Validating Internal Rating Systems

Validation of rating systems should be more than just annual testing to help banks comply with regulations. A continually improving rating system can reduce credit losses and sustain long-term profitability and competitive advantage.

BY BOGIE OZDEMIR

More banks are realizing the importance and urgency of improving and validating their internal credit rating systems. Some banks must do this to gain Basel II compliance, while many others want to do it to keep their compliance options open and to leverage best-practice risk management in making improved business decisions.

Improving business decisions is an important goal for validation programs because internal rating systems have become a core information system, feeding risk data into the MIS and economic-capital-based systems that banks use to make strategic and tactical decisions, including performance measurement, risk-based pricing, limits setting, and incentive compensation.

Validation programs are hard to get right, however, and involve a lot more than running one or two technical tests, such as a discriminatory power test. Statistical procedures are not enough to gain Basel II compliance. And they certainly are not enough to gain the risk management and
business benefits of an improved rating system—including better credit surveillance and a near-term reduction in credit losses.

Over the past five years, our effort to help banks build validation strategies has helped us understand the most common validation pitfalls and how banks can avoid them to turn their rating system investments to competitive advantage.

Surprisingly, the technical and data issues are rarely the real hurdle. Instead, the hurdle is much more like a classic management conundrum—setting the right goals, understanding the issues, putting the right people in place, and making sure different pieces of the organization work together to a common end.

Senior managers do not need a Ph.D. in mathematics to set their organizations on track to a sound validation and ratings improvement strategy, but they do need a solid game plan.

**Set the right goals. Do not aim only for compliance with Basel II, but shoot for business benefits.**

Basel’s Pillar 2 has prompted banks to improve their internal rating systems, but it also set many banks on the wrong track. Improving rating systems is seen as a cost of doing business, a compliance hurdle, rather than a source of potential business benefits.

Yet improving a bank’s risk rating system offers immediate business gains. A recent study suggests that moderately improving a rating system to better screen new midmarket loans can generate a $5.6 million gain for a medium-sized U.S. bank.¹ This number relates only to the improved loan screening during origination. It does not include the significant gains the bank would make from improved risk-based pricing, better capital allocation, and other competitive benefits.

Critically, gaining competitive benefits from the bank’s rating system depends not only on improved ratings accuracy, but also on the confidence that the bank and its senior management have in their rating system when making important decisions. In the aftermath of the current credit crisis, which is causing many banks to lose confidence in their lending practices, renewed trust in their systems may help management regain some of that lost willingness to lend.

Even those banks determinedly focused on compliance now need to move on from Pillar 1 and pay attention to Pillars 2 and 3. The Pillar 2 supervisory review makes many demands concerning credit stress testing and the quality of the bank’s capital adequacy assessments. It also obliges banks to forge a link between the ratings used to calculate regulatory capital and the credit risk information that management uses to make business decisions.

Pillar 3 compliance will concentrate minds even more.

---

**Outcomes Analysis**

**What Are the Key Technical Backtesting and Benchmarking Tests?**

**Backtesting**

1. **Discriminatory power of rating system.** These tests assess how good the bank’s system is at distinguishing default and recovery risk—in other words, whether defaults come from risk ratings with high estimated probabilities of default (PD). Do higher realized losses after a default correspond to high loss given default (LGD) estimates?

2. **Calibration test.** This test asks whether the PD and LGD levels assigned to a rating are consistent with the actual default and recovery experience. Are any apparent discrepancies statistically meaningful?

3. **Long-run PDs.** Do the PDs assigned to risk ratings reflect the bank’s long-run default experience? The challenge here is that the internal default rate time series, from which the PDs are estimated, may not cover even a single full economic cycle. How can the bank use internal and external data jointly and robustly to mend this problem?

4. **Downturn LGDs.** How can the bank adjust its LGD estimates to account for the interdependencies between default rates and recoveries? The correlation between PDs and LGDs is not taken into account when estimating Basel’s credit capital (or in many commercially available economic capital models). This results in an underestimation of the bank’s capital requirement. Using downturn LGD is one way to compensate for this underestimation.

5. **Realization of risk-rating philosophies.** The bank needs to validate whether it is actually following its stated risk-rating philosophy, whether this is “point in time” or “through the cycle.” An unnoticed divergence from the intended philosophy will have serious implications for how the bank interprets and manages its risk and capital.

6. **Risk homogeneity of obligors in each risk rating.** How does the bank know, for example, that a “risk rating 5” in one region or sector is the same as a “risk rating 5” in a different region or sector?
Banks will have to publish information drawn from their internal rating systems for the benefit of investors and bank analysts. That information had better be right because it will drive the bank’s share price up or down. Analysts also will monitor how trustworthy the bank’s credit information turns out to be over time, making an accurate and robust rating system a business imperative.

In truth, this is exactly what the regulators are hoping to see—banks translating Pillar 2 implementation into better risk management and business benefits, rather than writing Basel II down as a regulatory tax.

Take time to understand what validation really is.

It’s tempting to think that rating-system validation consists largely of a handful of technical tests. This is simply not true, not for regulators and not for banks keen to secure potential business benefits.

In fact, regulators use the term mosaic of evidence to refer to the many different dimensions of internal-rating-system validation. Much of that mosaic is not really technical testing at all. The mosaic includes the conceptual soundness of the rating models, how the rating system is operated day to day, and a series of very technical tests that must be interpreted within this wider context.

Figure 1 illustrates how all the components of validation work together. Take the example of a rating system that produces inconsistent probability-of-default estimates for similar borrowers because there is a problem with the rating system’s conceptual soundness. The specific reason for this might be that some of the ratings are based on a shorter time horizon than others (that is, they are more point-in-time) or there might be some systematic biases for obligors in certain industries. If this rating system allows lenders to override the system without sufficient limits or monitoring, it also will suffer from an operational weakness.

It’s likely that any outcomes analysis tests (see the box, Outcomes Analysis) run by this bank will show the rating system is not distinguishing effectively among borrowers or performing as well as external benchmarks. To respond to these technical test results and mend the system, managers will first have to grasp how the test results relate to potential conceptual and operational weaknesses.

This explains why validation is more than simply backtesting, and also why senior managers are a vital part of the validation process. They can encourage staff to stand back and take a critical view of how rating models are used in a business and operational context.

Building internal validation capabilities is an ongoing part of a bank’s core competence.

It is tempting for banks that regard validation procedures as a set of specialist technologies, applied once per year, to outsource rather than create their own validation skills. This is a mistake. Validation is a procedure that goes to the heart of the bank’s core competence: differentiating between borrowers.
Regulators insist that validation is an ongoing activity. So if banks do not run the show themselves, they must bring in expensive external consultants annually to run their black boxes. Of course, it can make sense to work with industry experts who already have the tools, data, and expertise needed to power up a ratings validation program. But banks should demand that these external experts help their staff build a validation factory that the bank can run itself.

Banks should require that external experts arrive with an open-book mentality, including offering open source code for validation models and tests, training bank staff in how to validate ratings, and showing how the results of tests can be built into a program of continuous rating system improvement.

Validation is not a backroom issue. Teams must challenge and improve organizational practices.

Validation is almost useless if it is limited to technical tests performed by “quants” with results that stay largely within the validation group, mainly because rating systems often go wrong for organizational and business reasons. To understand what is broken and how to put it right, banks have to build robust connections between validation personnel and other bank functions.

For example, validation tests often reveal that the actual default rate associated with the lowest origination (pass) rating grade in a portfolio is much higher than the probability-of-default estimate associated with that rating grade by the bank’s rating system. The reason lies with business pressures. Lenders are motivated to award pass grades and originate new loans even when the loans are higher risk than permitted by the bank’s official policies. To address this issue, the bank may need to reexamine its policies within a risk-return framework to determine whether the bank has established an economically rational cutoff point in terms of the riskiness of the lowest passing grade.

Questioning a bank’s rating system often provokes internal controversy. The Basel II regulators require that validation groups be able to “effectively challenge” the designers of the risk rating systems. That means validators must be given a minimum level of power and independence, for example, in terms of reporting lines.

It also is important to build the right mix of people into the validation group. Quantitative talent plays a role, but the group must also contain personnel from the front line of the credit group and from the business lines.

Lastly, validation groups should not simply be delivering test reports. They need to be talking to other bank functions and communicating with senior management in the language of the banking business. Otherwise, the team, in partnership with senior management, cannot persuade the wider bank to alter its practices to improve the accuracy and consistency of ratings.
continued from page 81

If default rates differ, it might be a sign that the risk ratings are not risk homogeneous. The bank needs to have a strategy for testing, quantifying, and mending this kind of problem.

**Benchmarking**

7. **Benchmarking performing loans.** These tests benchmark internal ratings against external ratings for the same borrower, where these are available, and against third-party models that produce PDs, LGDs, and implied ratings for unrated entities. The agreement between internal and external ratings should be monitored over time to identify systematic trends and changes in trends.

8. **Benchmarking defaulted loans.** Are the bank’s realized default rates, migration rates, and LGDs consistent with relevant industry experience?

9. **Benchmarking the backtesting performance.** It can be a big challenge to make sense of backtesting results. They are not meaningful in isolation. Banks need to benchmark the results to determine the performance of their rating systems on a relative scale. To make an apples-to-apples comparison, however, the bank must first adjust the results to account for differences in portfolio composition, risk-rating philosophy, and so on.

Validation is not a once-a-year activity. It is part of a continued improvement program.

Regulators require that certain validation activities be performed at least once a year (for example, outcomes analysis). With the right kind of validation factory in place, however, it is quick, easy, and cost effective for banks to use validation tools to continually monitor and improve their rating systems through the year.

This is especially true of the key performance indicators (KPIs) that regulators encourage banks to put in place. KPIs should cover many aspects of the credit rating system and credit risk reporting. For example, one class of KPI might track the tendency for the bank’s lending officers to override the bank’s rating system by lender and industry segment. That way, management can see where and why overrides are happening across the bank’s key portfolios.

Other KPIs might track the degree of dissonance between the bank’s internal ratings and benchmark or external agency ratings (or the ratings computed on a sample of the portfolio using a third-party model). For example, as macroeconomic conditions deteriorate in the current climate, can the bank spot and justify a gap that opens up between the bank’s internal ratings and those of the rating agencies? Figure 2 shows what shape this gap might take in various situations.

Management also may want to set up a KPI that compares realized default rates and realized loss-given-default rates for recently defaulted loans with the default and loss rates predicted by the bank’s rating system or with historical averages.

Increasingly, leading banks find it useful to monitor KPIs up to once a month and to run a large subset of validation procedures up to four times a year. For these banks, an initial ambition to build an annual Basel II rating validation procedure with business benefits has resulted in a continuous improvement program for their rating system.

**Conclusion**

Banks need to shed the idea that validation is simply a series of technical tests to help them comply with regulations. That idea will lead them to treat validation as a yearly expense rather than as a system of continual improvement for one of their core competitive competencies.

Too limited an approach may not even deliver compliance, given the regulators’ insistence that banks examine a wider mosaic of evidence, including rating concepts and operations. Furthermore, the limited approach will look increasingly irrelevant as the regulators put more emphasis on compliance with Pillars 2 and 3.

A more ambitious program means validation cannot be confined to the backroom but must reach out into the bank. Experience shows that senior management must help set the right goals, put the right people in the right place, and make sure the different pieces of the organization work together toward the common goal. That goal is not just regulatory compliance, but a continually improved rating system that reduces credit losses in the short term and sustains long-term bank profitability and competitive advantage.

---

**Bogie Ozdemir** is a vice president at Standard and Poor’s Risk Solutions. Contact him by e-mail at bogie_ozdemir@standardandpoors.com.

---

**Notes**


**References**
