

## IFRS 9 – Tests for PD Model Validation

Moldoveanu Marian Valentin, Despa Madalin-Mihai, Achim Luminita-Georgiana, Cazazian Maria Rafaela,  
Bucharest Academy of Economic Studies, Romania

---

**ABSTRACT:** *With the coming into force of the IFRS 9 standard, financial institutions have went from an incurred loss model to a forward looking model for the computation of impairment losses. As such, the IFRS 9 models use point-in-time (PIT) estimates of PDs and LGDs and provide a more faithful representation of the credit risk at a given PIT as they are based on past experiences as well as the most recent and forecasted economic conditions. However, given the short-term fluctuations in the macroeconomic conditions, the final outcome of the Expected credit loss (ECL) models is highly volatile due to their sensitivity to the business cycle. In order to prevent financial institutions' over or under provisioning after the models have been developed, they need to be adequately monitored and validated and if necessary re-calibrated in order to ensure that the outcome of the models are accurate.*

*As such, the paper focuses on the validation of PD models under the IFRS 9 standard, presenting the complexity and challenges of developing the PD models 9 and a selection of qualitative and quantitative techniques applicable in the monitoring or validation processes.*

**KEY WORD:** *IFRS 9, Framework, Model Validation*

---

Date of Submission: 04-04-2022

Date of Acceptance: 19-04-2022

---

### I. INTRODUCTION AND LITERATURE REVIEW

Banks had to invest in the development of new models, Institutions that were using internal rating models for their Pillar 1 and Pillar 2 capital quantifications and had an advantage as they were able to adapt the models, to satisfy the requirements of IFRS 9. On the other hand, banks that were using the standardized approach for their Pillar 1 and Pillar 2 capital quantification and that had no exposure to managing credit risk statistical models encountered significant issues e.g. data availability and data quality and a lack of model development and validation expertise. However, the IFRS 9 standard only sets the expectation that “an entity shall regularly review the methodology and assumptions used for estimating expected credit losses to reduce any differences between estimates and actual credit loss experience”. The Basel Committee on Banking Supervision (BCBS) and the European Banking Authority (EBA) have issued guidance on their expectations on the appropriate governance, including validation expectations of impairment models. Furthermore the Global Public Policy Committee (GPPC) paper describes how an appropriate methodology should look like for IFRS9. Currently these papers are considered reference points for the IFRS 9 validation Framework. The challenges of validating expected credit losses arise due to the introduction of the lifetime expected credit losses and staging concepts. Furthermore, among the most important pre-requirements is the data and the data quality framework. The data quality framework, should at a minimum ensure the accuracy and availability of data by tracing them to the source systems and performing reconciliations.

#### 1.1 Research Methodology and Data Analysis

In order to achieve the objectives of the paper, the study will be based on methods specific to scientific research. The fundamental purpose of the methodology is to help us understand, not so much the products of science, but the process of knowledge itself. The methodology of the scientific research that we will use in this paper combines the qualitative research with the quantitative one, stating that the efficiency of the results obtained from the research would be greater if an optimal combination between the qualitative research and the quantitative one was achieved, in order to meet the objectives, set.

The paper also puts emphasis on the importance of the control framework and encourages institutions to focus on the estimation and reporting of the ECL by establishing key performance indicators which can be used as tools for challenging the model's performance. In conjunction with the IFRS 9 requirements, the GPPC presents a non-exhaustive list of qualitative elements to be considered when developing/validating IFRS 9 models:

## 1) Regarding the expected credit loss methodology:

- Institutions are required to include in the models the connection between empirical facts caused by the economic reality. An example of non-compliance would be the use of fair value models to compute ECLs without adequately adjusting for changes in market values of interest rates and yields.
- Institutions are required to include forward looking components in the models, hence it is non-compliant to use expected losses calculated for regulatory purposes as ECL without assessing their compliance with the IFRS 9 requirements, furthermore the institutions should assess whether any adjustments should be applied before the models are fit for use under the IFRS 9 requirements e.g. based on best practices the PD models are adjusted to incorporate forward looking information, furthermore to strip out the throughout the cycle component while for the LGD models the downturn component and the inclusion of indirect costs is stripped out.
- The selection of risk drivers should be based on statistical tests (if possible) and their selection or exclusion should always be justified i.e. explanations should be provided especially in the case of variables selected based on expert judgement s. There is always a trade-off between discriminatory power, predictive power at grade level and stability. When it comes to the discriminatory power, the institution should ensure that the grades or pools defined (segments) share the same credit risk characteristics in order to ensure that changes in credit risk performance in a part of the portfolio will not be offset by the performance of other elements when their outputs are collectively assessed and measured.
- Institutions are required to include prepayment information in the models. I.e. include the effects of contractual repayments, prepayments as well as drawdowns.

## 2) Regarding the definition of default:

- Institutions should ensure the consistency of the definition of default trough time, i.e. The Definition of default use for modelling the probability of default for IFRS 9 purposes should be the same as the definition of default implemented and used by the institution in its live environment. In case there are differences between the definitions of default, the institution should carry analysis showing that the impact is not material and in case it is deemed material the institution is expected to be addressing the shortcomings through additional management overlays. The main reason behind this approach is that it is non-compliant for institutions to use a definition of default that generates fewer default events than actually monitored and observed through its credit risk management processes.
- As aforementioned, institutions are expected to investigate the differences and assess their impact on the staging distribution and ECL calculations. Furthermore, the institution is expected to align its IFRS 9 definition of default to the Information used for regulatory purposes should be assessed and if applicable adjusted to be fit for use under IFRS 9. The bank.
- Institutions should use the 90 days past due backstop as it is considered a benchmark in the financial industry. In order not to diverge from the 90 days past due backstop the institutions should demonstrate using reasonable and supportable information that a more lagging default criterion is more appropriate.

## 3) In relation to the Probability of default (PD):

- When basing the IFRS 9 models on already existing IRB Models, the institution should first determine if the models are fit for IFRS 9 purposes or they have to be adjusted in order to become IFRS 9 compliant.
- Institutions using a simplified approach for PD modelling should document and justify why this approach is reasonable. I.e. it would be inappropriate to consider a constant marginal rate of default over the remaining lifetime of a product without appropriate supporting analysis.
- Discriminatory power is essential for model accuracy, as it defines the link relationship between economic reality and the statistical model. A balance should be identified between the use of models and the incorporation of expert based opinions. “The entity should not obscure this information by grouping financial instruments with different risk characteristics. Examples of shared credit risk characteristics may include, but are not limited to, the: instrument type, credit risk rating, collateral type, date of initial recognition, remaining term to maturity, industry, geographical location of the borrower, the value of collateral relative to the financial asset if it has an impact on the probability of default occurring.”

The European Banking Authority Guidelines on Accounting for Expected Credit losses, EBA/GL/2017/06 in 12 May 2017, brings additional clarifications on the governance arrangement on validation, monitoring and review processes under “Principle 5 – ECL model validation”:

- 1) An adequate governance should be established (policies and procedures) to ensure at a minimum the following: accuracy and consistency of the models, risk rating systems and processes as well as an adequate estimation of all relevant risk components (PD, LGD, EAD). Furthermore, the role of professional judgement should be detailed along with the identified model limitations their impact and the mitigating actions considered to address them.
- 2) Model validation should be carried out at model development, as well as after the development of the model, through periodic validation and monitoring. Furthermore, specific considerations should be given when significant changes are made to the models, in order to ensure that the models continue to be fit for use.
- 3) The IFRS 9 models are expected to be updated frequently to ensure that changes in the macroeconomic conditions are factored into the models to comply with the point in time requirements as well as the use of most recent and updated information. Furthermore, the models should enable the incorporation the impact of changes in borrower riskiness (relative and absolute) and credit risk-related variables such as: PDs, LGDs, exposure amounts, collateral values, internal rating.

In practice, before the implementation of IFRS 9, some financial institutions did not have scorecards and for this reason they did not have adequate risk scoring mechanisms and 12 month PDs estimates. In such cases, institutions had to apply simplified assumptions in order to obtain an estimate for the PDs i.e. benchmarks obtained from the credit rating institutions or banking system information reported by national regulators. In most cases the derived PDs were based only on days past due information (without taking into consideration any unlikely to pay criteria), hence the segmentation criteria used as the basis of the lifetime estimates were days past due buckets. In such cases, the validation process focuses mostly on qualitative rather than quantitative criteria. Overfitting is one of the most common modelling mistakes i.e. the model explaining the particular sample, but ignoring other particularities of the entire population. With the help of out of time, out of sample validation tests the overfitting issue is identified and dealt with.

In the case of a 12 months PD, the following tests can be performed to assess the calibration of the model:

- Hosmer-Lemeshow or Chi-Square Test –comparing the observed versus predicted default rates for each pool or rating grade;
- Binomial Test - comparing the observed versus predicted default rates;
- Calibration Curve Shape Test – is a graphical method that can be used in corroboration with other tests.

It is based on establishing a confidence bound around the values the model predicted and depending on the number of instances for which actual outcomes lies outside this confidence bound, the result can be classified as a Red, Amber or Green. Regardless of the method used for PD lifetime calculation, it is important to start the validation process by assessing if the population of the development sample is comparable to the structure of the portfolio.

Another way of validating the lifetime PD estimates is by assessing the monotonicity of the lifetime’s curves. It is expected that the lifetime PD curves are ordered according to the ranking imposed by the underlying risk drivers hence the riskier rating must have a higher level than that of a curve with a lower risk profile, hence no intersection between the grade level PD lines is expected. The advantage of this validation method over the MSE is that it is assess the discriminatory power of the risk drivers over a lifetime horizon.

The following steps can be undertaken to assess monotonicity:

- A matrix is constructed; rows represent the risk drivers and the columns reflect the time horizon i.e. the years/months (most granular level based on which the PD was estimated).
- For all periods for which lifetime PD is estimated the value 0 is given to a monotonicity flag (MF) for the best risk driver level (the less risky rating grade).

As the risk drivers are ordered in ascending order based on their riskiness, each subsequent risk driver level is compared with the previous level:

$$MF_{i,j} = \begin{cases} 1, & PD_{i,j} > PD_{i-1,j} \\ 0, & PD_{i,j} \leq PD_{i-1,j} \end{cases}$$

Where:

I – row dimension of matrix;

j – column dimension of matrix;

PD – estimated cumulative PD

- The following formula is applied for each risk driver level (rating):

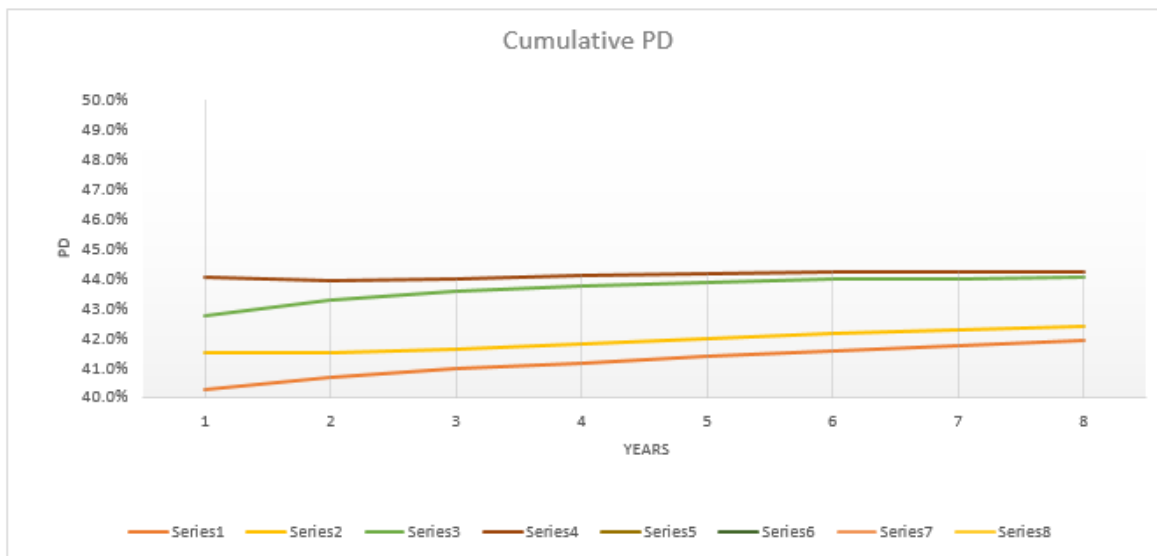
$$Monotonicity = \begin{cases} TRUE, & MIN(MF_{i,1} - MF_{i,n}) = 0 \\ FALSE, & MIN(MF_{i,1} - MF_{i,n}) <> 0 \end{cases}$$

- If all risk drivers levels have Monotonicity = TRUE then the lifetime PD curves are ordered according to risk level.

The following graph illustrates a valid monotonicity test.

Table 1: Monotonicity test

Year after reporting date									
Rating	1	2	3	4	5	6	7	8	Monotonicity
A	0	0	0	0	0	0	0	0	TRUE
B	0	0	0	0	0	0	0	0	TRUE
C	0	0	0	0	0	0	0	0	TRUE
D	0	0	0	0	0	0	0	0	TRUE



Source: own calculations

In case the curves intersect the risk drivers used for segmentation are not adequate for the proper discrimination between the levels of risk on a lengthy horizon. A possible cause could be the reduced discriminatory power of the rating system however, another possible explanation can be the fact that due to multiple calibration exercises performed over time the development/validation sample is no longer homogenous.

For the assessment of the marginal life-time PD obtained for each year Jeffrey’s test can be used. The rest requires them to split both the lifetime empirical curve and forecasted marginal PDs into yearly vintages.

The test compares the forecasted defaults with the observed defaults in a binomial model with independent observations. The null hypothesis is that the PD applied at the beginning of the relevant period in the sub-portfolio is greater than or equal to the modeled one. It is a one-sided hypothesis test.

Another statistical tests for the validation of the lifetime PD is the measurement of the “Z1” shift. “Z-shift measures the influence of the systematic component on credit risk and can be estimated by using a PIT transition matrix and an estimated TTC transition matrix. Under this assumptions, the lifetime PD estimation is based on the assumption that past behavior can be the basis of the prediction of the future behavior, as such the PDs must be as close as possible to ideal TTC. This validation method is recommended to be used for the lifetime PD computed using Markov chains to derive an aggregated TTC transition matrix, however if the lifetime PD is computed using the survival analysis, the data must be remodeled.

The test has the following set of hypotheses:

$$\begin{cases} H_0: R_{Observed} \leq \hat{R} - \text{past behaviour cannot be used to determine future behaviour} \\ H_1: R_{Observed} > \hat{R} - \text{past behaviour can be used to determine future behaviour} \end{cases}$$

<sup>1</sup> “Z” is defined as the systematic factor that influences credit risk and can be assumed to follow a standard normal distribution

Where:

$R_{Observed}$  – coefficient of determination between optimal Z-shift series and observed PIT-TTC error series;

$\hat{R}$  – coefficient of determination desired threshold (set by the financial institution);

In order to perform the test the following steps should be performed:

- (1) The TTC matrix is computed based on all transition events used for the lifetime PD estimation.

$$TTC_{matrix} = \begin{pmatrix} TTC_{11} & \cdots & TTC_{1j} & TTC_{1m} \\ \vdots & \ddots & \vdots & \vdots \\ TTC_{n1} & \cdots & TTC_{nj} & TTC_{nm} \end{pmatrix}$$

Where:

n – Number of rows of TTC matrix;

m – Number of columns of TTC matrix;

$TTC_{ij}$  – Transitions (%) from state “i” to state “j”;

- (2) The next step is the computation of the inverse of the standard normal cumulative probabilities of the TTC matrix. In case an absorbing states is defined, such as default, closed, prepaid, restructured etc. the corresponding rows will be excluded from the bin matrix. Ideally, no absorbing state should be defined in order for the full spectrum of possible transitions to be captured. The first column of the matrix doesn’t need a transformation, this is due to the fact that state 1 is limited to bin  $[\infty, \Phi^{-1}(TTC_{2j})]$ .

$$BIN_{matrix} = \begin{pmatrix} [\infty, \Phi^{-1}(TTC_{12})] & [\Phi^{-1}(TTC_{12}), \Phi^{-1}(TTC_{13})] & \cdots & [\Phi^{-1}(TTC_{1m}), -\infty] \\ \vdots & \vdots & \ddots & \vdots \\ [\infty, \Phi^{-1}(TTC_{n2})] & [\Phi^{-1}(TTC_{n2}), \Phi^{-1}(TTC_{n3})] & \cdots & [\Phi^{-1}(TTC_{nm}), -\infty] \end{pmatrix}$$

Where  $\Phi^{-1}(TTC_{ij}) = \Phi^{-1}(\sum_j^m TTC_{ij})$  – in other words the sum of TTC transition probabilities of each row, from “j” column to last column;

- (3) The initial shifted transition matrix for a Z-shift = 0 (no shift) is computed.

$$Shifted_{matrix} = \begin{pmatrix} 1 - \sum_{j=2}^m Shifted_{1j} & \Phi(BIN_{12} - ZShift) - \sum_{j=3}^m Shifted_{1j} & \cdots & \Phi(BIN_{1m} - ZShift) \\ \vdots & \vdots & \ddots & \vdots \\ 1 - \sum_{j=2}^m Shifted_{nj} & \Phi(BIN_{n2} - ZShift) - \sum_{j=3}^m Shifted_{nj} & \cdots & \Phi(BIN_{nm} - ZShift) \end{pmatrix}$$

- (4) PIT transition matrices are computed for each transition in the input data. The same level of granularity as the one used for PIT matrix computation must be used, for example if monthly data is used, then monthly PIT transition matrices must be computed.

$$PIT_k = \begin{pmatrix} PIT_{11} & \cdots & PIT_{1j} & PIT_{1m} \\ \vdots & \ddots & \vdots & \vdots \\ PIT_{n1} & \cdots & PIT_{nj} & PIT_{nm} \end{pmatrix}$$

Where “k” is the transition period.

- (5) A series of error matrices as squared difference between shifted transition matrix and each PIT matrix is computed. The errors are summed up on each row and then the final error is the sum of the sums.

$$ERROR_k = \begin{pmatrix} (Shifted_{11} - PIT_{11})^2 & \cdots & (Shifted_{1j} - PIT_{1j})^2 & (Shifted_{1m} - PIT_{1m})^2 \\ \vdots & \ddots & \vdots & \vdots \\ (Shifted_{n1} - PIT_{n1})^2 & \cdots & (Shifted_{nj} - PIT_{nj})^2 & (Shifted_{nm} - PIT_{nm})^2 \end{pmatrix}$$

$$SSE_k = \sum_{i=1}^n \sum_{j=1}^m ERROR_{ij}$$

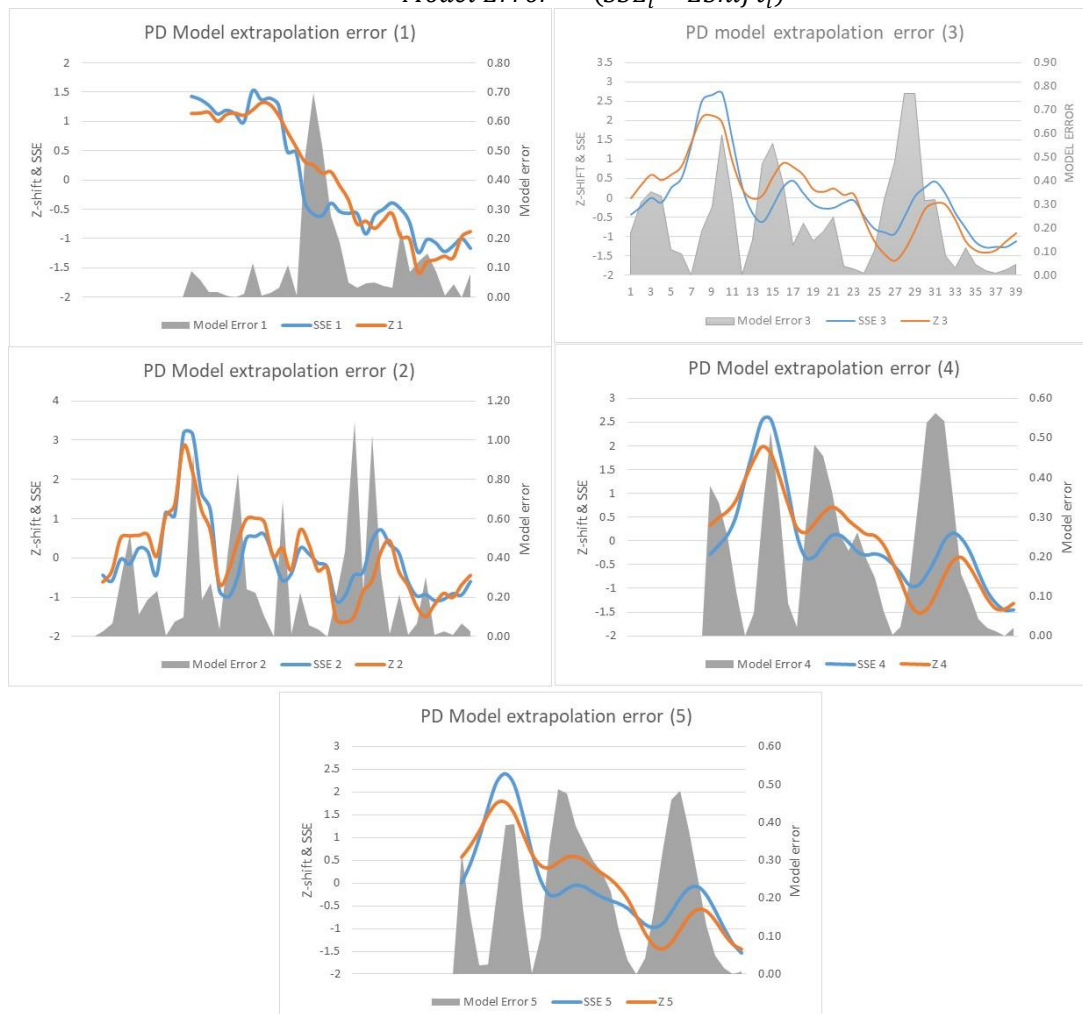
- a. The series of errors calculated at step 5 ( $SSE_k$ ) forms the empirical time series of differences between PIT and TTC that is embedded in PD lifetime calculation.
- (6) An optimization algorithm is applied to minimize each  $ERROR_k$  and reduce the distance between each PIT matrix and the estimated TTC matrix by using the Z-shift. The Generalized Reduced Gradient nonlinear algorithm available in excel can be used to find a Z-shift for each period. The restrictions are given by the shifted matrix and each PIT matrix and the objective is to minimize Z. The Z-shift series is then compared with the empirical time series.
- (7) To ensure a comparison between the two series a smoothening is applied in order to highlight the trend component in order to better reveal the underlying phenomenon and not to deprive the series of its specificity.
- (8) A linear regression is computed with the Z-shift series as dependent variable and empirical time series (SSE) as independent variable.

$$Z_{shift} = \alpha + \beta * SSE$$

- (9) The adjusted R square of the regression is  $R_{observed}$  used to test the hypothesis of the test. A high determination coefficient reveals that the SSE explains well the optimal TTC series of the portfolio and thus  $H_0$  can be rejected. The financial institutions should set the  $f \hat{R}$  threshold to be higher than 0.4.

Model error can be calculated by a square difference of each element of the SSE series and Z-shift series.

$$Model\ Error = (SSE_i - ZShift_i)^2$$



Source: own calculations



The above graph shows the application of the described algorithm on a PD lifetime on a portfolio composed which was grouped in 5 (PD model extrapolation 1 to 5) grades based on the days past due criteria, the estimation is performed over a 3 years' time horizon and incorporates all migration across the day past due grades. The adjusted R square obtained ( $R_{Observed}$ ). Is as follows:

Adjusted R Square Bucket 1	Adjusted R Square Bucket 2	Adjusted R Square Bucket 3	Adjusted R Square Bucket 4	Adjusted R Square Bucket 5
88.423%	76.529%	77.656%	77.937%	78.479%

### BIBLIOGRAPHY

- [1]. Sorin, Achim; Monica, Achim; Raluca, Streza, *Annals of the University of Oradea, Economic Science Series*;2008, Vol. 17 Issue 3, p907
- [2]. Adrian, T. and Shin, H. S., "Liquidity, Monetary Policy and Financial Cycles," *Current Issues in Economics and Finance*, Vol. 14, No. 1, January/February 2008.
- [3]. Basel Committee on Banking Supervision (1988), *International Convergence of Capital Measurement and Capital Standards*. <http://www.bis.org/publ/bcbs04a.htm>
- [4]. Basel Committee on Banking Supervision (2005), *International Convergence of Capital Measurement and Capital Standards: A Revised Framework*, BIS, Updated November 2005. <http://www.bis.org/publ/bcbs107.htm>
- [5]. Basel Committee on Banking Supervision (2005a), *Validation*, Newsletter No. 4. [http://www.bis.org/publ/bcbs\\_nl4.htm](http://www.bis.org/publ/bcbs_nl4.htm)
- [6]. Basel committee on banking supervision (2015), *Guidance on credit risk and accounting for expected credit losses*, <https://www.bis.org/bcbs/publ/d350.pdf>
- [7]. Beck, U., *Risikogesellschaft. Auf dem Weg in eine and ere Moderne*. Suhrkamp, Frankfurt a.M. 1986.
- [8]. Blochwitz, S. and Hohl, S., "Validation of Banks' Internal Rating Systems: A Supervisory Perspective," in Engelmann, B. and Rauhmeier, R., *The Basel II Risk Parameters*, Second Edition, Springer, 2011.
- [9]. Board of Governors of the Federal Reserve System, *Supervisory Guidance on Model Risk Management*, SR Letter 11–7, Washington, April 2011.
- [10]. Danielsson, J., "The Emperor has No Clothes: Limits To Risk Modelling," *Journal of Banking & Finance*, Elsevier, Vol. 26, No. 7, 1273–1296, July 2002.
- [11]. Derman, E. and Kani, I., "Riding on a Smile," *Risk*, Vol. 7, 32–39, 1994.
- [12]. European Banking Authority (2017), *Guidelines on credit institutions' credit risk management practices and accounting for expected credit losses*
- [13]. European Banking Authority (2017), *Guidelines on PD estimation, LGD estimation and the treatment of defaulted exposures*
- [14]. European Banking Authority (2016), *Guidelines on the application of the definition of default under Article 178 of Regulation (EU) No 575/2013*
- [15]. Global Public Policy Committee (2016), *The implementation of IFRS 9 impairment requirements by banks*
- [16]. Giddens, A., *The Consequences of Modernity*, Stanford University Press, Stanford, CA, 1990. JP Morgan, Report of JP Morgan Chase and Co. Management Task Force Regarding 2012
- [17]. International accounting standards Board (2014), *International Financial reporting standard 9 Financial instruments*
- [18]. Organisation for Economic Co-operation and Development, *OECD Principles of Corporate Governance*, Paris, 2004.
- [19]. Scandizzo, S., *Risk and Governance: A Framework for Banking Organisations*, Risk Books, London, 2013.
- [20].

Moldoveanu Marian Valentin, et. al. "IFRS 9 – Tests for PD Model Validation." *International Journal of Business and Management Invention (IJBMI)*, vol. 11(04), 2022, pp. 28-34. Journal DOI- 10.35629/8028