Development of Rating Models under IFRS 9

Ioan-Codruț ȚURLEA, PhD
Bucharest University of Economic Studies, Romania

Abstract

Before the release of the IFRS 9 standard Financial Instruments in 2014, the development of a rank ordering mechanism was mostly known through the Basel capital accords requirements for the computation of regulatory capital, as well as the economic capital models used for the estimation of internal capital needs. Most institutions would have been relying on application scorecards for ranking clients at application and assess their suitability to be granted a loan in line with their own risk tolerance. However, only a small number of institutions would have relied on behavioral scorecards. Both the Basel III Internal Rating Based Approach (IRBA) and IFRS 9 are principle based and offer their users a variety of modelling approaches. Hence, financial institutions are allowed to implement their own rating models. However, under IRBA the rating system must meet specific minimum requirements which are not required under IFRS 9.

The article focuses on highlighting a variety of rating methods and systems applicable under the IFRS 9 framework. Hence, it presents a series of statistical and non-statistical models for building and estimating the rating system. Furthermore, the benefits and drawbacks are presented for each approach. The paper concludes with an analysis of the models under the IFRS 9 framework.

Key terms: IFRS 9, Basel approach, credit risk, assessment, strategy

JEL Classification: M40, M41, G24

Introduction

This paper focuses on defining the statistical models that could be used by financial institutions to develop rating systems and offers an overview of the parametric and nonparametric methods which consider data availability. Additionally, it presents the benefits and drawbacks of each method.

Statistical rating systems are constructed using a selection of explanatory variables based on historical data points for non-defaulted obligors’ characteristics and behavioral assumptions. These data points are further used to establish a rating mechanism. Further analysis is carried out by institutions in order to estimate the probability of default for each of the specific grades.

In order to establish an obligor’s rating as accurately as possible, structural models use causal connections that rely on economic models. Hence, the models are based on the correlation between the obligor’s individual characteristics and the historical observed default rates.

Depending on the data availability throughout time, a variety of modelling approaches can be used to compute the ranking order for predicting the obligor performance.
Development of rating models under IFRS 9

Academic studies on the IFRS 9 standard *Financial Instruments* are scarce. Nobes (2013) developed one of the first studies on the classification of financial instruments under IFRS 9. In comparison to the IAS 39 *Financial Instruments: Recognition and Measurement*, IFRS 9 classifies financial instruments into three categories based on the evolution of the instrument’s riskiness since initial recognition – stage 1: instruments that are performing assets whose riskiness is similar to the one at initial recognition, with a 12-month expected credit loss (ECL) computed for these obligors; stage 2: instruments that are performing assets for which a significant increase in credit risk was identified and a lifetime expected credit loss is computed on the gross balance; stage 3: represents those assets that are defaulted and for which a lifetime expected loss is computed on the net balance.

In comparison to IAS 39, the computation of the IFRS 9 expected credit losses poses significant difficulties to the financial institutions due to the criteria incorporation of the significant increase identification in credit risk, as well as the incorporation of macroeconomic forward-looking information.

Moving away from the difficulties faced when calibrating the ECL, the basis for all modelling processes is the development of an accurate rating system, i.e., rank ordering mechanisms.

When it comes to the ranking systems and building scorecards, the viewpoint of Chawla *et al.* (2016) is that information related to the industry or geography should be included in the assessment, as there are significant differences between various, more risky, sectors and more cyclical economic responses. Another researcher, Edwards (2016), focuses on highlighting the importance of the term/maturity structure of the exposure, while Skoglund (2017) proposes various modelling approaches.

The current article presents multiple approaches for the developing of IFRS 9 compliant scorecards. In practice, financial institutions rely on the scorecards used for regulatory purposes, i.e., used for the Basel credit risk models. Most academic papers are focused on presenting the differences between the regulatory and IFRS 9 models. For example, Novotny-Farkas (2016) and, also, Miu and Ozdemir (2016) consider the main differences to be the calibration relying on prescribed floors for both *probability of default* (PD) and *loss given default* (LGD) parameters used for the regulatory capital calculation, while other studies, such as Reitgruber (2015), notes the models can satisfy both the Basel and IFRS 9 requirements allowing convergence across the institution.

From the author’s point of view, this article has an important impact on emphasizing the right perception regarding the development of rank ordering mechanisms having the purpose to obtain an adequate assignment stage. It addresses the synergies between the IFRS 9 and regulatory models regarding the ranking order mechanisms.

- **Statistic models and IFRS 9**

A rank ordering mechanism is necessary to satisfy IFRS 9 requirements regarding accurate assignment of PDs and LGDs and credit conversion factors (CCFs) for ECL calculation purposes and using a change in the rank ordering of an exposure as an identifier key of a significant increase in credit risk (SICR).

For internal rating based (IRB) purposes, a rank ordering mechanism is required for an appropriate assignment of exposures to a poor risk grade for the associated PD, LGD and CCF parameters with an increase in credit risk captured by grade migration due to a change in rank ordering. Hence, there is a strong overlap between the requirements of the rank ordering mechanism used for IFRS 9 purposes and the rank ordering mechanism used for IRB purposes.

Likewise, the IFRS 9 standard does not set up specific expectations on how the rating system should be structured. Furthermore, the International Accounting Standards Board (IASB) does not present any expectations around the methods and models applied when defining the rating system. In practice, most financial institutions rely on the pre-existing IRB definition provided by the European Banking Authority (EBA) and the requirements
defined in the Guidelines on credit institutions’ credit risk management practices EBA/GL/2017/06, namely that “the rating system is composed by all the methods, processes, controls, and data collection and IT systems that support the assessment of credit risk, the assignment of internal ratings, and the quantification of default and loss estimates”.

The methods presented in the article relate to the first component of the rating system, namely the internal ratings assignment through the use of a rank ordering mechanism.

From a regulatory perspective under a foundation or advance internal rating-based approach, institutions are expected to comply with the following grade requirements:

- To have at least seven rating grades for performing exposures (non-retail and retail) and one grade for default exposures;
- Initiations should create no undue or excessive concentrations in a single rating grade;
- Institutions should ensure a meaningful risk differentiation of risk achieved between rating grades;
- Plausible, intuitive, accurate and complete input data that is reflective of past experiences as well as future lending practices and credit risk management within the institution.

From the IFRS 9 perspective, an adequate definition and calibration of ratings are of the utmost importance, as it is considered the significant increase in credit risk (SICR) determination, i.e., transfer to stage 2, hence, for those exposures emphasizing the SICR criteria, a lifetime expected credit loss (ECL) computation would be required instead of 12 months.

The SICR criteria can be defined in absolute or relative terms, e.g., a change in PD since initial recognition, a rating downgrade, the exposure is more than 30 days past due. The SICR assessment is also influenced by the institution’s decision on the number of grades, i.e., it should enable an adequate differentiation between grades based on common risk characteristics, as well as ensure that the SICR could occur prior to the downgrade on a timely and accurate manner.

An adequate credit risk rating system allows for an appropriate differentiation of this risk between grades and captures both upwards and downwards migration and SICR, in order to accurately reflect the risk of individual facilities (for non-retail exposures) or obligors (for retail exposures), as well as the portfolio as a whole.

In this context, an effective and adequate credit risk rating system should allow institutions to identify both migration as well as the significant changes in this risk. Furthermore, in the case of externally rated obligors it would allow for a timely recognition of a downgrade made by the credit rating agency.

In accordance with best practices, the most common risk drivers used in scorecard development are past-due status, days past due information, loan-to-value and loan-to-income ratios, historical loss rates, product type, depreciation schedule type, down payments, prepayment information, market segment, geographical location, vintage (i.e., date of origination), collateral type, forbearance measures, as well as forward looking information, i.e., macro-economic factors. In case the model is meant to assess corporate or SME exposures, financial information is expected to be considered when the scorecards are built. Hence, scorecards would impact both the obligor and transaction risk information. However, the final decision regarding the driver’s selection needs business practices, portfolios’ specificities, as well the statistical testing results to be considered.

When grouping the obligors/the institution exposures one should ensure that the group remains homogenous in terms of credit risk drivers and that it is stable over time.

In order to meet the aforementioned objective, namely heterogeneity between grades and within the same grade, in a robust manner, institutions must consider among other risk drivers the correlation between macroeconomic factors and obligor attributes.

Ultimately, institutions should start the score card building process with a detailed historical patterns analysis, current trends, as well information regarding future credit risk practices which would allow the relevant risk drivers identification. Experience coming from non-modelling areas, i.e., business development, collection
strategies, pricing, etc., will serve for assessing the appropriateness of the statistically established risk drivers and for amending them accordingly based on expert judgement, in order to incorporate current and forward-looking information likely to affect those risk drivers.

In summary, the IFRS 9 and Basel guidance are not intended to limit the institution’s choice to a specific modelling approach, hence institutions should base their modelling choice on their internal data availability and internal constraints or limitations. The models discussed below can be tailored to allow institutions to avail of internal data estimates.

- **Regression analysis**

  Among the most used methods is regression analysis, i.e., the establishment of a linear relationship between the obligor’s characteristics and the default variable (based on observed default rates).

  1. **Regression analysis equation:**

     \[ y_i = \alpha x_i + \beta_i \]

     where:
     - \( y_i \) – defines if the obligor defaulted during the observation period \( t \) (\( y_i = 1 \)) or not (\( y_i = 0 \));
     - \( x_i \) – the vector of the observed obligor’s characteristics in period \( t \); for example, days past due, collateral type, product type, maturity, bankruptcy, forbearances measure, prepayment, etc.;
     - \( \alpha \) – the vector of the factors which capture the impact of the change in the characteristics of the default variable in period \( t \);
     - \( \beta_i \) – the residual variable.

     The equation above can be presented as an ordinary least squares (OLS) regression.

     2. **OLS equation:**

     \[ S_i = E(x_i | y_i) = \alpha \cdot x_i \]

     where:
     - \( S_i \) (the score) – is a continuous or discrete variable used as a ranking mechanism which takes values both greater than 1 and lower than 0.

     The equation depicts the obligor’s score based on the obligor’s individual characteristics.

     The main benefits of models depicted above are that these formulas are widely spread and therefore easy to compute and understand.

     The main drawback is that the variable \( \beta_i \) is heteroscedastic (different variances and standard errors of the variables \( i \)). Hence, the weighted least square estimator should be applied to compute \( \beta_i \) in order to try to ensure that the residuals are normally distributed. The selected variables should ensure the stability and predictive power of the model to the current and future portfolio’s structure, i.e., the \( \beta_i \) coefficient should ensure the existence of a significant relationship and not just a spurious correlation. In order to reduce these risks, a proper manner is to split in two, i.e., development and validation must be divided to ensure consistent results.

- **Discriminant analysis**

  Altman (1968) presents the discriminant analysis as a technique applied to corporate obligors in order to predict their bankruptcy. The linear discrimination function is based on the principle that the spread between the good and bad obligors (non-default and default) should be maximum, as well as between individual grades.

  3. **Discriminant analysis equation:**

     \[ S_i = \alpha \cdot x_i \]
S_i defines the discrimination variables, while the maximization process allows the α vector coefficients to be the optimal results. The coefficients are normalized. As such S_i is arbitrary, hence it cannot be assessed on a stand-alone basis. While the result of the linear regression can be interpreted as the higher score the higher risk.

Within the methods 1 and 2 the characteristics are deterministic on and the default state is presented by the random variable, while method 3 is the opposite.

Overall, the differences are mostly theoretical and in practice they are irrelevant. The benefits and drawbacks are the same as for the regression model.

### Logit and probit models

When developing the rating system, there are two alternative theoretical foundations: logit and probit binary econometric models. These assume on unobservable (latent) variable y which represents the obligor’s characteristics.

4. **Logit and probit equation 1:**

   \[ y_i = \alpha x_i + \beta_i \]

   The variables and parameters are defined as in model 1. However, the variable \( y_i \) is scaled as follows:

5. **Logit and probit equation 2:**

   \[
   \begin{cases}
   1, & y_i > 0 \\
   0, & \text{otherwise}
   \end{cases}
   \]

   As such the default event occurs when the latent variable is greater the 0, and the function becomes:

6. **Logit and probit equation 3:**

   \[
   P(y_i = 1) = P(\beta_i > -\alpha \cdot x_i) = 1 - F(-\alpha \cdot x_i) = F(-\alpha \cdot x_i)
   \]

   This assumes a symmetry density distribution around 0. Furthermore, it depends on the assumptions used around residuals, as follows:

   ✓ When the residuals are normally distributed, a probit model is considered.

7. **Probit model equation:**

   \[
   F(\alpha \cdot x_i) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha \cdot x_i} e^{-t^2/2} dt
   \]

   ✓ When the residuals follow a logistic distribution, a logit model is considered.

8. **Logit model equation:**

   \[
   F(\alpha \cdot x_i) = \frac{e^{\alpha \cdot x_i}}{1 + e^{\alpha \cdot x_i}}
   \]

   Another option would be to start the assessment from the estimation of the default probabilities. The following consideration must be taken into account: for a single obligor, default probabilities cannot be observed, only observed default rates can be computed at a portfolio or grade level. In this case, the observed default frequencies (\( p_i \)) can be interpreted as default probabilities. Thus, OLS estimation becomes:

9. **OLS for a single obligor equation:**

   \[ p_i = \alpha \cdot x_i + \beta_i \]

   Equation 8 is not bounded to values between 0 and 1, hence it cannot be used in the scoring assessment on its own. In order to scale \( \alpha \cdot x_i \) between 0 and 1, the linear expression is transformed by a nonlinear function \( F \).
10. Nonlinear function $F$ equation:

$$p_i = F(\alpha \cdot x_i)$$

By selecting a logistic link function, the equation becomes a logit model, while considering a normal distribution it becomes a probit model.

When using OLS, the coefficients should be heteroscedastic; this can be achieved through the use of the weighted least squares regression, which would lead to the observed default frequencies being grouped before the estimation. In practice, this can pose problems as it would require defining the size and number of pools or grades ahead of the analysis.

Another option is to use the maximum likelihood method instead of grouping them. Hosmer and Lemeshow (2000) have detailed the approach and the tests to validate the model and its variables.

In practice, the outcome of the two models is often similar as the distributions have a similar form, with the exception of the tails, the probit model having a fatter tail. However, the logit model is used more in practice as it is easier to understand.

11. Practice logit model equation:

$$\frac{P_i}{1 - P_i} = e^{\alpha \cdot x_i}$$

The left-hand side represents the odds, i.e., the relation between the default probability and the probability of survival. Hence, the variation of a single variable $x_i$ of one unit has an impact of $e^{\alpha i}$ on the odds, where $\alpha$ denotes the coefficient of the variable $x_i$ and the transformed coefficients $e^{\alpha_i}$ represent odds-ratios. They constitute the multiplicative impact of an obligor’s characteristic on the odds. The most important weakness of binary models is the interpretation of the coefficients, which is not straightforward.

The main strengths of the logit and probit models are their easy validation and interpretation of results, i.e., the output of the model can be interpreted directly as a default probability.

- **Panel models**

  The aforementioned-methods are cross-sectional, i.e., all variables were related to the same time horizon. However, institution avail of data spreading over multiple time periods. Hence, the cross-section data is spread to a panel dataset in order to increase the stability and accuracy of the model.

  Furthermore, by integrating the latest data, macroeconomic variables or stressed data, the model’s performance can be improved.

  The structure of a panel logit or probit model remains the same as presented in the previous section; the only difference consists in the incorporation of a time series indicator: $x_{it}$ instead of only $x_i$.

  In practice, when developing the model, the dynamic pattern of the variables might be ignored and simply fit the probit or logit model. Both probit and logit models require independent variables, criteria met by the cross-sectional data. However, the panel-data does not meet the criteria as observations from the same date and observations from the same obligor can be correlated.

  Chamberlain (1980) proposes to consider under the logit model the modified fixed-effect estimator by excluding all non-defaulting exposures. Additional analysis was carried out by Hosmer and Lemeshow (2000) on the binary panel analysis.

- **Hazard models**

  The models presented so far assess the riskiness of the obligor by computing a score that indicates whether the obligor defaults within the specified horizon. However, the models do not allow for a comparison with or an analysis on emphasizing the obligor’s risk evolution and it is not determined when the obligor defaults will occur.
In order to remediate the deficiencies of the aforementioned methods, Cox (1972) has proposed the use of hazard models which take into account the survival of obligors through time. These assume that the underlying hazard rate is presented as a semiparametric model relying on independent variables and does not rely on the nature or share of the underlying survival distribution.

12. **Hazard model equation:**

\[ h_i(t|x_i) = h_0(t) \cdot e^{\alpha \cdot x_i} \]

where:

- \( h_i(t|x_i) \) – the resultant hazard, considering the obligor’s covariation and survival time \( t \);
- \( h_0(t) \) – the baseline hazard when all independent variable values are set up to be 0. It can be seen as the average hazard rate of the obligor.

However, the equation assumes a multiplicative relationship between the hazard and the log-linear function of the explanatory variables. This relationship implies that outcome of the assessment is a continuous random variable, even though an obligor’s risk drivers are observed at discrete points in time and are not continuous throughout the period. Furthermore, the variables are treated as constant over time, while in general explanatory variables change.

As seen, the main strength of the model is the estimation of the survival period, while its weakness lies in the fact that the model is more complex than the previous presented approaches.

- **Neural networks**

Neural networks are an alternative to the parametrical models as they are non-parametrical and offer a more flexible approach between the independent and dependent variables.

This method is based on multiple nodes that send a specific output if they receive predefined information from the other nodes to which they are linked, i.e., the neural network uses a design sample in order to classify the obligors into grades or pools based on similar characteristics. Hence, the final network is defined by the interlinkage between the input and the output and through the intermediary nodes.

The neural networks easily model high complex non-linear relationships while ensuring high flexibility of the distributional assumptions and can be quickly adapted to new information. However, there is no predetermined mechanism to determine the optimum network as the connections are seen as black boxes that are difficult to interpret. Furthermore, the computation of default probabilities is limited.

- **Decision trees**

Decision trees are a further category of non-parametric models. The model is build based on a series of “if and then” conditions splitting the obligors into different groups. In the case of a binary classification tree, the population is subdivided in two subgroups and after each node unit the end note is reached. As such they are used for data containing a limited number of predictive variables. By doing so, all the variables are grouped and treated as categorical variables.

The main algorithms for building decision trees are proposed by Breiman *et al.* (1984) through the classification and regression trees algorithms and by McLoughlin and Kass (1978) through the CHAID algorithm chi-square automatic interaction detector, even though the algorithms use different criteria to identify the best grouping and to aggregate the categories which are not significantly different.

As in the case of neural networks, decision trees are not based on distributional assumptions and the outputs are easy to understand. However, the probability of default should be calculated in a separate step.
Conclusions

Rank ordering mechanisms are necessary to satisfy IFRS 9 requirements relating to the accurate assignment of PD, LGD and CCF parameters for ECL calculation and for the identification of a specific increase in credit risk. There is a strong overlap between the requirements of rank ordering mechanisms used for IFRS 9 purposes and the requirements of those used for pre-existing IRB purposes. Hence, financial institutions predominantly leverage existing IRB rank ordering mechanisms for IFRS 9 purposes.

IFRS 9 and IRB standards offer a high level of flexibility for the development of rank ordering mechanisms. This paper outlines a range of options used by financial institutions to develop such mechanisms.

Financial institutions’ choice of rank ordering methodology may be determined by several factors, such as the nature of the data available, time and resources constraints or an appropriate tradeoff between the model, the complexity and transparency.

Regardless of the methodology chosen, all approaches require the identification of appropriate risk drivers to describe the target variable, e.g., default for a PD estimate and a LGD estimate.

It is worth noting that, no matter how robust the selected modelling methodology is or how great the data availability, the rank ordering mechanism will only be robust if it is developed on a data sample that is representative for the exposures population on which it will be applied. Furthermore, even if a robust rank ordering mechanism will ensure that a rating system discriminates well, this does not necessarily mean that it will predict well. To ensure that the overall model predicts well, the output of the rank ordering mechanism must be appropriately calibrated.

Among the limitations of the study, we consider that it is highly important to elaborate a company cross analyses in order to emphasize both theoretical assumptions and practical duties at the same level of compliance. However, we should make sure whether any domain has a particular level of compliance in order to obtain an accurate perception regarding credit risks. Further research directions include extending the sample both on the national level and by including companies with similar characteristics from other countries, as well as alternative ways that could influence the companies’ decisions.

References


