

# ISDA

International Swaps and Derivatives Association



## ISDA-RMA-BBA Internal Ratings Validation Survey Executive Summary

A number of key messages have emerged as a result of the Survey. While these issues generally apply to all asset classes, their relative importance differs considerably. Below, we introduce each of the messages and briefly outline the implications for the corporate and retail asset classes respectively. More information on each of the issues is provided in the detailed thematic analysis that follows. It is important to note that the banks participating in the Survey represent to a certain extent the “cutting edge” of internal ratings development in their respective geographies. Consequently, the results of the Survey may tend to overstate the actual level of sophistication across the industry – a point industry bodies and regulators will need to keep in mind when addressing the implementation of the Basel II Accord.

### **1. Banks employ a wide range of techniques and diversity of practice with respect to internal ratings validation. Key differences can be seen with respect to the techniques used for corporate and retail ratings**

*Corporate Asset Class (including middle market).* Model types and validation techniques differ widely among institutions, within the corporate asset class. In general, banks use statistical models for parts of their portfolios, where the quantity of default data allows robust estimation. Expert judgment models are more common for those portfolios where default data is scarce. In some cases, hybrid models are used and / or vendor models complement the picture. Statistical techniques are more widespread in the middle market asset class, whereas expert judgment remains an important technique in the large corporate area. The techniques used to validate internal ratings and the assignment of PDs to individual rating classes also vary widely across institutions and geographies. Another aspect in this regard is the different level of experience and sophistication of various institutions with respect to internal ratings validation. Generally, institutions that have been using ratings for a long period of time and have built up internal data histories tend to utilize more quantitative/statistical techniques – where the data permits this approach. A large number of banks are beginning to address data inadequacies to support the validation regime for their internal ratings.

*Retail Asset Class.* For the retail asset class, statistical techniques are commonly used to assign a rating/ score. The greater availability of internal data history allows for more robust statistical validation tests, which are more widely used in the retail area than in other asset classes. Furthermore the Survey participants rely more

heavily on the statistical tests in this asset class and less on expert judgment. A small number of Survey participants have even set strict thresholds and triggers relating to monitoring of model performance and model re-development.

## **2. Ratings validation is not an exact science**

Even where banks employ statistical techniques to assess model performance during development or after implementation, with the exception of a small number of banks for retail models; they do not tend to use absolute triggers or thresholds. This entails that there is no absolute GINI coefficient, COC or ROC measure (or similar) that models need to reach in order to be considered adequate. In fact, absolute performance measures are seen by some banks as counter productive.

In cases where not enough internal default data exists, banks frequently resort to the use of external data, in particular default statistics published by the major rating agencies. How they use these default statistics (e.g. what time period to consider, which agency's data to use, whether it is appropriate to smooth the raw data, and if so, how this is achieved etc.) differs considerably from institution to institution and depends essentially on each individual bank's assessment of the most appropriate use of this external data. Benchmarking against external ratings also raises issues with regard to the unknown quality of external ratings, as well as methodology differences such as the time horizon under consideration, the default definition or the inclusion of LGD elements in the rating.

## **3. Expert judgment is of critical importance**

In a number of asset classes, particularly large corporates, but also banks and sovereigns, data scarcity makes it practically impossible to develop statistically based internal ratings models. Some banks use statistical techniques to establish the suitability of particular factors to risk rank customers, but stop short of determining the relative weights of the factors through statistical techniques. Even where weights are specified, there is usually a judgmental overlay to allow a modification of the model rating based on the assessment of a ratings expert (account officer, credit analyst or the like). Only in retail models is this judgmental overlay generally not applied to modify ratings but rather to modify the credit decision proposed by the model.

There is concern among the Survey participants that much of the discussion around internal ratings validation is centered on statistical techniques and absolute trigger ratios. Large proportions of banks' exposures are – and will be for the foreseeable future – however, covered by expert-judgment type rating systems and there is a feeling that not enough time has been spent on discussing acceptable validation techniques for these types of systems.

#### **4. Data issues center around quantity not quality**

While a large number of banks commented on the inadequacy of their internal data with respect to internal ratings validation, data quality was not seen as the major obstacle to validation in the long term. Most banks have initiated projects to collect the necessary data in a consistent manner across the entire institution, which should provide sufficiently robust data for validation purposes going forward.

Unlike data quality, the quantity of data – in particular of default data – poses a real problem for most institutions. This is particularly the case for the corporate, bank, sovereign and specialized lending exposure classes, whereas default data for retail and middle market exposures is generally considered to be adequate. For some exposure classes – notably banks and sovereigns – it is likely that there will never be enough default data to allow robust statistical estimates of default rates with a granular rating scale. Techniques other than statistical analysis will therefore be necessary to assess the adequacy of banks' rating systems and PD estimates.

A number of institutions have recently joined data pooling initiatives, both for PD and LGD data. It is, however, not yet possible to evaluate the applicability of these data pools for ratings validation going forward. Industry participants remain skeptical as to the ultimate usefulness of these initiatives with respect to overcoming the issues with data quantity outlined above.

#### **5. Definite regional differences exist with respect to internal ratings and their validation**

The structure of rating systems and the resulting validation techniques show definite regional differences. This is true both for corporate and retail exposure classes.

While statistical (scorecard) techniques are widely used for retail exposures across all institutions, these tend to be product-specific in the US and UK, while the focus in Continental Europe is clearly on customer scores/ratings. In addition, scorecards in the US/UK tend to be re-developed much more often, using the most recent available data, than those on the Continent, where robustness of ratings and the long-term stability of factors and their respective weights are of higher priority to institutions. This often has direct implications for the statistical measures used to assess model performance (e.g. the GINI coefficient), as the longer-term, more stable models tend to show lower GINIs than those models using the latest available data.

A similar divide can be observed for corporate ratings. While both North American and European banks tend to use expert judgment models for the assessment of their large corporate portfolios, the structures of these ratings (and consequently the validation techniques) differ significantly. In North America there are generally no fixed weightings for the factors to be assessed by the experts, whereas most European banks (in this case including most of the UK banks) set specific weights for each of the factors to be considered.

While vendors models based on equity market information (like KMV) or balance sheet information (like Moody's RiskCalc) are widely used by participants for their corporate and middle market portfolios, the application differs between North America and the rest of the world. In North America these models tend to be an integral part of the rating assignment process and are often used in a hybrid approach in conjunction with expert judgment. In Europe, such external vendor models are more often used as a benchmark or validation of the rating derived by the internal rating model. Market based methodologies do not generally lend themselves as sources for ratings assessments outside of North America, as fewer companies use the equity markets to raise funds.

## **6. Further work is necessary with respect to defining standards for stress testing**

Stress testing is widely practised within the industry, however, no uniform approach exists regarding the type of stress testing undertaken, its frequency, or actions undertaken in response to stress testing results.

Most banks currently undertake stress testing at a portfolio level, with risk ratings being a key input into stress testing scenarios for economic capital requirements. There is, however, uncertainty as to the level of additional stress testing potentially required under Basel II, as the QIS 3 Technical Guidance indicates a potential stress testing requirement for rating model inputs. This form of stress testing is not currently undertaken (or indeed planned) by a large number of institutions and there is some debate in the industry as to its potential relevance.

Further clarification on this issue on the part of the regulator as well as debate on the part of the industry is therefore considered necessary.

In summary the Survey has shown that rating validation is not an exact science but that banks continue to see expert judgment as a key component of the validation process. As a result, banks use a wide variety of methods for ratings validation, depending on the availability of default data and additional information such as external ratings. Data is a key concern for most Survey participants, however the problem is more the scarcity of default data than the quality and integrity of data in general.

The above key conclusions need to be considered in the context of the standards set out by the Basel Committee on Banking Supervision within both the [Second Consultative Paper \(CP2\)](#) issued in January 2001 and the technical guidance issued in relation to the third [Quantitative Impact Study \(QIS3\)](#) in October 2002. It should be noted that while the study was undertaken on the basis of the requirements as set out in the QIS Technical guidance, the answers are equally applicable to the requirements set out in the [Third Consultative Paper \(CP3\)](#) issued on April 29, 2003.