

known. The most important type of benchmarking of a rating system is to ask whether another rater or rating method attaches the same rating to a particular obligor or facility. Regardless of the rating approach, the benchmark can be either a judgmental or a model-based rating. Examples of such benchmarking include:

- Ratings reviewers who completely re-rate a sample of credits rated by individuals in a judgmental system.
- An internally developed model is used to rate credits rated earlier in a judgmental system.
- Individuals rate a sample of credits rated by a model.
- Internal ratings are compared against results from external agencies or external models.

Because it will take considerable time before outcomes will be available, using alternative ratings as benchmarks will be a very important validation device. Such benchmarking must be applied to all rating approaches, and the benchmark can be either a model or judgment. At a minimum, banks must establish a process in which a representative sample of its internal ratings is compared to third-party ratings (*e.g.*, independent internal raters, external rating agencies, models, or other market data sources) of the same credits.

Benchmarking also includes activities designed to draw broader inferences about whether the rating system—as opposed to individual ratings—is working as expected. The bank can look for consistency in ranking or consistency in the values of rating characteristics for similarly rated credits. Examples of such benchmarking activities include:

- Analyzing the characteristics of obligors that have received common ratings.
- Monitoring changes in the distribution of ratings over time.
- Calculating a transition matrix calculated from changes in ratings in a bank's portfolio and comparing it to historical transition matrices from internal bank data or publicly available ratings.

While benchmarking activities allow for inferences about the correctness of the ratings system, they are the not same thing as back-testing. The benchmark itself is a prediction and may be in error. If benchmarking evidence suggests a pattern of rating differences, it should lead the bank to investigate the source of the differences. Thus, the benchmarking process illustrates the possibility of feedback from ongoing validation to model development,

underscoring the characterization of validation as a process.

Back Testing

S. Banks must develop statistical tests to back-test their IRB rating systems.

S. Banks must establish internal tolerance limits for differences between expected and actual outcomes.

S. Banks must have a policy that requires remedial actions be taken when policy tolerances are exceeded.

The third component of a validation process is back-testing, which is the comparison of predictions with actual outcomes. Back-testing of IRB systems is the empirical test of the accuracy of the parameter values, PD and LGD, associated with obligor and loss severity ratings, respectively. For IRB rating systems, back-testing addresses the combined effectiveness of the assignment of obligor and loss severity ratings and the calibration of the parameters PD and LGD attached to those ratings.

At this time, there is no generally agreed-upon statistical test of the accuracy of IRB systems. Banks must develop statistical tests to back-test their IRB rating systems. In addition, banks must have a policy that specifies internal tolerance limits for comparing back-testing results. Importantly, that policy must outline the actions that would be taken whenever policy limits are exceeded.

As a combined test of ratings effectiveness, back-testing is a conceptual bridge between the ratings system architecture discussed in this chapter and the quantification of parameters, discussed in Chapter 2. The final section of Chapter 2 discusses back-testing as one type of quantitative test required to validate the quantification of parameter values.

III. Quantification of IRB Systems

Ratings quantification is the process of assigning numerical values to the four key components for internal ratings-based assessments of credit-risk capital: probability of default (PD), the expected loss given default (LGD), the expected exposure at default (EAD), and maturity (M). Section I establishes an organizing framework for considering IRB quantification and develops general principles that apply to the entire process. Sections II through IV cover specific principles or supervisory standards that apply to PD, LGD, and EAD respectively. The maturity component, which is much less dependent on statistical estimates and the use of data, receives somewhat different treatment in section V.

Validation of the quantification process is covered in section VI.

A. Introduction

Stages of the Quantification Process

With the exception of maturity, the risk components are unobservable and must be estimated. The estimation must be consistent with sound practice and supervisory standards. In addition, a bank must have processes to ensure that these estimates remain valid.

Calculation of risk components for IRB involves two sets of data: the bank's actual portfolio data, consisting of current credit exposures assigned to internal grades, and a "reference data set," consisting of a set of defaulted credits (in the case of LGD and EAD estimation) or both defaulted and non-defaulted credits (in the case of PD estimation). The bank estimates a relationship between the reference data set and probability of default, loss severity, or exposure; then this estimated relationship is applied to the actual portfolio data for which capital is being assessed.

Quantification proceeds through four logical stages: obtaining reference data; estimating the reference data's relationship to the parameters; mapping the correspondence between the reference data and the portfolio's data; and applying the relationship between reference data and parameters to the portfolio's data. (Readers may find it helpful to refer to the appendix to this chapter, which illustrates how this four-stage framework can be applied to ratings quantification approaches in practice.) An evaluation of any bank's IRB quantification process focuses on understanding how the bank implements each stage for each of the key parameters, and on assessing the adequacy of the bank's approach.

Data—First, the bank constructs a reference data set, or source of data, from which parameters can be estimated.

Reference data sets include internal data, external data, and pooled internal/external data. Important considerations include the comparability of the reference data to the current credit portfolio, whether the sample period "appropriately" includes periods of stress, and the definition of default used in the reference data. The reference data must be described using a set of observed characteristics; consequently, the data set must contain variables that can be used for this characterization. Relevant characteristics might include external debt ratings, financial measures, geographic regions, or any other factors that are believed to be