we could build the models on a few recessions and test on one recession, we could be convinced of the effectiveness of the model. However, the typical 12- or 27-month tests capture only gross errors in the models.

The same situation occurred with Basel II. The solution to being unable to prove the effectiveness of the Basel II formula was to build a secondary stress test model. If a second model of a different type provides a similar answer, then we have greater confidence in the initial result. Examiners appear to be suggesting a similar process for CCAR and DFAST.

Using nomenclature common in scoring, the institution's best stress test model is called the Champion, but a second Challenger model should always be under development with the goal of supplanting the Champion at some point. The Challenger must not be last year's failed model, but rather the best attempt at creating something better. The best practice is for Champion and Challenger models to be of different types so that structural weaknesses in one may be avoided in the other.



Along with Champion and Challenger models, top institutions report that examiners have asked for Confirming models to be maintained. The Confirming model should be a fallback model in case the Champion fails to pass validation and the Challenger is not yet ready. Logically the Confirming model may be simpler than the Champion, so it may not be fully compliant with the latest requests from examiners, but it should be easy to maintain and provide a baseline answer.

In addition to Champion, Challenger, and Confirming, we have also heard of model validation being asked to develop Confirming models. The validators' Confirming model provides another point of comparison when assessing the Champion model's effectiveness. In some cases, the validators' Confirming model has been seen to be better than the original Champion model, but the separation of roles mandates that this Confirming model never be used as part of the final CCAR submission.

Whether an institution needs 2, 3, or 4 models depends strongly on its size and systemic importance, but having a benchmark model (Challenger) is no longer discretionary for any sized institution.

Grading the Models—“All Models Are Wrong..”

The process of grading the models is another minefield. Following the traditional audit framework, only the binary grades, pass and fail, are given to the models. But this has proven to be insufficient. The aphorism famously stated by George Box, that

“all models are wrong, but some are useful,” reflects the complexity in assessing a model. There are a couple implications to this statement. First, this shows that grading a model is not a black-and-white, binary task. If we accept that all models are wrong and no model can provide the truth, then we should focus on assessing how close a model is to the truth, and how much useful information it can provide us. This means a letter grade system would be more suitable than a pass/fail system. A letter grade system would also be able to provide more insights and valuable feedback to the model developers. The second implication from Box’s statement is that models should be assessed for their usefulness to a specific purpose.

The attempt to fit a pass/fail structure to model review can lead to unintended results. Validation needs to avoid the temptation to pass a model just because no better model was available. At the opposite extreme, audit cannot take the position that any exception means that the model must fail. Clearly, we must accept that all models will have exceptions with varying severity. We must avoid giving a simple “Pass” grade to a model with many near-fatal exceptions, because developers can reach the erroneous conclusion that the model is acceptable. Likewise, we cannot be so harsh that a model that produces sufficiently accurate results is failed on purely on technicalities.

This brings us back to the issue of the roles of model validation and internal audit in grading the models. The role of model validation is rather straightforward—it is responsible for all full validations of models, with rigor in quantitative techniques. The role of internal audit, on the other hand, is trickier. We believe that audit should manage checklists of the development and validation process rather than revalidating the models. Audit’s quants should certify the model development and validation processes. If the model appears insufficient to audit, audit might choose to fail the process that created or approved the model rather than the model itself.

However, banks struggle to interpret the statement on page 19 of the SR11-7 Attachment from the Federal Reserve:

“Internal audit's role is not to duplicate model risk management activities. Instead, its role is to evaluate whether model risk management is comprehensive, rigorous, and effective. To accomplish this evaluation, internal audit staff should possess sufficient expertise in relevant modeling concepts as well as their use in particular business lines. If some internal audit staff perform certain validation activities, then they should not be involved in the assessment of the overall model risk management framework.”

Therefore, the Fed clearly states that internal audit should not duplicate the work performed by other parts of the model risk management framework, which includes model validation. However, the Fed does require internal audit to provide assessment on the rigor and effectiveness of MRM, and acknowledges that expertise is required on modeling concepts and their use in business lines. This has caused

confusion among the banks regarding how much quantitative assessment should be performed by audit. In order to truly “evaluate whether model risk management is comprehensive, rigorous, and effective,” audit does need to be far more involved with quantitative methods than its role historically entailed. This has inevitably led audit departments hire more quants, especially if regulators have criticized the lack of rigor in audit. But these quants in the audit department are, after all, quants (if audit can find them, that is)—quants who are drawn to validating the models, if not building them.

The core question is what level of review these quants should be performing. Audit could follow a checklist approach, simply confirming that developers and validators have faithfully followed all necessary steps in their respective processes. The opposite extreme is for audit’s reviewers to delve into the details of the application of the chosen modeling approach to the available data, and all of the various applicable validations. Banks would benefit for clearer guidelines from regulators. Or are the regulators experimenting with a new hybrid approach to audit, watching how banks respond to the requirements to develop best practices?

Regulators, the 4th Line

Banks have begun to appreciate the increased sophistication of the regulators. Tom Dahlin of Citizens commented: “Over the past 3 years from the governance perspective, the richest conversations have been ones with the regulators—those conversations have been the most engaging, the most thorough, and covered the most ground in a short period of time, which has been my one big surprise that I’ve encountered here. They are not the regulators that I was familiar with 10 years ago. They are much sharper and coordinating much better than before. They are doing a better job than I’ve ever seen. The questions and responses we are getting back from them... I think they are right on the money. The questions are very pointed, very thoughtful, and they are asking the right questions.”

This change in examiner interactions has created a greater partnership with regulators as a fourth line of defense for the banks. Leading organizations treat the examiners as partners in the process, discussing strengths, weaknesses, and plans for improvement in constructive discussions. One lead model developer commented, “[Examiners] want to know that you are taking this seriously and not keeping secrets. We have learned how to engage the examiners. Our interactions with them are really healthy.”

These are encouraging improvements from both sides, but there are further improvements among the regulators and the banks that could significantly drive the practice of model risk management forward. For example, even though it seems that the regulators provide suggestions on the overall structure of model risk management within banks, they have not given clear, prescriptive guidelines that the banks can adhere to. In part, we expect that regulators are searching for the best

answers too, but a more public discussion of best practice in model risk management, similar to what has been provided for model development¹ would be a boost to the industry.

Conclusion

In recent years, banks have begun to realize that they need dedicated teams to perform all the work required for CCAR and DFAST. During this time we have seen migration from sometimes rudimentary first-try models to more effective approaches, the validation process has become much more rigorous, and businesses are looking to leverage all this development in day-to-day portfolio and account management.

As these teams strive toward best practice, we have gathered the following insights that we believe are key to better practice in model risk management in the context of CCAR and DFAST:

1. Models need grades rather than pass/fail.
2. Pre-determine the failure criteria.
3. Don't wait until the end. Validate and audit as you go. Give feedback early enough for remediation. Make sure that the models address the regulators' MRIA's (Matters Requiring Immediate Attention) and MRA's (Matters Requiring Attention).
4. Hire the right staff. Banks need quants for both validation and audit.

Banks have since spent enormous sums on these projects in order to pass the stress tests with the regulators. If the banks can successfully implement the above functions within the Model Risk Management framework, they will become much more cost-effective, and with better results. At any rate, however, we should always be asking ourselves this question: Is what's being accomplished worth the expense? How much is too much? These concerns are best satisfied through effective planning, definition of roles and criteria, and coordination across teams. A more disciplined, smarter approach can dramatically reduce the time, money and heartaches for banks.

Additionally, we have seen the regulators gradually moving the goal posts on what is sufficient for CCAR models. Unlike Basel II, where clear model criteria were provided, CCAR appears to be a process of continual enhancement. With CCAR, "good enough" may be many years in the future, from the most to the least sophisticated lenders. In part, this is due to the lack of multiple economic cycles in the data. Both banks and regulators still have much to learn about how to build and deploy effective stress test models.

Clearly, the regulators have concluded that through the mortgage crisis we have insufficient models and model risk management. Something more was necessary. Time will tell if we have reached the correct level.

References

1. Board of Governors of the Federal Reserve System, 2013, "Capital Planning at Large Banking Holding Companies: Supervisory Expectations and Range of Current Practice," August.
2. Board of Governors of the Federal Reserve System, 2014, "Supervisory Guidance on Dodd-Frank Act Company-Run Stress Testing for Banking Organizations with Total Consolidated Assets of More Than \$10 Billion but Less Than \$50 Billion (SR14-3)," March 6.
3. Board of Governors of the Federal Reserve System, 2014, "Supervisory Guidance for Data, Modeling, and Model Risk Management Under the Operational Risk Advanced Measurement Approaches," June 30.
4. Office of the Comptroller of the Currency, U.S. Department of the Treasury, 2011, "Supervisory Guidance on Model Risk Management," April 4.
5. Board of Governors of the Federal Reserve System, Office of the Comptroller of the Currency, 2011, "Supervisory Guidance on Model Risk Management (SR11-7 Attachment)," April 4.
6. Federal Deposit Insurance Corporation, 2013, "Stress Testing and Model Governance," August 6.

To appear in RMA Journal 2015
Copyright © 2013 Prescient Models LLC. All rights reserved.

