EUROPEAN CENTRAL BANK

BANKING SUPERVISION

# Instructions for reporting the validation results of internal models

IRB Pillar I models for credit risk

# Contents

# 1 Background and rationale

## 1.1 Introduction

Regulators and supervisors attribute a crucial role to the assessments conducted by internal validation functions throughout the life cycle of internal models. The term "internal validation function" encompasses the personnel of a credit institution who are responsible for validating internal models and reporting the results of the internal validation process.

In accordance with the requirements set out in Regulation (EU) No 575/2013[1] (the Capital Requirements Regulation or CRR), credit institutions must ensure that their Pillar I internal models for credit, market, counterparty credit and operational risk are subject to a validation process with the aim of verifying the overall adequacy, robustness and reliability of the internal estimates used to calculate own funds requirements.

The review carried out by the internal validation function is the first point of reference for supervisors when performing internal model-related tasks such as model approvals (initial approvals, model changes, extensions and roll-outs) and ongoing model monitoring.

The instructions in this document shall assist significant credit institutions supervised by the European Central Bank (ECB) under the Single Supervisory Mechanism (SSM) in complying with the decisions requiring the reporting of additional supervisory information from the validation function for credit risk models (hereinafter: Decisions). To this end the instructions (i) describe the background for the requested reporting of the significant credit institutions' validation results and (ii) explain the supplementary validation reporting using common metrics to assess Pillar I models for credit risk introduced by the Decisions. This supplementary reporting is without prejudice to credit institutions' own methods for validating internal models.

These instructions do not constitute a legal act, and they do not have any binding legal effect. Nothing within their wording, context or substance should be construed otherwise. These instructions are neither intended to replace or overrule applicable European Union (EU) law nor national law.

---

[1] Regulation (EU) No 575/2013 of the European Parliament and of the Council of 26 June 2013 on prudential requirements for credit institutions and investment firms and amending Regulation (EU) No 648/2012.

## 1.2 Regulatory framework

The CRR lays down minimum requirements with regard to the validation of Pillar I models for different risk categories.[2]

In accordance with the requirements set out in the CRR, the European Banking Authority (EBA) has drafted regulatory technical standards (RTS on the specification of the assessment methodology for competent authorities regarding compliance of an institution with the requirements to use the IRB Approach in accordance with Articles 144(2), 173(3) and 180(3)(b) of Regulation (EU) No 575/2013)[3] Until they are adopted, the provisions of the draft RTS are considered good practices for interpretative purposes.

The above-mentioned draft RTS attribute considerable importance to the assessments conducted by internal validation functions and provide competent authorities (CAs) with detailed instructions on how to examine the governance, methods and procedures of internal validation functions, as well as the soundness of the reporting process in terms of validation results.

In addition, Section 5 of the general topics chapter of the ECB guide to internal models[4] details processes and activities of the internal validation function on the basis of the requirements outlined in the CRR. The ECB guide to internal models provides clarity on how the ECB understands (i) the level of validation and associated responsibilities, (ii) the content and frequency of the tasks of the validation function, and (iii) reporting and follow-up processes.

## 1.3 The role of credit institutions' validation reporting in the ECB's ongoing monitoring of internal models

A sound validation function is crucial to ensuring the reliability of internal models and their ability to accurately compute capital requirements. It is the responsibility of the credit institution to ensure that its internal models are fully compliant with all regulatory requirements.

The ongoing assessment of permission to use internal approaches is governed by Article 101(1) of Directive 2013/36/EU (the Capital Requirements Directive or CRD),[5] which requires CAs to review credit institutions' compliance with the applicable requirements on a regular basis, and at least every three years.

---

[2]  For details of the main requirements as regards credit risk, see Articles 174(d), 185 and 188 of the CRR.

[3]  Draft EBA/RTS/2016/03 published on 21 July 2016.

[4]  See the "ECB Guide to internal models – General topics chapter", as published on the ECB Banking Supervision website on 15 November 2018.

[5]  Directive 2013/36/EU of the European Parliament and of the Council of 26 June 2013 on access to the activity of credit institutions and the prudential supervision of credit institutions and investment firms, amending Directive 2002/87/EC and repealing Directives 2006/48/EC and 2006/49/EC.

For the purposes of Article 101(1) of the CRD, ongoing supervision of models includes analysis of credit institutions' model validations, as stated in the ECB guide to banking supervision[6], as well as the SSM supervisory manual.[7]

Each credit institution produces validation reports on internal models in accordance with the scope and standards set out in the applicable regulation (such as the CRR) and the relevant RTS. The ultimate goal of those validation reports is to allow the credit institution's senior management to understand the performance and weaknesses of their internal models.

All individual validation reports should be shared with the ECB in accordance with the Decision, i.e. no later than one month after the validation reports are finalised and have been approved by the credit institution's management in accordance with the institution's internal policies.

---

[6]   See the "Guide to banking supervision", as published on the ECB Banking Supervision website in November 2014.

[7]   See the "SSM Supervisory Manual - European banking supervision: functioning of the SSM and supervisory approach", as published on the ECB Banking Supervision website in March 2018.

# 2 Supplementary validation reporting on credit risk

Credit institutions are required to develop validation processes to validate their internal models.[8] While all validation processes must comply with the same regulatory requirements, the ability to perform comparisons across models and institutions remains limited.

Against this background, supplementary validation reporting requested by a supervisory decision enables the ECB to carry out ongoing model monitoring mandated by Article 101(1) of the CRD by allowing an initial assessment of the performance of different models on the basis of a common set of statistical tests and analyses (validation tools) for Pillar I models relating to credit risk. The internal validation function within each credit institution is responsible for implementing this validation reporting.

Most of the validation tools focus on quantitative aspects of internal models. To provide clarity and facilitate the implementation of the validation tools and reduce the room for interpretation, detailed instructions on their implementation are provided below.

Under no circumstances should the validation tools referred to in this document be regarded as the minimum set of analyses needed to prove compliance with the applicable requirements set out in the relevant legislation. The validation of internal models remains the responsibility of the institution in question, which should determine the level of sophistication needed, taking into account the complexity and materiality (both current and prospective) of the internal models and their range of application.

This chapter outlines the description of the information which credit institutions supervised by the ECB with at least one approved internal model for credit risk are asked to provide based on specific validation tests and analyses (also referred to as "validation tools"). Those tests and analyses are grouped together by risk parameter: PD, LGD, expected loss best estimate ($EL_{BE}$), LGD for defaulted assets (LGD in-default), the CCF and the slotting criteria for specialised lending exposures. For each risk parameter, the validation tools are linked to individual areas of investigation.

The sections detailing the various validation tools are all structured in the same way. First, the objective of the analysis is described, followed by a description of the tool. Credit institutions are then provided with detailed guidance on implementing the tool and the scope of its application. Lastly, institutions are given details of how to report the results of these tests and analyses.

---

[8] For details of the main requirements, see Article 185 of the CRR.

The ECB is conscious of the limitations of statistical tests and the importance of the underlying assumptions, particularly in the case of small samples. Those limitations will be taken into account when interpreting the results.

This chapter is organised as follows: Section 2.1 describes the scope of the various validation tools; Section 2.2 contains reporting instructions; Section 2.3 contains definitions to be applied to all credit risk validation tools; Section 2.4 describes the general model and the portfolio information that is to be provided; Sections 2.5 to 2.9 describe the validation tools for PD, LGD, $EL_{BE}$, LGD in-default and the CCF; and Section 2.10 deals with validation tools aimed at assessing the reliability of slot assignment of specialised lending exposures based on the slotting approach.

## 2.1 Scope of application

Unless stated otherwise, the validation tools described in this section should be applied separately to each of the rating systems[9] approved by the CAs at the start of the relevant observation period for reporting (see Section 2.2). Analysis should include all facilities or customers that fall within the range of application[10] of the rating system at that point in time.

Rating systems approved for the calculation of own funds requirements for equity exposures (with the exception of the PD/LGD method) and securitisation positions are excluded from the scope of this supplementary validation reporting.

Where a rating system comprises different statistical models and other mechanical methods ("models" in this document[11]) for the assignment of final PD, LGD, $EL_{BE}$, LGD in-default, CCF estimates or risk weights/expected loss amounts[12] to facilities or obligors for the purpose of calculating own funds requirements, the analysis is performed at the same level as the credit institution's internal validation[13] (see Figure 1 for an illustrative example). Consequently, it is not necessary to regard each individual calibration segment within a model design as a separate model to which the reporting instructions are to be applied. Where the same model (for any risk parameter) is used across several rating systems, it is not necessary to perform

---

9   "Rating system", as defined by Article 142(1) of the CRR, means "all of the methods, processes, controls, data collection and IT systems that support the assessment of credit risk, the assignment of exposures to rating grades or pools, and the quantification of default and loss estimates that have been developed for a certain type of exposures".

10  Article 143(3) of the CRR states that the range of application of a rating system must "comprise all exposures of the relevant type of exposure for which that rating system was developed".

11  For the purposes of this document, the term "model" is defined using the definition set out in Section 2.4 of the EBA Guidelines on PD estimation, LGD estimation and the treatment of defaulted exposures (EBA/GL/2017/16; published on 20 November 2017). Specifically, a PD model is defined as "all data and methods used as part of a rating system within the meaning of Article 142(1) point (1) of Regulation (EU) No 575/2013, which relate to the differentiation and quantification of own estimates of PD and which are used to assess the default risk for each obligor or exposure covered by that model". Likewise, an LGD model relates to "LGD […] used to assess the level of loss in the case of default for each facility covered by that model". An illustrative depiction of rating systems, models and calibration segments can be found in Figure 1.

12  In the case of rating systems for specialised lending based on the slotting approach.

13  Please note that nothing in this document should be interpreted as requiring institutions to alter the level at which validation techniques are applied in their internal validation.

additional analysis at the level of the aggregated portfolio (across rating systems). In the particular case of LGD, EL$_{BE}$, LGD in-default and CCF models which are applied to more than one rating system, the credit institution may apply the validation reporting only to the model's aggregated portfolio, provided that this is justified to the ECB (see Figure 2 for an illustrative example).

**Figure 1**

Possible structure of a rating system
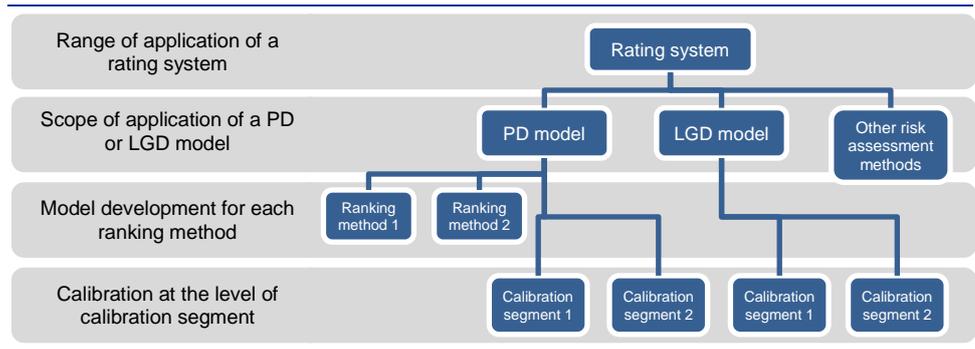


Source: EBA Guidelines on PD estimation, LGD estimation and the treatment of defaulted exposures.

Supplementary validation templates are to be completed at a consolidated level for approved models, unless explicitly stated otherwise in the Decision. However, credit institutions should be able to produce validation templates at an individual level for all of the approved models or rating systems that are used by their various legal entities for the calculation of own funds requirements at an individual level.[14]

Where institutions in a given country are part of a cross-guarantee scheme (pursuant to Article 4(1)(127) of the CRR) and a rating system has been approved at an individual level, the supplementary validation reporting for each member of that cross-guarantee scheme can be replaced by the parent institution's validation reporting at a sub-consolidated level following confirmation by the ECB.

Where a pool model approach[15] is used, the supplementary validation reporting applies to the rating system of each individual participant in the pool.

---

14  For the purposes of this reporting, data can be segmented at the level of the legal entity on the basis of the consolidated portfolio.

15  Article 179(2) of the CRR allows institutions to use data that are pooled across several institutions, provided that the conditions set out in that article are fulfilled.
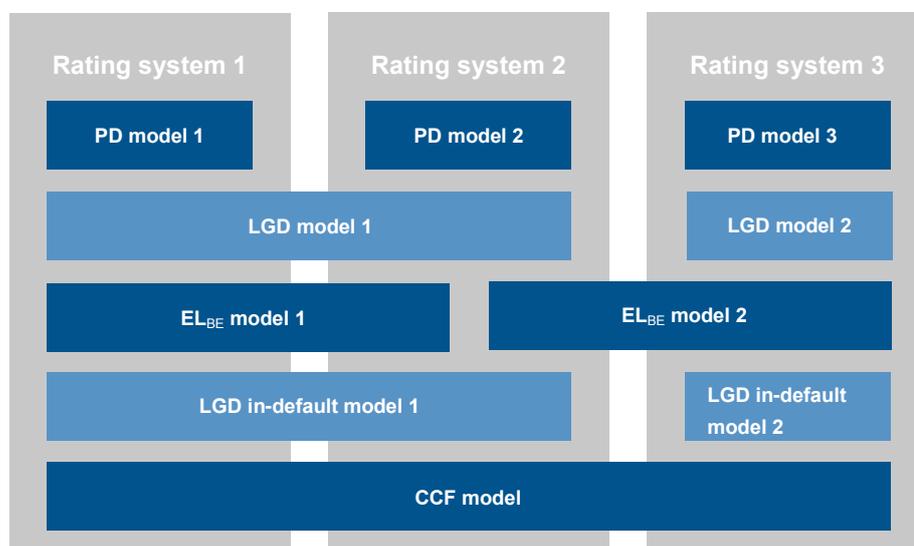
Source: ECB Banking Supervision.

Figure 2 provides an illustrative overview of a generic model landscape. Individual models, each of which comprises ranking methods and calibration segments, may fall within the scope of more than one rating system. This practice may be more common for LGD and CCF models than for PD models. In the same vein, $EL_{BE}$ and LGD in-default models need not necessarily have the same range of application. That is reflected in the structure of this document, which deals with the assessment of such models in separate sections (see Sections 2.7 and 2.8 respectively).

## 2.2    Reporting instructions

In line with the Decision, credit institutions should submit the results of these validation exercises when they send their annual report documenting the outcomes of their internal model validation. Those results should be sent to the ECB using the templates provided, in accordance with the instructions outlined in this document.

The Decision determines also the level (individual, consolidated) at which the validation templates for such reporting shall be completed. Unless explicitly stated otherwise in the Decision, the templates shall be completed at consolidated level.

Validation tools are to be applied on the basis of the one-year observation period[16] used by credit institutions for the internal validation of the relevant rating system or model. If a credit institution does not use a one-year observation period for its internal validation, validation tools are to be applied on the basis of a one-year period ending on the same date as the observation period used by the credit institution. In

---

[16]    See Section 2.3 for a definition of this observation period.

general, the annual end date of the chosen observation period should be kept constant across reporting years (i.e. back-to-back observation periods). If, for any reason, the institution increases the frequency of its validation activities such that its observation period is less than one year in length, the observation period for the validation tools should be adjusted accordingly to ensure that a consistent end date is used.

Credit institutions are invited to use free text fields in the reporting templates to comment on the test results and highlight related analysis documented in their internal validation reports. Institutions are asked to include references to relevant sections of their corresponding internal validation reports for each area of investigation.[17]

The templates to be used for reporting results for the validation tools are attached to this document. The templates should be submitted using the naming convention "[LEICode]_[ModelType]_[ModelID]_[ReferenceDate]_[VersionNumber]", where:

- [LEICode] denotes the LEI code of the institution submitting the template at the highest level of consolidation for which the model is used;

- [ModelType] is "PD", "LGD", "ELBE", "LGDD" (for LGD in-default), "CCF" or "Slotting", depending on the type of model to which the template relates;

- [ModelID] is the unique model identifier agreed by the institution and the ECB;

- [ReferenceDate] is the end of the relevant observation period, as defined in Section 2.3, which should be expressed as DDMMYYYY (e.g. 31122018);

- [VersionNumber] is a number indicating the order of submissions for a given reference date (e.g. 1 for the first submission for 31 December 2018, 2 for the second submission, if needed, and so on).

## 2.3    Definitions

The definitions below apply throughout this document (unless explicitly stated otherwise):

(a) *Relevant observation period*: the uniform one-year period on which all data and information that are needed to perform the validation are based (unless stated otherwise). This period is normally identical to the observation period that the credit institution uses for its internal validation of the relevant model. If the credit institution uses a period of a different length, this will be a one-year period ending on the same date as the institution's observation period (see Section 2.2).

---

17   Such references should be at the level of the area of investigation (e.g. discriminatory power) and should be included regardless of whether the same validation techniques are applied in the institution's internal validation (e.g. the area under the ROC curve (AUC) in the case of discriminatory power).

(b)  *Customers*: non-defaulted obligors (before data exclusions) with positive exposures[18] or committed but undrawn amounts at the beginning of the relevant observation period. For the purposes of this document, a customer may be either an obligor or, if applicable, a facility. Customers and their ratings are considered before substitution effects due to credit risk mitigation. Note that customers and their ratings are considered after rating transfers, since the transfer of a rating (see point (f) of Section 2.5.1) is not a credit risk mitigation technique.

(c)  *Defaults*: customers with at least one default (start) event (according to the institution's internal definition of default) during the relevant observation period (also after migration to a different model or method for determining own funds requirements). This does not, therefore, include technical defaults (see point (d)). Multiple defaults during an observation period are only counted once.

(d)  *Technical defaults*:[19] defaults classified as "technical past due situations" that occur when customers comply with their contractual obligations, but defaults arise owing to data or system errors (including manual errors), a failure of the payment system or processing delays encountered by the credit institution. This definition does not include wrong credit decisions.

(e)  *Different rating model or method for the calculation of regulatory capital requirements*: either (i) a methodology for determining the PD or risk weight (in case of the slotting approach) for a customer that differs from the methodology under consideration (i.e. a different PD model or different slotting method) or (ii) a methodology for determining own funds requirements that differs from the methodology under consideration (i.e. the standardised approach). For instance, a customer that was rated using PD model A at the beginning of the relevant observation period and was rated using PD model B at the end of that observation period is considered to have migrated to a different model. As another example, a customer that was rated using a PD model at the beginning of the relevant observation period and whose own funds requirements at the end of that observation period were determined using the standardised approach is considered to have migrated to a different method of calculating own funds requirements.

(f)  *Portfolio*: the range of application of the rating system or model under investigation (see Section 2.1). The size of the portfolio (in terms of the

---

[18]  If the institution's internal definition of default allows for positive exposures that are not defaultable, (e.g. non-credit obligations), the words "with positive exposures" should be replaced with "with exposures that are defaultable according to the institution's internal default process".

[19]  This definition should be applied as stated. It is not intended to replace, overrule or affect applicable EU or national law.

number of customers[20] for PD models and slotting approach or the number of facilities for LGD, $EL_{BE}$, LGD in-default and CCF models) is denoted by M.

(g) *Original exposure*: the total drawn and undrawn exposure amount pre-CCF and before substitution effects due to credit risk mitigation.

(h) *Initial validation/development*: the observation period for the initial model validation that is relevant to the latest model approval (initial model approval or material model change pursuant to Article 143(3) of the CRR). If that reference period was before the start of this supplementary reporting, it can be replaced with the one-year period immediately preceding the current relevant observation period. Note that where the calculation of a specific test described in the sections below is based on the initial validation/sample, the sample used for the initial validation/development should be analogous to the sample used for the relevant observation period as outlined in the sections below.

(i) *Closed recovery process*: Defaulted facilities are considered to have a closed recovery process if one of the following conditions is met:

  (i) The institution does not expect to implement any further recovery measures in relation to the defaulted customer/facility.

  (ii) The defaulted customer/facility remains in this status for a period of time that exceeds the maximum duration of the recovery process as specified in the institution's internal policies.

  (iii) The defaulted customer/facility has been written-off or fully repaid.

  (iv) The defaulted customer/facility has been reclassified as non-defaulted.

## 2.4    General information

For all credit risk models, general model information, validation information and portfolio information are collected on the basis of all customers or facilities in the portfolio at the end of the relevant observation period.

---

[20] See the definition of "customers" in point (b) above.

### 2.4.1 General model information

Reports for PD, LGD, EL$_{BE}$, LGD in-default, CCF models and slotting approach must include:

- the country code of the institution's group head (two-letter ISO code);

- the LEI code of the institution's group head;

- the name of the institution;

- the unique identifier (model ID) for each model, as agreed between the credit institution and the ECB, and as used in the file name of the template submitted (see Section 2.2);

- the start and end dates of the relevant observation period as defined in Section 2.3.

### 2.4.2 Validation information

Reports for PD, LGD, EL$_{BE}$, LGD in-default, CCF models and slotting approach should include the following information:

- an indication of whether a material model change was implemented within the observation period;

- the internal validation function's overall assessment of the model.

The "validation function's overall assessment" corresponds to the validation function's assessment of the adequacy of the internal model documented in the internal validation report, translated into the following standardised scale:

1. *Adequate with no deficiencies*: No deficiencies detected by the validation function (i.e. no follow-up needed).

2. *Adequate with minor deficiencies*: Minor deficiencies detected that do not lead to any significant bias for risk estimates.

3. *Major deficiencies identified*: Identified deficiencies indicate a significant bias for risk parameter estimates, such as a potential quantitative impact on the risk-weighted exposure amount (RWEA) equal to or above +/- 5% but below +/- 10%.

4. *Severe deficiencies identified*: Identified deficiencies indicate a severe bias for risk parameter estimates, such as a potential quantitative impact on the RWEA equal to or above +/- 10%.

For the specific model types indicated below, reports should also include the following information:

- details of decisions made as regards the model version used (at the beginning or end of the observation period) as defined in the introductions to Sections 2.5 (PD) and 2.9 (CCF);

- an indication of whether the model design allows the downturn component to be separated from the estimated LGD described in Section 2.6.2.1;

- an indication of whether the definition of rating/facility grades/pools changed between the beginning and the end of the observation period (relevant for PD, LGD and CCF models).

In addition to the above, reports for all model types should include the following information for each investigated area:

- the name of the internal validation report in which the relevant validation area is addressed (if that area is part of the validation process);

- a reference to the relevant section of that validation report (if that area is part of the validation process);

- details of the relevant page number(s) in the validation report (if that area is part of the validation process);

- confirmation as to whether the relevant area is part of the validation process;

- the institution's own comments on the area of investigation (optional).

### 2.4.3  Portfolio information

Reports for PD, LGD, EL$_{BE}$, LGD in-default, CCF models and slotting approach should provide the following information on the application portfolio for both the beginning and the end of the relevant observation period:

- RWEA for the portfolio;

- portfolio's exposure at default (EAD), defined as the estimated exposure after CCF (see Article 166 of the CRR; PD, LGD, CCF models and slotting approach only);

- portfolio's exposure value on the relevant reference date in accordance with Article 166 of the CRR (EL$_{BE}$ and LGD in-default models only);

- portfolio's exposure value for defaulted customers in accordance with Article 166 of the CRR (PD and slotting approach only);

- number of customers (PD and slotting approach only);

- number of facilities (LGD, EL$_{BE}$, LGD in-default and CCF models only);

- number of rating grades[21] used by the model for non-defaulted exposures
  (PD models only);

- number of facility grades or pools used by the model (LGD, EL$_{BE}$, LGD
  in-default and CCF models only);

- the type of modelling approach used to segment the model (LGD, CCF,
  EL$_{BE}$, LGD in-default models only); i.e., whether the model is based on
  (i) more than 20 facility grades/pools or it is a continuous model, or (ii) less
  than or equal to 20 facility grades/pools;

- number of defaults (PD and slotting approach only);

- whether the estimates are calculated at the beginning of the observation
  period or one year before default (CCF and LGD models only).

## 2.5 Probability of default

The aim of the validation tools below is to monitor the performance of PD models in
the following areas of investigation:

(a) rating process;

(b) predictive ability (or calibration);

(c) discriminatory power (or rank-ordering performance);

(d) stability.

The validation tools are applied at rating grade[21] level, at portfolio level or both, as
specified in the relevant sections below. The individual rating grades should be those
used to calculate own funds requirements for non-defaulted exposures. Institutions
which have more rating grades than the reporting templates allow or which use
models with continuous PD should use the rating scale employed by the institution
for validation and reporting purposes. In such cases, the PD derived from the model
is mapped to that rating scale and the results for predictive ability and stability are
reported on the basis of number-weighted average PD per grade. If a rating scale is
used under these conditions, this is to be indicated in the relevant comment field for
that area of investigation.

Supplementing the scope of application set out in Section 2.1, the institution has the
option to report all validation tools for PD models using a model that is in production
at the end of the relevant observation period, but only if:

- the model has undergone a material change that was approved by the
  competent authorities during the relevant observation period; and

---

21  For the reporting of PD models, the instructions uniformly refer to rating grades. In case an institution
estimates PD by pool in accordance with Article 180(2)(a) of the CRR, the term "rating grade" should
be replaced by "pool".

- the institution's validation report is also based on that model in production at the end of the relevant observation period.

### 2.5.1 Specific definitions

The definitions below apply to the whole of Section 2.5:

(a) *PD*: the (final) PD used for the calculation of own funds requirements (including any regulatory floors, add-ons, appropriate adjustments, margin of conservatism, overrides and mapping to master scales). Please refer to the definition of "customers" in Section 2.3 for details of the treatment of credit risk mitigation.

(b) *Customers with missing ratings or ratings forced to default values*: customers that did not have a valid rating (i.e. the rating process had not been concluded and finalised) at the start of the relevant observation period but fall within the range of application of the model under consideration, and customers with PD estimates assigned to predefined conservative ratings owing to a lack of data on any model risk driver. This term does not, however, include customers whose ratings are based, in part, on missing information.

(c) *Data exclusions due to process deficiencies*: customers that are not part of the sample owing to process deficiencies. This includes customers that should have been rated using the rating model under consideration at the beginning of the relevant observation period, but were not (including, among others, customers with missing ratings as defined in point (b)) and customers that were rated using the rating model under consideration but were excluded from the sample for process-related reasons (e.g. incorrect segmentation).

(d) *Customers with outdated ratings or financial statements*: in order to reduce effects arising from a specific (conservative) treatment of outdated information, institutions must, for the purposes of this exercise, exclude the following customers from the validation sample:

  (i) *Customers with outdated ratings*: customers with ratings assigned more than 12 months before the start of the relevant observation period.

  (ii) *Customers with ratings based on outdated financial statements*: customers with ratings that are based on financial statements with reference dates more than 24 months before the start of the relevant observation period.[22]

---

[22] Note that analysis of this process deficiency is performed if financial statements are used as inputs (risk factors) for the PD model. This deficiency is expected to apply only to parts of non-retail rating systems.

(e) *Customers with overrides*: customers with manual adjustments to the ratings proposed by the model that are based on subjective criteria.[23] These include, in particular, rating changes or changes to the model's input that are executed manually owing to subjective criteria that are not covered satisfactorily by the model. In contrast, a rating transfer (as defined in point (f) below) is not considered to be an override.

(f) *Customers with transferred ratings*: customers to which a rating of a third party[24] has been transferred because of the existence of an appropriate guarantee which ensures that there is no difference in risk between the customer and the related party. This does not include ratings which take account of the rating of the parent company as a risk factor as part of the model, or where parent ratings are taken into account as an indication of the overriding of a customer's rating (see point (e) above).

(g) *Number of customers (*N*)*: the number of customers[25] in the portfolio after data exclusions due to process deficiencies and excluding customers with outdated ratings or financial statements and customers with transferred ratings (see points (a), (c), (d) and (f) above).

(h) *Migration matrix*: If $K$ is the number of rating grades used by the model for non-defaulted exposures, the migration matrix is a matrix with $K$ rows and $K + 3$ columns showing the frequency with which customers migrate from rating grades $1$ to $K$ (based on the number of customers $N$) to one of the following statuses in the course of the relevant observation period:[26]

   (i) rating grades $1$ to $K$ (based on the number of customers $N$);

   (ii) default grade, including defaulted customers that have left the rating system or model during the observation period;

   (iii) non-defaulted customers for which a different rating system, model or method is used to determine own funds requirements;[27]

   (iv) non-defaulted customers that have terminated their business relationship with the credit institution during the relevant observation period. This category contains all remaining customers that are not included in the other categories defined above.

---

[23] "Non-automated manual adjustments to the ratings" refer to situations in which human judgement may override inputs or outputs in the grade and pool assignment process in accordance with Article 172(3) of the CRR.

[24] Whereas the rating of the third party is assigned internally in accordance with the rating system for which the institution has received permission.

[25] See the definitions of "customers" and "portfolio" in Section 2.3.

[26] Note that the status of the customer in between the start and end of the observation period is not reflected in the migration matrix.

[27] See the definition in point (e) of Section 2.3.

In addition to the specific definitions above, the definitions contained in points (a) to (h) of Section 2.3 also apply.

### 2.5.2 Qualitative validation tools

The analysis of qualitative aspects of PD models is aimed at ensuring the appropriateness of the rating process and the default recognition process.

Results should be reported for the following validation tools:[28]

1.  The size of the portfolio ($M$) (see point (f) of Section 2.3).

2.  Rating process statistics regarding the occurrence of:

    (a)  outdated ratings or financial statements (see point (d) of Section 2.5.1);

    (b)  transferred ratings, if not included in (a) above (see point (f) of Section 2.5.1);

    (c)  data exclusions due to process deficiencies in accordance with the definitions in point (c) of Section 2.5.1.

3.  The average PD and number of defaults that occurred during the relevant observation period among customers excluded under point 2(a) and (b) above.

4.  The number of customers ($N$) (see point (g) of Section 2.5.1).

5.  Qualitative measures:

    (a)  occurrence of overrides, on the basis of the definition in point (e) of Section 2.5.1;

    (b)  occurrence of technical defaults (see point (d) of Section 2.3).

All of the summary statistics in this section should be computed on the basis of the portfolio's composition at the beginning of the observation period.

### 2.5.2.1 Rating process statistics

Objectives of the tool

The objective of this validation tool is to verify the appropriateness of important aspects of the rating process applied by the credit institution when assigning obligors to a specific rating grade or PD level.

Description

The statistics considered here concern the relative frequency of rating process deficiencies and transferred ratings as defined in Section 2.5.1.

---

[28]  See the relevant definitions in Section 2.5.1. Note that the rating process statistics mentioned in the list are exclusions from basic set $M$ and that the number of clients in the validation sample ($N$) is derived from basic set $M$ by subtracting all data exclusions listed under "rating process statistics".

Calculate the $M_{ex}/M$ summary statistic, where M denotes the number of customers before exclusions as defined in point (f) of Section 2.3 and $M_{ex}$ denotes the number of customers with a rating process deficiency, with a transferred rating or with outdated ratings or financial statements respectively at the beginning of the relevant observation period. See Figure 3 for an illustration of the different subsets for the calculation of the $M_{ex}/M$ summary statistic.

**Figure 3**

Illustration of the data basis for qualitative rating process statistics and the resulting back-testing sample
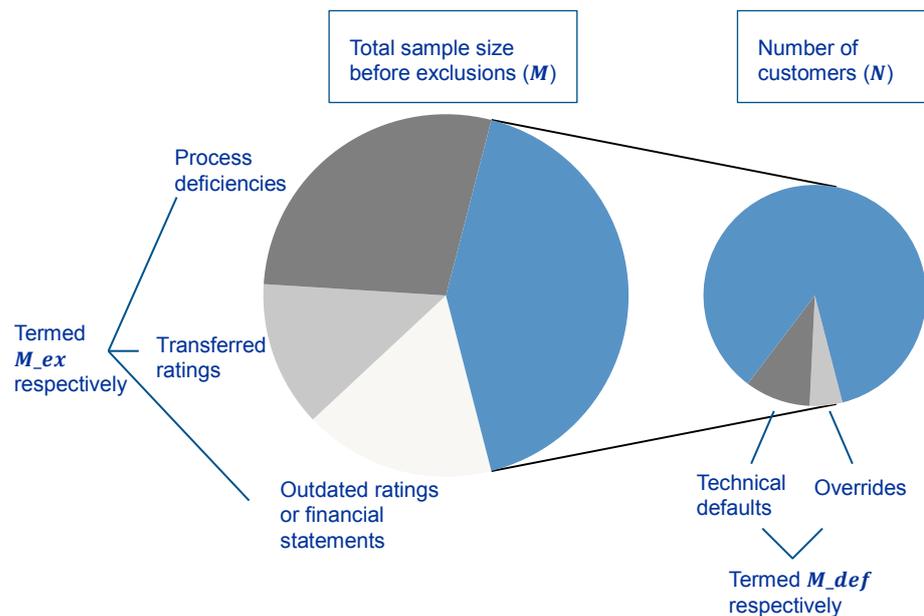


Source: ECB Banking Supervision.

Summary statistics are calculated at portfolio level for the following process deficiencies:

1.  Customers with outdated ratings or financial statements[29] (see point 2(a) in Section 2.5.2);

2.  Customers with transferred ratings[30] (see point 2(b) in Section 2.5.2);

3.  Data exclusions due to process deficiencies (see point 2(c) in Section 2.5.2).

The report should include results for:

-   the size of the portfolio (M) – i.e. before data exclusions;

---

[29]   Note that analysis of this process deficiency is performed if financial statements are used as inputs (risk factors) for the PD model. This deficiency is expected to apply only to parts of non-retail rating systems.

[30]   Care must be taken to ensure that all customers with transferred ratings at the beginning of the relevant observation period which have defaulted during the observation period are taken into account (even in the event of the disbanding of the group structure).

- the number of customers with outdated ratings or financial statements ($M_{ex}$) and their PD at the beginning of the relevant observation period, the number of defaults among those customers that occurred during the relevant observation period, and an indication whether or not the model design allows for outdated ratings or financial statements;

- the number of customers which have a transferred rating ($M_{ex}$) and their PD at the beginning of the relevant observation period, the number of defaults among those customers that occurred during the relevant observation period, and an indication whether or not the assignment process allows for rating transfer;

- the number of customers which are excluded from the validation sample owing to a rating process deficiency ($M_{ex}$);

- the number of customers in the validation sample ($N$) – i.e. after data exclusions.


## 2.5.2.2    Occurrence of overrides

The objective of this validation tool is to verify PD models' appropriateness by analysing the occurrence of overrides.

The statistic considered here is the relative frequency of customers with overrides.

Calculate the $M_{def}/N$ summary statistic, where *N* denotes the number of customers as defined in point (g) of Section 2.5.1 and $M_{def}$ denotes the number of customers with overrides in $N$.[31]

The summary statistic is calculated at portfolio level.

The number of customers with overrides ($M_{def}$) and an indication whether or not the assignment process allows for overrides are reported.


## 2.5.2.3    Occurrence of technical defaults

The objective of this validation tool is to verify the appropriateness of default recognition by analysing the occurrence of technical defaults.

The statistic considered here is the relative frequency of technical defaults, as defined in point (d) of Section 2.3.

Calculate the $M_{def}/N$ summary statistic, where $N$ denotes the number of customers as defined in point (g) of Section 2.5.1 and $M_{def}$ denotes the number of technical defaults.[32]

---

[31]    Please note that ratings with overrides are not excluded from the sample (see Section 2.5.1).

The summary statistic is calculated at portfolio level.

The number of customers with technical defaults ($M\_def$) and an indication whether or not such defaults are assessed by the institution are reported.

### 2.5.3 Predictive ability

The analysis of predictive ability (or calibration) is aimed at ensuring that the PD parameter adequately predicts the occurrence of defaults – i.e. that PD estimates constitute reliable forecasts of default rates.

The input parameters and the results of the Jeffreys test are reported, while the data basis is the number of customers ($N$), as defined in point (g) of Section 2.5.1.

#### 2.5.3.1 PD back-testing using a Jeffreys test

The objective of the Jeffreys test[33] is to assess the predictive ability of PD estimates at the level of individual rating grades and at portfolio level.

The Jeffreys test compares forecasted defaults with observed defaults in a binomial model with independent observations under the null hypothesis that the PD applied in the portfolio/rating grade at the beginning of the relevant observation period is greater than the true one (one-sided hypothesis test). The test statistic is the PD of the portfolio/rating grade. Given the Jeffreys prior for the binomial proportion, the posterior distribution is a beta distribution with shape parameters $a = D + 1/2$ and $b = N - D + 1/2$. Here, $N$ is the number of customers in the portfolio/rating grade and $D$ is the number of those customers that have defaulted within that observation period. The p-value (i.e. the cumulative distribution function of the aforementioned beta distribution evaluated at the PD of the portfolio/rating grade) serves as a measure of the adequacy of estimated PD.

Calculate the p-value $\beta_{D+\frac{1}{2},N-D+\frac{1}{2}}(PD)$, where $\beta_{D+\frac{1}{2},N-D+\frac{1}{2}}$ is the distribution function of the beta distribution with shape parameters $a = D + 1/2$ and $b = N - D + 1/2$, for the portfolio and each rating grade.

The test is performed

1.   at portfolio level and

2.   at the level of all individual rating grades.

---

[32]   Please note that the number of technical defaults is based on the number of customers ($N$), as defined in Section 2.5.1. For example, a technical past due situation concerning a customer with an outdated rating does not contribute to the ratio, as the customer was excluded.

[33]   See Brown, L., Cai, T. and DasGupta, A. (2001), "Interval Estimation for a Binomial Proportion", *Statistical Science*, Vol. 16(2), pp. 101-117.

The test is based on the number of customers ($N$) as defined in point (g) of Section 2.5.1. The relevant PDs are the PDs for the individual rating grades for non-defaulted customers at the beginning of the relevant observation period. The number of tests to be conducted is equal to the number of rating grades for non-defaulted exposures plus one (for the overall portfolio).

The report should include:

- the name of the rating grade;

- PD at the beginning of the relevant observation period;

- the number of customers ($N$);

- the number of defaulted customers ($D$);

- the p-value $\beta_{D+\frac{1}{2},N-D+\frac{1}{2}}(PD)$, where $\beta_{D+\frac{1}{2},N-D+\frac{1}{2}}$ is the distribution function of the beta distribution with shape parameters $a = D + 1/2$ and $b = N - D + 1/2$;

- the original exposure[34] at the beginning of the relevant observation period.

## 2.5.4    Discriminatory power

The analysis of discriminatory power is aimed at ensuring that the ranking of customers that results from the rating process appropriately separates riskier and less risky customers. There are several equivalent measures of discriminatory power, which can be defined in a number of ways. In order to make definitions as precise as possible and calculations as simple as possible, the measure used in this section is the AUC, which is defined and calculated in terms of the Mann-Whitney U statistic (see annex, Section 3.1).

The validation tool in this section is based on a comparison between the discriminatory power of a PD model (current AUC) and the discriminatory power achieved by the same PD model at the time of the initial validation during development.

The AUC, as referred to in this section, should be computed as follows. In portfolios where final PD is obtained by mapping a continuous PD or score to a PD scale with discrete rating grades, those rating grades should be used for calculating the AUC. Where final PD is itself continuous, the AUC should be calculated by mapping PD to a relevant master scale (used for validation and reporting purposes) and using those rating grades to obtain the AUC. Thus, the scale used for the analysis of discriminatory power is identical to the scale used for assessing predictive ability (see Section 2.5.3).

---

[34]    The original exposure is employed in the calculation of the concentration measures described in Section 2.5.5.3.

### 2.5.4.1 Current AUC vs. AUC at initial validation/development

**Objectives of the tool**

The current discriminatory power is benchmarked against the discriminatory power measured at the time of the initial validation in the course of the model's development.

**Description**

The AUC for the relevant observation period is compared with the AUC at the time of the initial validation[35] during development via hypothesis testing based on a normal approximation, assuming a deterministic AUC at the time of development. The null hypothesis of the test is that the AUC at the time of development is smaller than the AUC for the relevant observation period.

**Scope**

The test is applied at portfolio level on the basis of the number of customers (N) as defined in Section 2.5.1:

1.  for the relevant observation period vs. the time of the initial validation; or

2.  (if deemed necessary[36]) for the aggregation (on the basis of the method described in the annex, Section 3.1) of the relevant observation period and the two preceding one-year observation periods vs. the aggregation of the relevant (one-year) observation period at the time of the initial validation and the two preceding one-year observation periods (i.e. three one-year periods).

The preparation of data should be consistent across all of the observation periods considered.

**Implementation**

Calculate the test statistic:

$$S = \frac{\text{AUC}_{\text{init}} - \text{AUC}_{\text{curr}}}{s},$$

where $\text{AUC}_{\text{init}}$ denotes the AUC at the time of development, $\text{AUC}_{\text{curr}}$ denotes the AUC for the relevant observation period and $s = \sqrt{s^2}$ denotes the estimated standard deviation of $\text{AUC}_{\text{curr}}$ (see annex, Section 3.1).

If the AUC at the time of development is based on a sample comprising multiple observation periods, $\text{AUC}_{\text{init}}$ is calculated in accordance with the aggregation method described in the annex, Section 3.1.

The data basis for this test consists of all customers (N) as defined in Section 2.5.1.

**Reporting of the results**

The report should include:

*   values for $\text{AUC}_{\text{init}}$, $\text{AUC}_{\text{curr}}$ and the estimated variance ($s^2$);

*   values for the test statistic $S$ and the p-value $1 - \Phi(S)$, where $\Phi$ denotes the distribution function of the standard normal distribution;

---

[35] See the definition of "initial validation/development" in point (h) of Section 2.3.

[36] For portfolios with a low number of defaults in the relevant observation period, institutions may decide to apply the test over an aggregated three-year period in order to achieve more robust results.

- details of whether calculations are based on multiple (i.e. three) observation periods;

- details of the sample used to calculate the initial AUC: the time period of the validation sample (start date and end date), the number of customers and the variance of the AUC in the validation sample.


### 2.5.5 Stability

The analyses in this section provide insight with regard to the stability of rating model outputs over the one-year observation period.

The stability of PD estimates is assessed using the following validation tools:

1. customer migrations;

2. stability of the migration matrix;

3. concentration in rating grades.


### 2.5.5.1 Customer migrations

Objectives of the tool

The objective of this validation tool is to analyse the migration of customers across rating grades during the relevant observation period.[37]

Description

A migration matrix is used to calculate the relative frequency of customer migrations over a certain number of rating grades. Two statistics ($MWB$) are calculated in order to summarise upgrades and downgrades, i.e. respective values above and below the diagonal of the migration matrix.

Implementation

The migration matrix shows the frequency with which customers migrate from one status $1$ to $K$ to another status $1$ to $K$ in the course of the relevant observation period (where $K$ is the number of rating grades for non-defaulted exposures). The data basis is the number of customers ($N$) as defined in points (g) and (h)(i) of Section 2.5.1.

The summary statistics to compute are as follows:

For customers in rating grades $1$ to $K$ at the beginning of the relevant observation period, the matrix weighted bandwidth ($MWB$) metric for upper off-diagonal transitions (column $j$ > row $i$) and lower off-diagonal transitions (column $j$ < row $i$) respectively, where:

- upper $MWB = \left(\frac{1}{M^{norm,u}}\right) \sum_{i=1}^{K-1} \sum_{j=i+1}^{K} |i - j| \cdot N_i \cdot p_{ij}$;[38]

---

[37] Note that the status of the customer in between the start and end of the observation period is not reflected in the migration matrix.

- lower $MWB = \left(\frac{1}{M^{norm,l}}\right) \sum_{i=2}^{K} \sum_{j=1}^{i-1} |i-j| \cdot N_i \cdot p_{ij}$;

- $M^{norm,u} = \sum_{i=1}^{K-1} [max(|i-K|, |i-1|) \cdot (N_i \sum_{j=i+1}^{K} p_{ij})]$;

- $M^{norm,l} = \sum_{i=2}^{K} [max(|i-K|, |i-1|) \cdot (N_i \sum_{j=1}^{i-1} p_{ij})]$;

- $N_i$ is the number of customers in rating class $i$ at the beginning of the observation period;

- $p_{ij}$ is the relative frequency of transitions between rating classes $i$ and $j$ – i.e. $p_{ij} = N_{ij}/N_i \; for \; N_i > 0$;

- $N_{ij}$ is the number of customers that are in rating class $i$ at the beginning of the relevant observation period and in rating class $j$ at the end of that observation period.

<table>
<tr><td>Scope</td><td>The statistics are calculated at portfolio level.</td></tr>
<tr><td>Reporting of the results</td><td>The report should include results for:</td></tr>
</table>

- the $MWB$ metric for upper off-diagonal transitions (upper $MWB$) and lower off-diagonal transitions (lower $MWB$).

## 2.5.5.2 Stability of the migration matrix

**Objectives of the tool**

The objective of this validation tool is to verify the monotonicity of off-diagonal transition frequencies in the migration matrix by means of z-tests, thereby identifying possible portfolio shifts.

**Description**

Consider the entries in the migration matrix corresponding to the status "rating grades 1 to K" at the beginning and at the end of the relevant observation period on the basis of the number of customers (N) as defined in points (g) and (h)(i) of Section 2.5.1. The fact that rating migrations follow a multinomial distribution can be exploited by pairwise z-tests exploiting the asymptotic normality of the test statistic. Let $p_{i,j}$ denote the (observed) relative frequency of transition (i.e. the relative frequency of customer migrations) between rating grade $i$ (at the beginning of the relevant observation period) and rating grade $j$ (at the end of that observation period). The null hypothesis of the tests is either $H_0: p_{i,j} \geq p_{i,j-1}$ or $H_0: p_{i,j-1} \geq p_{i,j}$ depending on whether the $(i,j)$ entry in the migration matrix is below or above the main diagonal.

To complement the transition frequencies per rating grade, calculate the relative frequency of customers migrating to default grade (as defined in point (h)(ii) of Section 2.5.1), to a different rating system, model or method (as defined in point (h)(iii) of Section 2.5.1) and the relative frequency of customers where the business

---

[38] Note that the term $N_i \cdot p_{ij}$ corresponds to the number of customers migrating from rating grade $i$ to rating grade $j$ – i.e. $N_i \cdot p_{ij}$ is an integer.

relationship has been terminated during the relevant observation period (as defined in point (h)(iv) of Section 2.5.1).

For each rating grade $i = 1, \ldots K$, the following test statistics are to be computed:

$$z_{i,j} = \frac{p_{i,j+1} - p_{i,j}}{\sqrt{\dfrac{p_{i,j}(1 - p_{i,j})}{N_i} + \dfrac{p_{i,j+1}(1 - p_{i,j+1})}{N_i} + 2\dfrac{p_{i,j}p_{i,j+1}}{N_i}}}$$

for $1 \leq j < i$ (lower off-diagonal) and

$$z_{i,j} = \frac{p_{i,j-1} - p_{i,j}}{\sqrt{\dfrac{p_{i,j}(1 - p_{i,j})}{N_i} + \dfrac{p_{i,j-1}(1 - p_{i,j-1})}{N_i} + 2\dfrac{p_{i,j}p_{i,j-1}}{N_i}}}$$

for $K \geq j > i$ (upper off-diagonal). Here, the following definitions apply:

- $p_{i,j}$ denotes the relative frequency of transition between rating grade $i$ and rating grade $j$ (as described above);

- $N_i$ is the number of customers in rating grade $i$ at the beginning of the relevant observation period, as defined in Section 2.5.1.

As the test statistic is asymptotically normal, the p-value which is used for the assessment is calculated as the cumulative distribution function of the standard normal distribution evaluated using $z_{i,j}$ – i.e. $\Phi(z_{i,j})$, where $\Phi$ denotes the cumulative distribution function of the standard normal distribution.

If the test statistic $z_{i,j}$ is not well defined for given frequencies of transition $p_{i,j}$ or $N_i$, the values for $z_{i,j}$ and $\Phi(z_{i,j})$ are reported as "missing" in the reporting template (i.e. the respective fields are left empty).

The test is performed for each combination of rating grades $1$ to $K$ at the beginning and end of the relevant observation period.

The report should include:

- results for $p_{i,j}$ – the relative frequency of transition between rating grade $i$ (at the beginning of the relevant observation period) and rating grade $j$ (at the end of that observation period);

- results for $p_{i,j}$ – the relative frequency of transition between rating grade $i$ (at the beginning of the relevant observation period) and each of the classes defined in points (h)(ii) to (h)(iv) in Section 2.5.1 (at the end of that observation period);

- for $i \neq j$, values for the test statistic $z_{i,j}$ and the p-value $\Phi(z_{i,j})$, where $\Phi$ denotes the cumulative distribution function of the standard normal distribution.

### 2.5.5.3 Concentration in rating grades

The objective of this validation tool is to assess whether rating grades have meaningful dispersion – i.e. to benchmark the current concentration level against the concentration level measured at the time of the initial validation[39] in the course of the model's development. The level of concentration is calculated both in terms of the percentage of customers and in terms of exposure.

Comparison of the Herfindahl Index at the beginning of the relevant observation period and the Herfindahl Index at the time of the initial validation during development via hypothesis testing based on a normal approximation assuming a deterministic Herfindahl Index at the time of the model's development. The null hypothesis of the test[40] is that the current Herfindahl Index is lower than the Herfindahl Index at the time of development.

Calculate the coefficient of variation and the Herfindahl Index as:

$$CV_{curr} = \sqrt{K \sum_{i=1}^{K} \left(R_i - \frac{1}{K}\right)^2},$$

$$HI_{curr} = 1 + \log\left(\frac{CV_{curr}^2 + 1}{K}\right) / \log(K),$$

where:

- $K$ is the number of rating grades for non-defaulted exposures;

- $R_i$ is the relative frequency of rating grade $i$ at the beginning of the relevant observation period.

Calculate the p-value $1 - \Phi\left(\sqrt{K-1}(CV_{curr} - CV_{init})/\sqrt{CV_{curr}^2(0.5 + CV_{curr}^2)}\right)$, where $\Phi$ denotes the cumulative distribution function of the standard normal distribution and $CV_{init}$ is the coefficient of variation at the time of the initial validation.

The test is performed on the number of customers ($N$) as defined in Section 2.5.1 at the following levels:

- number-weighted: in this case, $R_i = N_i / \sum_{j=1}^{K} N_j$, where $N_i$ is the number of customers in rating grade $i$;

- exposure-weighted (only calculation of $HI_{curr}$): in this case, $R_i = E_i / \sum_{j=1}^{K} E_j$, where $E_i$ is the original exposure of customers in rating grade $i$.

---

[39] See the definition of "initial validation/development" in point (h) of Section 2.3.

[40] See Miller, G.E. (1991), "Asymptotic test statistics for coefficients of variation", *Communications in Statistics – Theory and Methods*, Vol. 20(10), pp. 3351-3363.

The report should include:

- the number-weighted $HI_{curr}$ for the current sample and the number-weighted $HI_{init}$ for the initial validation sample, as well as the p-value as described above;

- the exposure-weighted $HI_{curr}$ for the current sample;

- details of the sample used to calculate $CV_{init}$: the time period of the validation sample (start date and end date), the number of customers in the validation sample and the number of rating grades of the model in the initial validation sample.

## 2.6 Loss given default

The aim of the validation tools outlined in this section is to monitor LGD models' performance in the following areas of investigation:

(a) predictive ability (or calibration);

(b) discriminatory power;

(c) qualitative validation tools.

Those validation tools are applied at facility grade or pool level, at portfolio level or both, as specified in the relevant sections below. Institutions which have more facility grades or pools than the reporting templates allow or which use models with continuous LGD should apply the validation tools on the basis of the 12 standardised segments defined by LGD estimates in this section.

### 2.6.1 Specific definitions

The definitions below apply to the whole of Section 2.6:

(a) *Estimated LGD*: the LGD that would have been used to calculate own funds requirements at the beginning of the year in which the default occurred[41] if the LGD model applied at the end of the observation period (with its current scope) had been in force at that time. This means that LGD is calculated using the model in place at the end of the observation period, but on the basis of customer and facility-specific input data (collateral and collateral valuations, risk factors, etc.) relating to the beginning of the year in which the default occurred.[42] Alternatively, LGD

---

[41]  Where a facility did not exist at the beginning of the year in which the default occurred, LGD is calculated using the currently applied LGD model, but on the basis of facility-specific input data (collateral and collateral valuations, risk factors, etc.) relating to the moment when the credit decision in respect of that facility was approved.

[42]  Facilities should not be excluded from the sample. If in doubt, estimates should be produced using appropriate approximations.

one year before the individual default can be estimated, using the model in place at the end of the observation period. Note that, for the purposes of this reporting, the floor specified in Article 164(4) of the CRR is not applied to those LGD estimates – i.e. the floor needs to be extracted from the estimate. When the composition of the portfolio at the beginning or the end of the observation period is analysed (performing exposures only), as is the case with the qualitative tools in Section 2.6.4, estimated LGD refers to assigned LGD at the beginning of the observation period (and, respectively, at the end of the observation period) that is used to calculate own funds requirements (but without the floor specified in Article 164(4) of the CRR as mentioned above).

(b) *Realised LGD*: the realised loss rate (in terms of economic loss as defined in Article 5(2) of the CRR) that is calculated on the basis of the definition (including material direct and indirect costs, discounting, etc.) which is used by the LGD model under investigation. Institutions are expected to be able to calculate realised LGD for all reported facilities, as the recovery process will have been closed (see point (i) of Section 2.3) in the relevant observation period.[43] Furthermore, facilities should not be excluded from the sample. Institutions should apply the multiple-default approach defined in their internal procedure.

(c) *LGD estimates missing or forced to default values*: (i) LGD model estimates that cannot be assigned to facilities within the scope of the model owing to a lack of data on any model risk driver and are forced to take on predefined values, including the use of any kind of fall-back value (e.g. conservative values), and (ii) estimates forced to take on predefined caps or floors defined by the institution.

(d) *Number of facilities (back-testing)*: the number of facilities in the portfolio for which a recovery process has been closed in the course of the relevant observation period.[44] This includes facilities that were in default at the beginning of that observation period, as well as facilities that have defaulted at some point during that observation period. This definition only applies to Sections 2.6.2, 2.6.3 and 2.6.4.4.

(e) *Number of facilities (application portfolio)*: the total number of non-defaulted facilities within the scope of the rating system or model

---

[43] In case historic recovery cash-flows cannot be clearly allocated to single facilities for the calculation of the realised LGD, it is suggested for the purpose of the supplementary reporting that institutions apply the collateral/recovery assignment scheme that is aligned with LGD estimation as part of the calculation of capital requirements.
Institutions should calculate realised LGD at facility level for each default. In exceptional cases where (i) the recovery is performed not at individual facility level, but at a more aggregated level, (ii) this practice is legally enforceable, and is enforced in practice, and (iii) the same practice is applied in the institution's internal validation (e.g. where several facilities of the same or different types are secured using the same collateral), realised LGD can be calculated at a more aggregated level than individual facility level.

[44] See the definition of "closed recovery process" in point (i) of Section 2.3.

under investigation (see Sections 2.1 and 2.3) at a given point in time. This definition only applies to Sections 2.6.4.1 and 2.6.4.2.

(f) *Number of facilities (defaulted)*: the total number of defaulted facilities in the portfolio (i.e. facilities for which the recovery process has not been closed) at a given point in time (beginning or end of the observation period). This definition only applies to Section 2.6.4.3.

(g) *Collateralisation rate*: for the purposes of this reporting, the ratio of the value of collateral and guarantees to the on-balance-sheet amount. This valuation should follow the valuation rules set out for the calculation of regulatory capital and be reported as specified in COREP template C 08.01 (credit and counterparty credit risks and free deliveries: IRB approach to capital requirements), columns 150-210 (section on "credit risk mitigation techniques taken into account in LGD estimates excluding double default treatment")[45], for the relevant reference point.

In addition to the specific definitions above, the definitions contained in points (a) to (h) of Section 2.3 also apply.

## 2.6.2    Predictive ability

The analysis of predictive ability (or calibration) is aimed at ensuring that the LGD parameter adequately predicts the loss rate in the event of a default – i.e. that LGD estimates constitute reliable forecasts of realised loss rates.

The results of LGD back-testing using a t-test should be reported, while the data basis should consist of the number of facilities (back-testing) as defined in point (d) of Section 2.6.1.

---

[45]    See Annex I of Commission Implementing Regulation (EU) No 680/2014 of 16 April 2014.

**Figure 4**

Illustration of the back-testing sample used for LGD models

This figure shows the construction of the sample used to assess the predictive ability of LGD models. It uses the example of two generic facilities (A and B) whose recovery process has been closed in the relevant one-year observation period, indicating the estimated LGD that is to be back-tested in each case.
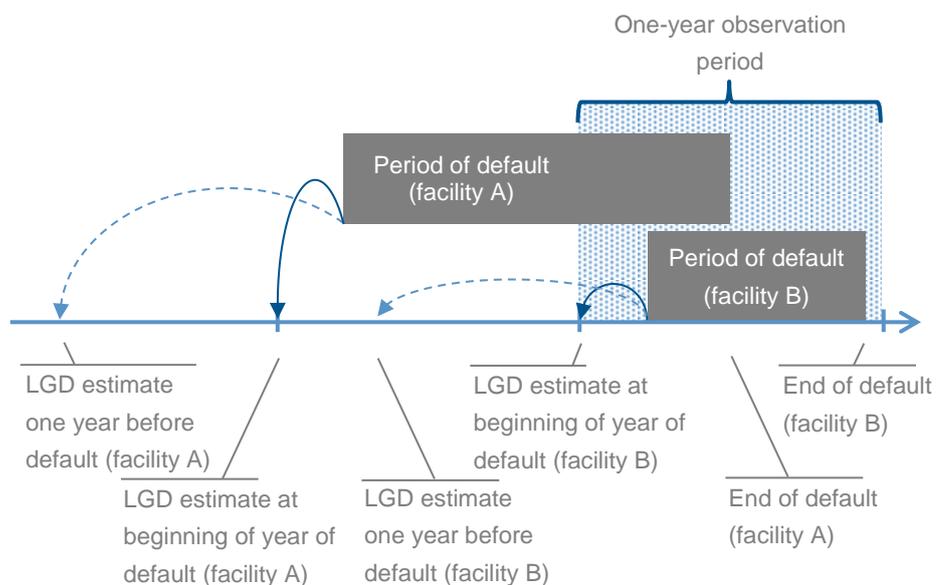


Source: ECB Banking Supervision.

Figure 4 shows the construction of the LGD back-testing sample for two generic facilities. Facilities form part of the back-testing sample if the end of the recovery process falls within the relevant one-year observation period. As outlined in point (a) of Section 2.6.1, the estimated LGD that is to be back-tested for an individual facility relates to either (i) the point in time one year before the individual default or (ii) the beginning of the year in which the default event occurred.

### 2.6.2.1 LGD back-testing using a t-test

Objectives of the tool

The objective of this validation tool (LGD back-testing using a one-sample t-test for paired observations) is to assess the predictive ability of LGD estimates at portfolio level, as well as at grade/pool or segment level.

Description

The one-sample t-test for paired observations compares estimated LGD with realised LGD under the null hypothesis that estimated LGD is greater than true LGD (one-sided hypothesis test) assuming independent observations. Under the null hypothesis, the test statistic is asymptotically Student-t distributed with $(N - 1)$ degrees of freedom, where $N$ denotes the number of facilities (back-testing).

The data basis for the t-test consists of all facilities (back-testing) as defined in point (d) of Section 2.6.1. While the observation period in which the recovery process ends (due to curing, liquidation, etc.) will be the same for each observation, the

length of the recovery process will typically be different for each of the closed defaults (i.e. the defaults might have occurred in different years).

In addition, on the same data basis, estimated LGD without a downturn component is analysed at portfolio level where the model's design allows for the separation of that component.

Calculate the t-test statistic as follows:

$$T = \sqrt{N} \frac{\frac{1}{N}\sum_{i=1}^{N}\left(LGD_i^R - LGD_i^E\right)}{\sqrt{s_{LGD}^2}} \ ,$$

$$s_{LGD}^2 = \frac{\sum_{i=1}^{N}\left(\left(LGD_i^R - LGD_i^E\right) - \frac{1}{N}\sum_{j=1}^{N}\left(LGD_j^R - LGD_j^E\right)\right)^2}{N-1},$$

where:

- $N$ is the number of facilities (back-testing) as defined in point (d) of Section 2.6.1;

- $LGD_i^E$ denotes the estimated LGD for facility $i$ as defined in point (a) of Section 2.6.1;

- $LGD_i^R$ denotes the realised LGD for facility $i$ as defined in point (b) of Section 2.6.1.

Calculate the p-value $1 - S_{N-1}(T)$, where $S_{N-1}$ is the cumulative distribution function of the Student t-distribution evaluated using the test statistic (T) with $(N-1)$ degrees of freedom.

LGD back-testing is to be performed at both of the following levels:

1. portfolio level; and

2. facility grade/pool or segment level.

As regards the second of those, institutions should apply one of the following two approaches:

(a) If the model is based on 20 facility grades/pools or less, the test is to be performed at the facility grade/pool level used in the institution's internal validation.

(b) Otherwise (including in the case of continuous LGD models), the institution should use 12 predefined "LGD segments" on the basis of the following criteria:

   *Segment 1*: facilities i with 0% ≤ $LGD_i^E$ < 5%;

   *Segment 2*: facilities i with 5% ≤ $LGD_i^E$ < 10%;

   *Segment 3*: facilities i with 10% ≤ $LGD_i^E$ < 20%;

...

(*10% LGD steps from Segment 3 to Segment 11*);

…

*Segment 12*: facilities i with $100\% \leq LGD_i^E$.

The report should include:

- the number of facilities (back-testing) as defined in point (d) of Section 2.6.1;

- the number-weighted average of estimated LGD without a downturn component (if available);

- number-weighted averages for estimated and realised LGD;

- a contingency table with frequencies of estimated LGD and realised LGD. If the model is based on more than 20 facility grades/pools or is a continuous model, the contingency table consists of the 12 predefined segments for estimated and realised LGD respectively. If the model is based on 20 facility grades/pools or less, estimated LGD is segmented on the basis of the estimated LGD of the facility grade or pool. Realised LGD is segmented on the basis of the estimated LGD of the facility grades or pools in ascending order;[46]

- the value of the test statistic ($T$), the estimated variance ($s_{LGD}^2$) and p-value $1 - S_{N-1}(T)$ as described above.

### 2.6.3    Discriminatory power

The analysis of discriminatory power is aimed at ensuring that LGD models are able to discriminate between facilities with high and low values for LGD. The measure used in this section to assess the discriminatory power of LGD models is the generalised AUC. That validation tool is based on a generalisation of the classical AUC that can be applied to multi-class problems. More information on the statistics referred to below can be found in the annex, Section 3.2.

Note that, for simplification purposes, the calculation is performed not at the level of individual facilities, but at the level of aggregated segments or grades/pools, as outlined below.

---

[46] As an example, consider a model with only two facility grades: "F1" and "F2". Realised LGD is segmented into three classes, whereby the first class comprises all realised LGD values that are less than or equal to the estimated LGD of grade F1, the second class comprises all realised LGD values that are less than or equal to the estimated LGD of grade F2 which are not part of the first class, and the third class comprises all realised LGD values that are greater than the estimated LGD of grade F2. For more information, see annex, Section 3.2.

## 2.6.3.1 Current gAUC vs. gAUC at initial validation/development

The current discriminatory power of the LGD model is benchmarked against the discriminatory power measured at the time of the initial validation in the course of the model's development.

The generalised AUC (gAUC) for the relevant observation period is compared with the gAUC at the time of the initial validation[47] during development via hypothesis testing based on a normal approximation, assuming a deterministic gAUC at the time of development. The null hypothesis of the test is that the gAUC at the time of development is smaller than the gAUC for the relevant observation period.

The test is applied at portfolio level on the basis of the number of facilities (back-testing) as defined in point (d) of Section 2.6.1, – i.e. the sample consists of all facilities for which the recovery process has been closed within the relevant observation period. While the observation period in which the recovery process ends (due to curing, liquidation, etc.) will be the same for each observation, the length of the recovery process will typically be different for each of the closed defaults (i.e. the defaults might have occurred in different years).

Calculate the test statistic:

$$S = \frac{gAUC_{init} - gAUC_{curr}}{s},$$

where $gAUC_{init}$ denotes the gAUC at the time of the initial validation, $gAUC_{curr}$ denotes the gAUC for the relevant observation period and $s = \sqrt{s^2}$ denotes the estimated standard deviation of $gAUC_{curr}$.

The gAUC is calculated on the basis of facility grades or pools as the ordinal segmentation of LGD. Facility grades or pools are defined in the same way as in the institution's internal validation. If the model is based on more than 20 facility grades/pools or it is a continuous LGD model, the test is performed using 12 predefined "LGD segments" on the basis of the following criteria:

*Segment 1*: facilities i with $0\% \leq LGD_i^E < 5\%$;

*Segment 2*: facilities i with $5\% \leq LGD_i^E < 10\%$;

*Segment 3*: facilities i with $10\% \leq LGD_i^E < 20\%$;

...

(*10% LGD steps from Segment 3 to Segment 11*)

…

*Segment 12*: facilities i with $100\% \leq LGD_i^E$.

---

[47] See the definition of "initial validation/development" in point (h) of Section 2.3.

For more information on the calculation of the gAUC, see the specification in the annex, Section 3.2.

The test is applied at portfolio level.

The report should include:

- values for $\text{AUC}_{\text{init}}$, $\text{AUC}_{\text{curr}}$ and the estimated variance ($s^2$);

- values for the test statistic $S$ and the p-value $1 - \Phi(S)$, where $\Phi$ denotes the cumulative distribution function of the standard normal distribution;

- information on the sample used to calculate the initial gAUC: the time period of the validation sample (start date and end date), the number of facilities and the variance in the validation sample;

- a contingency table for estimated and realised LGD (see Section 2.6.2.1).

## 2.6.4     Qualitative validation tools

The analysis of qualitative aspects of LGD models is aimed at ensuring the appropriateness of LGD's assignment to the application portfolio and the distribution of LGD for facilities under analysis.

The following are reported:

1. LGD assignment process statistics regarding the relative frequency of LGD estimates with missing or forced to default values in the application portfolio (see point (c) of Section 2.6.1);

2. LGD application portfolio distribution by facility grade/pool or by predefined LGD segment;

3. Statistics on the LGD defaulted portfolio (recovery process not closed);

4. Statistics on the LGD defaulted portfolio (recovery process closed within the observation period).

The statistics in Sections 2.6.4.1 and 2.6.4.2 are computed on the basis of the composition of the application portfolio at the beginning and end of the observation period (see point (e) of Section 2.6.1). The statistics in Section 2.6.4.3 are calculated on the basis of the number of defaulted facilities whose recovery process has not been closed in the relevant observation period (see point (f) of Section 2.6.1). The statistics in Section 2.6.4.4 are calculated on the basis of the number of facilities used for the back-testing analysis (see point (d) of Section 2.6.1).

### 2.6.4.1      LGD assignment process statistics

**Objectives of the tool**

The objective of this validation tool is to verify the appropriateness of important aspects of the LGD values assigned to the portfolio within the scope of the model – more specifically, facilities that originally present missing estimates.

**Description**

The statistic considered here is the relative frequency of a specific LGD model deficiency, namely the occurrence of missing estimates or estimates that are forced to take on default values as defined in point (c) of Section 2.6.1.

**Implementation**

Calculate the $M\_miss/M$ summary statistic, where $M$ denotes the number of facilities (application portfolio)[48] and $M\_miss$ denotes the number of facilities with missing LGD estimates as defined in point (c) of Section 2.6.1 at the beginning of the relevant observation period.

**Scope**

Summary statistics are calculated at portfolio level at the beginning of the observation period.

**Reporting of the results**

The report should include results for:

- the number of facilities (application portfolio) ($M$) and the number of facilities with missing LGD values ($M\_miss$) at the beginning of the relevant observation period.


### 2.6.4.2      LGD application portfolio distribution

**Objectives of the tool**

The objective of this validation tool is to analyse the distribution of the application portfolio on the basis of key drivers of LGD, such as the collateralisation rate and original exposure.

**Description**

All statistics defined below should be calculated on the basis of the number of facilities (application portfolio), as defined in point (e) of Section 2.6.1, minus the number of facilities defined in point (c) of Section 2.6.1 (i.e. facilities with LGD estimates missing or forced to default values). The statistics considered here are (i) the number of facilities, (ii) the average estimated LGD,[49] (iii) the average collateralisation rate and (iv) the original exposure. The statistics are reported by facility grade/pool or by predefined LGD segment. If the model is based on 20 facility grades/pools or less, results are reported at the facility grade/pool level used in the institution's internal validation. Otherwise (including in the case of continuous LGD models), results are reported using the 12 predefined LGD segments. The statistics for the facilities are reported at the beginning and end of the relevant observation period.

---

[48]    See point (e) of Section 2.6.1.

[49]    See the definition of "estimated LGD" contained in point (a) of Section 2.6.1. If the definition of facility grades/pools did not change between the beginning and the end of the relevant observation period (as reported in Section 2.4.2), the facility grades/pools should be identical to those in Section 2.6.2.1.

Report at the beginning and end of observation period the number of facilities, the number-weighted average collateralisation rate, the number-weighted average of estimated LGD and the original exposure for each facility grade or pool. Estimated LGD for the facilities is calculated as defined in Section 2.6.1.

In addition, calculate the Population Stability Index (PSI) for estimated LGD, which is defined as:

$$\text{PSI} = \sum_{i=1}^{K} (p_{i,2} - p_{i,1}) \ln\left(\frac{p_{i,2}}{p_{i,1}}\right).$$

Here, $p_{i,j}$ denotes the relative frequency of the observed value $i$ in sample $j$, where $j$ refers to the beginning of the relevant observation period ($j = 1$) and the end of the relevant observation period ($j = 2$) respectively, and where $K$ is the number of facility grades/pools or segments.

These statistics are calculated at facility grade/pool level if the model is based on 20 facility grades/pools or less. Otherwise, they are calculated using the 12 predefined LGD segments for the two samples at the beginning and end of the relevant observation period.

The Population Stability Index is calculated at portfolio level.

The report should include:

- the name of the facility grade/pool or segment;

- the number of facilities (application portfolio) minus the number of facilities with LGD estimates missing or forced to default values, the number-weighted average of estimated LGD, the average collateralisation rate and the original exposure of the application portfolio by facility grade/pool or by predefined LGD segment.[50] Results should be reported both at the beginning and at the end of the relevant observation period;

- the Population Stability Index (which should be calculated on the basis of the number of facilities per facility grade/pool or segment at the beginning and end of the observation period).

### 2.6.4.3 LGD defaulted portfolio (recovery process not closed)

The objective of this validation tool is to analyse important aspects of facilities that have defaulted but whose recovery process has not been closed.

---

[50]   If the model is based on 20 facility grades/pools or less, results are reported at facility grade/pool level on the basis of the grades/pools used in the institution's internal validation. Otherwise (including in the case of continuous LGD models), results are reported using the 12 predefined LGD segments.

| Description | The statistics considered here are the number of facilities, the average time in default, the exposure value[51], $EL_{BE}$ and the RWEA used for own funds requirements. These statistics are measured at the beginning and end of the observation period for facilities that have defaulted but whose recovery process has not been closed. |
|---|---|
| Implementation | Calculate the number of facilities (defaulted), the number-weighted average time in default in days, the exposure value, $EL_{BE}$ and the RWEA for facilities which have defaulted but whose recovery process has not been closed, with each being measured at the beginning and end of the relevant observation period. |
| Scope | Summary statistics are calculated at portfolio level at the beginning and end of the observation period. |
| Reporting of the results | The report should include results for: |

- RWEA, the exposure value, $EL_{BE}$, the number of facilities (defaulted) and the number-weighted average time in default in days – each at the beginning and at the end of the relevant observation period.

## 2.6.4.4 LGD defaulted portfolio (recovery process closed within the observation period)

| Objectives of the tool | The objective of this validation tool is to analyse the average duration of the recovery process for defaulted facilities whose recovery process has been closed within the observation period. |
|---|---|
| Description | The statistic considered here is the average duration of the recovery process in days for all defaulted facilities whose recovery process has been closed within the observation period. |
| Implementation | Calculate the number-weighted average duration of the recovery process in days for defaulted facilities whose recovery process has been closed during the relevant observation period (i.e. the number of facilities (back-testing) as defined in point (d) of Section 2.6.1 and used in Section 2.6.2). |
| Scope | The summary statistic is calculated at portfolio level for facilities whose recovery process has been closed during the relevant observation period. |
| Reporting of the results | The report should include results for: |

- the number-weighted average duration of the recovery process in days for defaulted facilities whose recovery process has been closed during the relevant observation period.

---

[51] The portfolio's exposure value on the relevant reference date in accordance with Article 166 of the CRR.

## 2.7 Expected loss best estimate

The aim of the validation tool outlined in this section is to monitor the performance of $EL_{BE}$ models in the area of predictive ability (or calibration).

The validation tool is applied at facility grade or pool level and at portfolio level, as specified below. Institutions which have more facility grades or pools than the reporting template allows or which use models with continuous $EL_{BE}$ should apply the validation tool on the basis of 12 standardised segments defined by $EL_{BE}$. In general, the sample is equivalent to that used in Section 2.6.

### 2.7.1 Specific definitions

The definitions below apply to the whole of Section 2.7:

(a) *$EL_{BE}$*: the $EL_{BE}$ for defaulted exposures[52] that would have been used to calculate own funds requirements for a given estimation date if the $EL_{BE}$ model applied at the end of the observation period had been in force at that time. This means that $EL_{BE}$ is calculated using the $EL_{BE}$ model in place at the end of the observation period, but on the basis of customer and facility-specific input data (collateral and collateral valuations, risk factors, etc.) relating to the estimation/reference date in question.[53]

(b) *Realised LGD*: the realised loss rate (in terms of economic loss as defined in Article 5(2) of the CRR) that is calculated on the basis of the definition (including material direct and indirect costs, discounting, etc.) which is used by the $EL_{BE}$ model under investigation.[54] Institutions are expected to be able to calculate realised LGD for all reported facilities, as the recovery process will have been closed in the relevant observation period. Furthermore, facilities should not be excluded from the sample. Institutions should apply the multiple-default treatment defined in their internal procedure.

(c) *Number of facilities (back-testing)*: the number of facilities in the portfolio for which a recovery process has been closed in the course of the relevant observation period.[55] This includes facilities that were in default at the

---

[52] As referred to in Article 153(1)(ii), Article 154(1)(i), Article 158(5) and Article 181(1)(h) of the CRR.

[53] Facilities should not be excluded from the sample. If in doubt, estimates should be produced using appropriate approximations.

[54] In case historic recovery cash-flows cannot be clearly allocated to single facilities for the calculation of the realised LGD, it is suggested for the purpose of the supplementary reporting that institutions apply the collateral/recovery assignment scheme that is aligned with the estimation of $EL_{BE}$ as part of the calculation of capital requirements.
Institutions should calculate realised LGD at facility level for each default. In exceptional cases where (i) the recovery is performed not at individual facility level, but at a more aggregated level, (ii) this practice is legally enforceable, and is enforced in practice, and (iii) the same practice is applied in the institution's internal validation (e.g. where several facilities of the same or different types are secured using the same collateral), realised LGD can be calculated at a more aggregated level than individual facility level.

[55] See the definition of "closed recovery process" in point (i) of Section 2.3.

beginning of that observation period, as well as facilities that have defaulted at some point during that observation period.

(d) *Number of facilities (application portfolio)*: the total number of defaulted facilities within the scope of the rating system or model under investigation (see Sections 2.1 and 2.3) at a given point in time. This definition is used only for the reporting of portfolio information (see Section 2.4.3).

(e) *Estimated LGD in-default:* the LGD for defaulted exposures[56] that would have been used to calculate own funds requirements for a given estimation date if the LGD in-default model applied at the end of the observation period had been in force at that time. This means that LGD in-default is calculated using the LGD in-default model in place at the end of the observation period, but on the basis of customer and facility-specific input data (collateral and collateral valuations, risk factors, etc.) relating to the estimation/reference date in question.[57] Estimated LGD in-default is used as a reference value reported for all facilities within the scope of the $EL_{BE}$ model under investigation. If the $EL_{BE}$ and LGD in-default models do not have the same scope (see Figure 2), individual LGD in-default estimates may originate from different LGD in-default models.

In addition to the specific definitions above, the definitions contained in points (a) to (d) and (g) of Section 2.3 also apply.

## 2.7.2 Predictive ability

The analysis of predictive ability (or calibration) is aimed at ensuring that the $EL_{BE}$ parameter adequately predicts the loss rate in the event of a default – i.e. that $EL_{BE}$ values constitute reliable forecasts of realised loss rates.

Institutions report the results of $EL_{BE}$ back-testing using a t-test, whereby the data basis is the number of facilities (back-testing) as defined in point (c) of Section 2.7.1.

---

[56] As referred to in Article 153(1)(ii), Article 154(1)(i) and Article 181(1)(h) of the CRR.

[57] Facilities should not be excluded from the sample. If in doubt, estimates should be produced using appropriate approximations.

**Figure 5**

Illustration of the back-testing sample used for $EL_{BE}$ models

This figure shows the construction of the sample used to assess the predictive ability of $EL_{BE}$ models. It uses the example of two generic facilities (A and B) whose recovery process has been closed in the relevant one-year observation period, indicating the $EL_{BE}$ that is to be back-tested in each case.
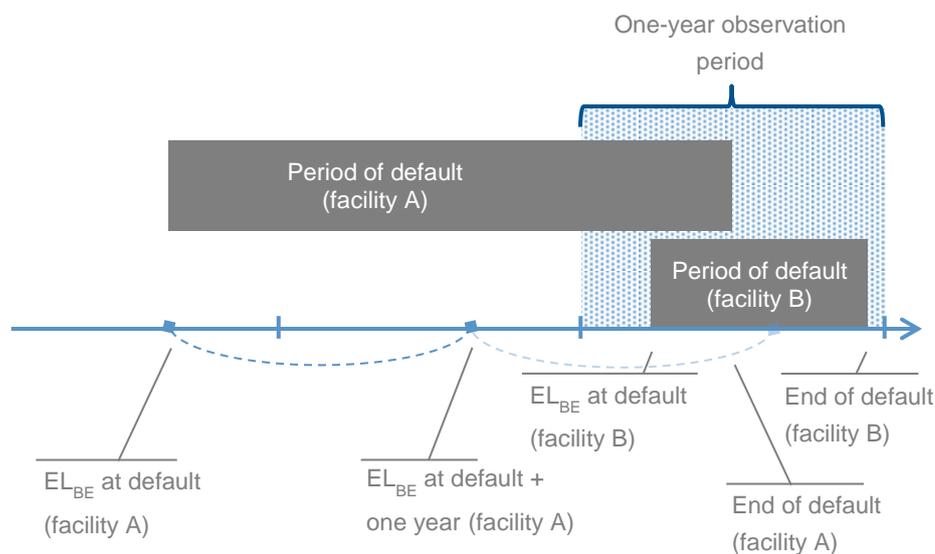


Source: ECB Banking Supervision.

Figure 5 shows the construction of the $EL_{BE}$ back-testing sample for two generic facilities. Facilities form part of the back-testing sample if the end of the recovery process falls within the relevant one-year observation period. As outlined in Section 2.7.2.1, $EL_{BE}$ for an individual facility is back-tested at various points in default: at the time of default, as well as one, three, five and seven years after that default.

### 2.7.2.1 $EL_{BE}$ back-testing using a t-test

*Objectives of the tool*

The objective of this validation tool ($EL_{BE}$ back-testing using a one-sample t-test for paired observations) is to assess the predictive ability of $EL_{BE}$ at portfolio level, as well as at grade/pool or segment level, at various reference points in default.

*Description*

The one-sample t-test for paired observations compares $EL_{BE}$ with realised LGD under the null hypothesis that $EL_{BE}$ is equal to realised LGD (two-sided hypothesis test), assuming independent observations. Under the null hypothesis, the test statistic is asymptotically Student-t distributed with $(N - 1)$ degrees of freedom, where $N$ denotes the number of facilities (back-testing).

The data basis for the t-test consists of all facilities (back-testing) as defined in point (c) of Section 2.7.1. Those estimates are compared with realised values at various reference points in default – i.e. at the time of default, as well as one, three, five and seven years after that default. For both $EL_{BE}$ and realised LGD, all input parameters should relate to the reference point (particularly the exposure value,

which may change owing to payments after default). For realised LGD, only recoveries realised between the reference point after default and the closing of the recovery process are to be taken into account.[58]

Similarly, for all facilities (back-testing) included in the relevant databases for the t-test, estimated LGD in-default is analysed at the $EL_{BE}$ portfolio level, as well as at grade/pool or segment level. Note that the $EL_{BE}$ and LGD in-default models need not necessarily have the same scope (see Figure 2).

Calculate t-test statistics at the time of the individual facility's default, as well as one, three, five and seven years after that default, on the following basis:

$$T = \sqrt{N} \frac{\frac{1}{N}\sum_{i=1}^{N}\left(LGD_i^R - EL_{BE_i}\right)}{\sqrt{s_{EL_{BE}}^2}} \, ,$$

$$s_{EL_{BE}}^2 = \frac{\sum_{i=1}^{N}\left(\left(LGD_i^R - EL_{BE_i}\right) - \frac{1}{N}\sum_{j=1}^{N}\left(LGD_j^R - EL_{BE_j}\right)\right)^2}{N-1},$$

where:

- $N$ is the number of facilities (back-testing) as defined in point (c) of Section 2.7.1;

- $EL_{BE_i}$ denotes the $EL_{BE}$ value for facility $i$;

- $LGD_i^R$ denotes the realised LGD for facility $i$.

Calculate the p-value $2 \cdot (1 - S_{N-1}(|T|))$, where $S_{N-1}$ is the cumulative distribution function of the Student t-distribution evaluated at the test statistic $(T)$ with $(N-1)$ degrees of freedom.

$EL_{BE}$ back-testing is performed at the time of an individual facility's default, as well as one, three, five and seven years after that default, at both of the following levels:

1. portfolio level; and

2. facility grade/pool or segment level.

As regards the second of those, institutions should apply one of the following two approaches:

(a) If the model is based on 20 facility grades/pools or less, the test is to be performed at the facility grade/pool level used in the institution's internal validation.[59]

---

[58] If, for example, a facility in the back-testing sample defaulted two years before the workout was closed, values will be reported at the time of default ("year zero") and one year after the default date ("year one"), but not three, five or seven years after the default date. Any payments between year zero and year one change the exposure value after one year in default. These payments should only be taken into account for realised LGD at year zero – not at year one (as only payments after any reference point should be taken into account).

Otherwise (including in the case of continuous $EL_{BE}$ models), the institution should use 12 predefined "$EL_{BE}$ segments" on the basis of the following criteria:

*Segment 1*: facilities i with $0\% \leq EL_{BE_i} < 5\%$;

*Segment 2*: facilities i with $5\% \leq EL_{BE_i} < 10\%$;

*Segment 3*: facilities i with $10\% \leq EL_{BE_i} < 20\%$;

...

(*10% EL*$_{BE}$ *steps from Segment 3 to Segment 11*);

…

*Segment 12*: facilities i with $100\% \leq EL_{BE_i}$.

**Reporting of the results**

The report should include the name of the facility grade/pool or segment, and results for all reference points, as described above, on:

- the number of facilities (back-testing) as defined in point (c) of Section 2.7.1;

- number-weighted average $EL_{BE}$ and realised LGD;

- the test statistic ($T$), the estimated variance ($s^2_{EL_{BE}}$) and p-value $2 \cdot (1 - S_{N-1}(|T|))$;

- number-weighted average estimated LGD in-default.

## 2.8     LGD in-default

The aim of the validation tool outlined in this section is to monitor the performance of LGD in-default models in the area of predictive ability (or calibration).

That validation tool is applied at facility grade or pool level and at portfolio level, as specified below. Institutions which have more facility grades or pools than the reporting template allows or which use models with continuous LGD in-default should apply the validation tools on the basis of 12 standardised segments defined by LGD in-default estimates. In general, the sample is equivalent to that used in Section 2.6.

---

[59]     Note that this scale must be used consistently across all reference points (i.e. the grades/pools must not change over time, even though they may not necessarily be populated for every point in time by virtue of the model's design).

### 2.8.1 Specific definitions

The definitions below apply to the whole of Section 2.8:

(a) *Estimated LGD in-default*: the LGD in-default[60] that would have been used to calculate own funds requirements for a given estimation date if the LGD in-default model applied at the end of the observation period had been in force at that time. This means that estimated LGD in-default is calculated using the LGD in-default model in place at the end of the observation period, but on the basis of customer and facility-specific input data (collateral and collateral valuations, risk factors, etc.) relating to the estimation/reference date in question.[61]

(b) *Realised LGD*: the realised loss rate (in terms of economic loss as defined in Article 5(2) of the CRR) that is calculated on the basis of the definition (including material direct and indirect costs, discounting, etc.) which is used by the LGD in-default model under investigation.[62] Institutions are expected to be able to calculate realised LGD for all reported facilities, as the recovery process will have been closed in the relevant observation period. Furthermore, facilities should not be excluded from the sample. Institutions should apply the multiple-default treatment defined in their internal procedure.

(c) *Number of facilities (back-testing)*: the number of facilities in the portfolio for which a recovery process has been closed in the course of the relevant observation period.[63] This includes facilities that were in default at the beginning of that observation period, as well as facilities that have defaulted at some point during that observation period.

(d) *Number of facilities (application portfolio)*: the total number of defaulted facilities within the scope of the rating system or model under investigation (see Sections 2.1 and 2.3) at a given point in time. This definition is used only for the reporting of portfolio information (see Section 2.4.3).

(e) *$EL_{BE}$*: the $EL_{BE}$[64] that would have been used to calculate own funds requirements for a given estimation date if the $EL_{BE}$ model applied at the end of the observation period had been in force at that time. This means

---

[60] As referred to in Article 153(1)(ii), Article 154(1)(i) and Article 181(1)(h) of the CRR.

[61] Facilities should not be excluded from the sample. If in doubt, estimates should be produced using appropriate approximations.

[62] In case historic recovery cash-flows cannot be clearly allocated to single facilities for the calculation of the realised LGD, it is suggested for the purpose of the supplementary reporting that institutions apply the collateral/recovery assignment scheme that is aligned with estimating LGD in-default as part of the calculation of capital requirements.
Institutions should calculate realised LGD at facility level for each default. In exceptional cases where (i) the recovery is performed not at individual facility level, but at a more aggregated level, (ii) this practice is legally enforceable, and is enforced in practice, and (iii) the same practice is applied in the institution's internal validation (e.g. where several facilities of the same or different types are secured using the same collateral), realised LGD can be calculated at a more aggregated level than individual facility level.

[63] See the definition of "closed recovery process" in point (i) of Section 2.3.

[64] As referred to in Article 153(1)(ii), Article 154(1)(i), Article 158(5) and Article 181(1)(h) of the CRR.

that $EL_{BE}$ is calculated using the $EL_{BE}$ model in place at the end of the observation period, but on the basis of customer and facility-specific input data (collateral and collateral valuations, risk factors, etc.) relating to the estimation/reference date in question.[65] $EL_{BE}$ is used as a reference value reported for all facilities within the scope of the LGD in-default model under investigation. If the $EL_{BE}$ and LGD in-default models do not have the same scope (see Figure 2), individual $EL_{BE}$ values may originate from different $EL_{BE}$ models.

In addition to the specific definitions above, the definitions contained in points (a) to (d) and (g) of Section 2.3 also apply.

## 2.8.2 Predictive ability

The analysis of predictive ability (or calibration) is aimed at ensuring that the LGD in-default parameter[66] adequately predicts the loss rate in the event of a default – i.e. that LGD in-default estimates constitute reliable forecasts of realised loss rates.

Institutions report the results of LGD in-default back-testing using a t-test, whereby the data basis is the number of facilities (back-testing) as defined in point (c) of Section 2.8.1.

---

[65] Facilities should not be excluded from the sample. If in doubt, estimates should be produced using appropriate approximations.

[66] See Section 2.1.

**Figure 6**

Illustration of the back-testing sample used for LGD in-default models

This figure shows the construction of the sample used to assess the predictive ability of LGD in-default models. It uses the example of two generic facilities (A and B) whose recovery process has been closed in the relevant one-year observation period, indicating the estimated LGD in-default that is to be back-tested in each case.
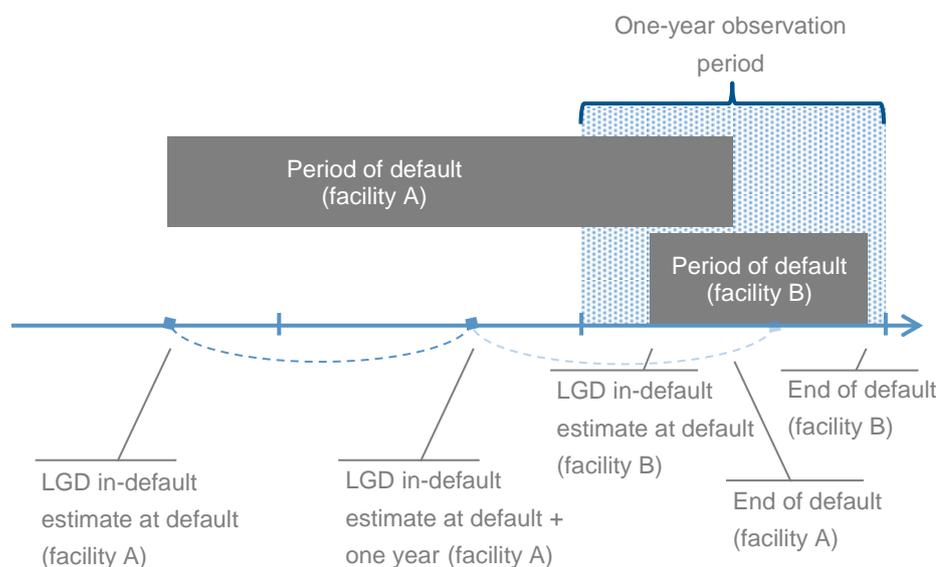


Source: ECB Banking Supervision.

Figure 6 shows the construction of the LGD in-default back-testing sample for two generic facilities. Facilities form part of the back-testing sample if the end of the recovery process falls within the relevant one-year observation period. As outlined in Section 2.8.2.1, estimated LGD in-default for an individual facility is back-tested at various points in time in default: at the time of default, as well as one, three, five and seven years after that default.

### 2.8.2.1 LGD in-default back-testing using a t-test

Objectives of the tool

The objective of this validation tool (LGD in-default back-testing using a one-sample t-test for paired observations) is to assess the predictive ability of LGD in-default estimates at portfolio level, as well as at grade/pool or segment level, at various reference points in default.

Description

The one-sample t-test for paired observations compares estimated LGD in-default with realised LGD under the null hypothesis that estimated LGD in-default is greater than realised LGD (one-sided hypothesis test), assuming independent observations. Under the null hypothesis, the test statistic is asymptotically Student-t distributed with $(N - 1)$ degrees of freedom, where $N$ denotes the number of facilities (back-testing).

The data basis for the t-test consists of all facilities (back-testing) as defined in point (c) of Section 2.8.1. Those estimates are compared with realised values at various reference points in default – i.e. at the time of default, as well as one, three, five and seven years after that default. For both LGD in-default and realised LGD, all

input parameters should relate to the reference point (particularly the exposure value, which may change owing to payments after default). For realised LGD, only recoveries realised between the reference point after default and the closing of the recovery process are to be taken into account.[67]

Similarly, for all facilities (back-testing) included in the relevant databases for the t-test, $EL_{BE}$ is analysed at the LGD in-default portfolio level, as well as at grade/pool or segment level. Note that the $EL_{BE}$ and LGD in-default models need not necessarily have the same scope (see Figure 2).

Implementation

Calculate t-test statistics at the time of the individual facility's default, as well as one, three, five and seven years after that default, on the following basis:

$$T = \sqrt{N} \frac{\frac{1}{N}\sum_{i=1}^{N}(LGD_i^R - LGD_i^D)}{\sqrt{s_{LGD^D}^2}} \,,$$

$$s_{LGD^D}^2 = \frac{\sum_{i=1}^{N}\left(\left(LGD_i^R - LGD_i^D\right) - \frac{1}{N}\sum_{j=1}^{N}\left(LGD_j^R - LGD_j^D\right)\right)^2}{N-1},$$

where:

- $N$ is the number of facilities (back-testing) as defined in point (c) of Section 2.8.1;

- $LGD_i^D$ denotes the estimated LGD in-default for facility $i$;

- $LGD_i^R$ denotes the realised LGD for facility $i$.

Calculate the p-value $(1 - S_{N-1}(\text{T}))$, where $S_{N-1}$ is the cumulative distribution function of the Student t-distribution evaluated at the test statistic (T) with $(N-1)$ degrees of freedom.

Scope

LGD in-default back-testing is performed at the time of the individual facility's default, as well as one, three, five and seven years after that default, at both of the following levels:

1. portfolio level; and

2. facility grade/pool or segment level.

As regards the second of those, institutions should apply one of the following two approaches:

---

[67] If, for example, a facility in the back-testing sample defaulted two years before the workout was closed, values will be reported at the time of default ("year zero") and one year after the default date ("year one"), but not three, five or seven years after the default date. Any payments between year zero and year one change the exposure value after one year in default. These payments should only be taken into account for realised LGD at year zero – not at year one (as only payments after any reference point should be taken into account).

(a) If the model is based on 20 facility grades/pools or less, the test is to be performed at the facility grade/pool level used in the institution's internal validation.[68]

(b) Otherwise (including in the case of continuous LGD in-default models), the institution should use 12 predefined "LGD in-default segments" on the basis of the following criteria:

Segment 1: facilities i with $0\% \leq LGD_i^D < 5\%$;

Segment 2: facilities i with $5\% \leq LGD_i^D < 10\%$;

Segment 3: facilities i with $10\% \leq LGD_i^D < 20\%$;

...

(10% LGD in-default steps from Segment 3 to Segment 11);

…

Segment 12: facilities i with $100\% \leq LGD_i^D$

The report should include the name of the facility grade/pool or segment, and results for all reference points, as described above, on:

- the number of facilities (back-testing) as defined in point (c) of Section 2.8.1;

- number-weighted average estimated LGD in-default and realised LGD;

- the test statistic ($T$), the estimated variance ($s_{LGD^D}^2$) and p-value $1 - S_{N-1}(T)$;

- number-weighted average EL$_{BE}$.

## 2.9    Credit conversion factor

The aim of the validation tools outlined in this section is to monitor CCF models' performance in the following areas of investigation:

(a) CCF assignment process;

(b) predictive ability (or calibration);

(c) discriminatory power (or rank-ordering performance);

(d) qualitative validation tools.

---

[68]   Note that this scale must be used consistently across all points in time in default (i.e. the grades/pools must not change over time, even though they may not necessarily be populated for every point in time by virtue of the model's design).

Supplementing the scope of application set out in Section 2.1, the institution has the option to report all validation tools for CCF models using a model that is in production at the end of the relevant observation period, but only if:

- the model has undergone a material change that was approved by the competent authorities during the relevant observation period; and

- the institution's validation report is also based on that model in production at the end of the relevant observation period.

### 2.9.1 Specific definitions

The definitions below apply to the whole of Section 2.9:

(a) *Estimated CCF*: the CCF that is used to calculate own funds requirements, either at the beginning of the relevant observation period (cohort approach) or 12 months prior to the individual default (fixed-horizon approach[69]).[70] For Section 2.9.5, where the portfolio's composition at the beginning and end of the observation period is analysed (both performing and non-performing exposures), this relates to the assigned CCF that is used to calculate own funds requirements at that point in time.

(b) *Realised CCF*: the realised conversion factor that is calculated on the basis of the definition used by the CCF model (estimated CCF) under investigation.[71]

(c) *Estimated exposure at default*: the estimated amount that is drawn at the time of default before substitution effects due to credit risk mitigation.

(d) *Data exclusions due to process deficiencies*: defaulted facilities that are not part of the sample owing to process deficiencies. This includes facilities that should have been evaluated by the model under consideration at the reference point for estimation, but were not, and facilities that were evaluated by the rating model under consideration, but were excluded from the sample for process-related reasons (e.g. incorrect segmentation).

---

[69] In the case of the fixed-horizon approach, if customers or facilities were not part of the portfolio at that time, the CCF is calculated using the CCF model that is applied at the beginning of the relevant observation period, but on the basis of customer and facility-specific input data (line usage, risk factors, etc.) as they were at the precise moment the facility entered the portfolio.

[70] The decision made in this regard should, if possible, be aligned with the institution's approach to the modelling and validation of the CCF parameter.

[71] Institutions should calculate the realised CCF at facility level for each default. In exceptional cases where (i) the recovery is performed not at individual facility level, but at a more aggregated level, (ii) this practice is legally enforceable, and is enforced in practice, and (iii) the same practice is applied in the institution's internal validation (e.g. where several facilities of the same or different types are secured using the same collateral), the realised CCF can be calculated at a more aggregated level than individual facility level.

(e) *Data exclusions due to outlier treatment for realised CCF*: defaulted facilities that are not part of the sample owing to the exclusion of outliers in the internal validation.[72] Facilities that are affected by replacement of outliers, such as floors, but are included in the internal validation sample, are not covered by this definition.

(f) *Facilities with missing CCF or EAD estimates*: facilities that do not have a CCF or EAD estimate at a given point in time, but fall within the scope of the model under consideration. In particular, this term does not include facilities whose CCF or EAD estimates are based on missing or partly missing information.

(g) *Facilities covered by an EAD approach*: facilities that are covered by a direct EAD estimate (e.g. estimates in the "region of instability").

(h) *Number of facilities (back-testing)*: the total number of facilities that have defaulted[73] in the relevant observation period, after data exclusions due to process deficiencies (see point (d)), outlier treatment (see point (e)) and missing estimates (see point (f)).

(i) *Number of facilities (application portfolio)*: the total number of facilities within the scope of the rating system or model under investigation (see Sections 2.1 and 2.3) at a given point in time. This definition only applies to Section 2.9.5.

(j) *Line usage*: defined as a facility's current drawn amount as a proportion of the sum of its drawn and undrawn amounts.

In addition to the specific definitions above, the definitions contained in points (a) to (h) of Section 2.3 also apply.

## 2.9.2    Qualitative validation tools (back-testing portfolio)

The analysis of qualitative aspects of CCF models is aimed at ensuring the appropriateness of the assignment process for this parameter.

Results should be reported for the following validation tools:[74]

1. CCF assignment process statistics regarding the occurrence of:

    (a) data exclusions due to process deficiencies that are carried out in the institution's internal validation (see point (d) of Section 2.9.1);

---

[72]  It is not expected that such exclusions are foreseen by the institution.

[73]  See point (c) of Section 2.3.

[74]  See the relevant definitions in Section 2.9.1. Note that the process statistics referred to in this list are exclusions from basic set $M^b$ and that the number of facilities in the validation sample (N) is the result of subtracting all data exclusions listed under "CCF assignment process statistics" from basic set *M*.
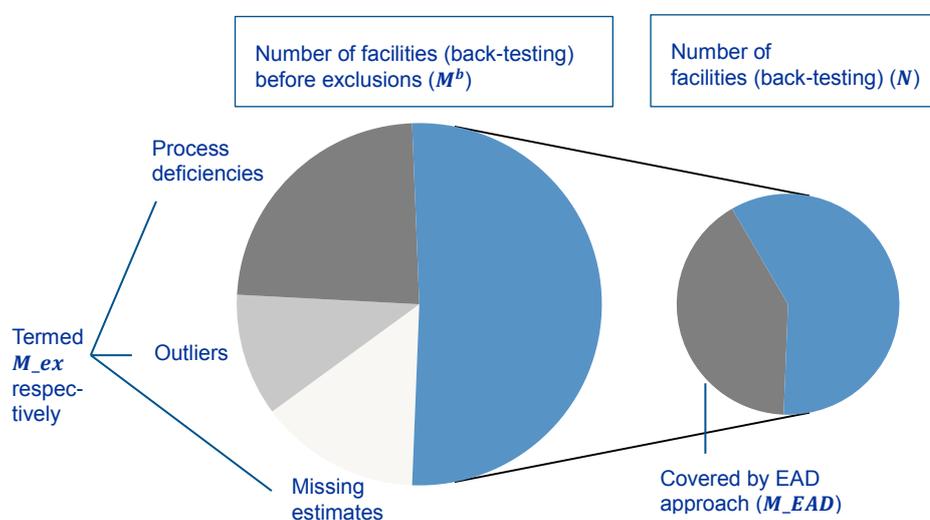
(b)   data exclusions due to outlier treatment for realised CCF (see point (e) of Section 2.9.1);

(c)   facilities with missing CCF or EAD estimates (see point (f) of Section 2.9.1).

2.   Qualitative measures:

(a)   facilities covered by an EAD approach (see point (g) of Section 2.9.1).

**Figure 7**

Illustration of the data basis for qualitative rating process statistics and the resulting back-testing sample



Source: ECB Banking Supervision.

The summary statistics presented in this section are computed on the basis of the portfolio's composition at the reference point for the estimation (see point (a) of Section 2.9.1).

### 2.9.2.1   CCF assignment process statistics

Objectives of the tool

The objective of this validation tool is to verify CCF models' appropriateness by analysing the occurrence of data exclusions.

Description

The statistics considered here indicate the relative frequency of facilities that are excluded due to process deficiencies or due to outlier treatment for realised CCF respectively, and facilities with missing CCF or EAD estimates.

Implementation

Calculate the $M\_ex/M^b$ summary statistic, where $M^b$ denotes the number of facilities envisaged for the back-testing, but before exclusions, and $M\_ex$ denotes the number of excluded facilities in question at the beginning of the relevant observation period (see Figure 7).

Summary statistics are calculated at portfolio level for the following deficiencies:

1.  data exclusions due to process deficiencies that are carried out in the institution's internal validation (see point (d) of Section 2.9.1);

2.  data exclusions due to outlier treatment for realised CCF (see point (e) of Section 2.9.1);

3.  facilities with missing CCF or EAD estimates (see point (f) of Section 2.9.1).

Reports should indicate the number of facilities (back-testing) before exclusions ($M^b$), the respective number of facilities ($M\_ex$) which are excluded from the validation sample, and the number of facilities (back-testing) $N$.

### 2.9.2.2 Facilities covered by an EAD approach

The objective of this validation tool is to analyse the relative frequency of facilities covered by a direct EAD estimate.

The statistic considered here is the relative frequency of facilities that are covered by an EAD approach as defined in point (g) of Section 2.9.1.

Calculate the $M\_EAD/N$ summary statistic, where $N$ denotes the number of facilities (back-testing) after data exclusions as defined in Section 2.9.1 and $M\_EAD$ denotes the number of facilities which are covered by an EAD approach at the reference point for the estimation (see point (a) of Section 2.9.1).

The summary statistic is calculated at portfolio level.

Reports should indicate the number of facilities that are excluded as a result of being covered by an EAD approach ($M\_EAD$).

### 2.9.3 Predictive ability

The analysis of predictive ability (or calibration) aims to ensure that the CCF risk parameter facilitates a good prediction of EAD. Where facilities are covered by an EAD approach (see point (g) of Section 2.9.1) a simplified analysis is applied.

The results of the following tests are reported:

1.  back-testing of the CCF using a t-test;

2.  back-testing of EAD using a t-test.

**Figure 8**

Illustration of the back-testing sample used for CCF models

**This figure shows the construction of the sample used to assess the predictive ability of CCF models. It uses the example of two generic facilities (A and B) whose recovery process has begun in the relevant one-year observation period, indicating the estimated CCF that is to be back-tested in each case.**
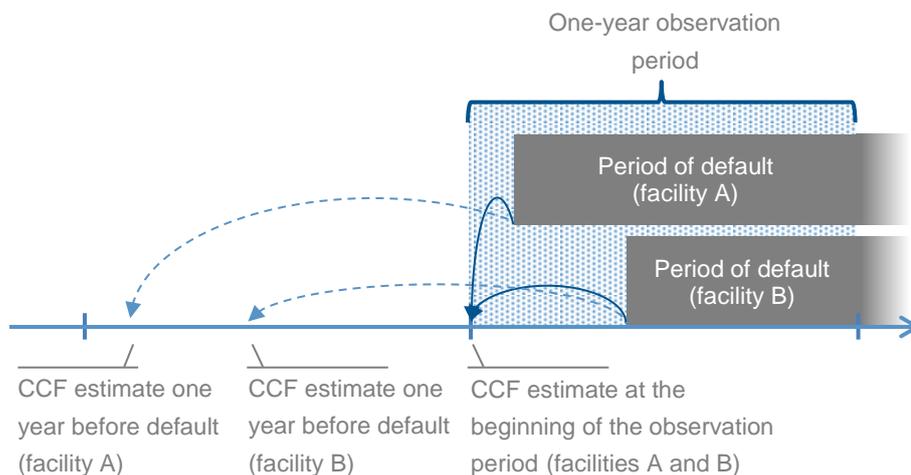
Figure 8 shows the construction of the CCF back-testing sample for two generic facilities. Facilities form part of the back-testing sample if the default occurs in the relevant one-year observation period. As outlined in point (a) of Section 2.9.1, the estimated CCF to be back-tested for an individual facility relates to either (i) the point in time one year before the facility's default or (ii) the beginning of the relevant one-year observation period.

### 2.9.3.1 CCF back-testing using a t-test

*Objectives of the tool*

The objective of this validation tool (CCF back-testing using a one-sample t-test for paired observations) is to assess the predictive ability of CCF estimates at facility grade or pool level (in accordance with Article 182(1)(a) of the CRR).

*Description*

The one-sample t-test for paired observations compares the estimated CCF with the realised CCF under the null hypothesis that the estimated CCF is greater than the true one (one-sided hypothesis test), assuming independent observations. Under the null hypothesis, the test statistic is asymptotically Student-t distributed with $(R - 1)$ degrees of freedom, where $R = N - M\_EAD$ is the number of facilities (back-testing) minus those facilities which are covered by an EAD approach.

The data basis for the t-test consists of all facilities that have defaulted during the relevant observation period (see the definition of the number of facilities (back-testing) ($N$) in point (h) of Section 2.9.1). Facilities that are affected by outlier treatment, such as floors, but are included in the internal validation sample, form part of the relevant data basis.

Calculate the t-test statistic as follows:

$$T = \sqrt{R} \cdot \frac{\sum_{i=1}^{R}(CCF_i^R - CCF_i^E)}{\sqrt{s_{CCF}^2}},$$

$$s_{CCF}^2 = \frac{\sum_{i=1}^{R}\left((CCF_i^R - CCF_i^E) - \frac{1}{R}\sum_{j=1}^{R}(CCF_j^R - CCF_j^E)\right)^2}{R-1},$$

where:

- $R$ is the number of facilities (back-testing) minus those facilities that are covered by an EAD approach (see point (g) of Section 2.9.1);

- $CCF_i^E$ is the estimated CCF for facility i (see point (a) of Section 2.9.1);

- $CCF_i^R$ is the realised CCF for facility i (see point (b) of Section 2.9.1).

Calculate the p-value $1 - S_{R-1}(T)$, where $S_{R-1}$ is the cumulative distribution function of the Student's t-distribution evaluated at the test statistic ($T$) with ($R - 1$) degrees of freedom, and $R = N - M\_EAD$ is the number of facilities (back-testing) minus those which are covered by an EAD approach.

The tests are performed at both of the following levels:

1. portfolio level; and

2. facility grade/pool or segment level.

As regards the second of those, institutions should apply one of the following two approaches:

(a) If the model is based on 20 facility grades/pools or less, the test is performed at the facility grade/pool level used in the institution's internal validation.

(b) Otherwise (including in the case of continuous CCF models), the institution should use 12 predefined "CCF segments" on the basis of the following criteria:

*Segment 1*: facilities i with 0% $\leq CCF_i^E$ < 5%;

*Segment 2*: facilities i with 5% $\leq CCF_i^E$ < 10%;

*Segment 3*: facilities i with 10% $\leq CCF_i^E$ < 20%;

...

(*10% CCF steps from Segment 3 to Segment 11*)

…

*Segment 12*: facilities i with 100% $\leq CCF_i^E$.

The report should include:

- name of the facility grade/pool or segment;

- the number of facilities (back-testing) minus those covered by an EAD approach;

- number-weighted averages for both estimated and realised CCF;

- the percentage of realised CCF values that are floored and (if applicable) the floor which is used;

- on the basis of the same observations used for the t-test, information about the distribution of realised CCF (taking into account the treatment of outliers; see Section 2.9.1): minimum, 5% quantile, 10% quantile, 25% quantile, 50% quantile, 75% quantile, 90% quantile, 95% quantile, maximum, and exposure-weighted[75] average realised CCF at time of default;

- the test statistic ($T$), the estimated variance ($s_{CCF}^2$) and p-value $1 - S_{R-1}(T)$, as described above;

- the percentage of defaulted facilities that are excluded owing to the institution's treatment of outliers as defined in point (e) of Section 2.9.1.

### 2.9.3.2 EAD back-testing for facilities covered by an EAD approach

Where facilities are covered by an EAD approach (see point (g) of Section 2.9.1), a simplified analysis is carried out. These are facilities where the institution uses a direct EAD estimate (e.g. in the "region of instability"). This may, for example, include facilities that are fully drawn at a given point in time, but for which an EAD different to the outstanding amount is estimated.

This back-testing should assess the predictive ability of direct EAD estimates.

The one-sample t-test for paired observations compares the estimated exposure with the realised drawn amount at the time of default $D^R$ under the null hypothesis that the estimated exposure is greater than the drawn amount (one-sided hypothesis test), assuming independent observations. Under the null hypothesis, the test statistic is asymptotically Student-t distributed with ($M\_EAD - 1$) degrees of freedom, where $M\_EAD$ denotes the number of facilities (back-testing) that are covered by an EAD approach.

---

[75] Exposure weighted by the committed but undrawn credit amount.

Perform a t-test based on $D_i^R - EAD_i^E$, where $D_i^R$ denotes the drawn amount at the time of default of facility $i$. The test is based on the following equation:

$$T = \sqrt{M\_EAD} \cdot \frac{\frac{1}{M\_EAD} \sum_{i=1}^{M\_EAD} (D_i^R - EAD_i^E)}{\sqrt{s_{EAD}^2}} \, ,$$

$$s_{EAD}^2 = \frac{\sum_{i=1}^{M\_EAD} \left( (D_i^R - EAD_i^E) - \frac{1}{M\_EAD} \sum_{j=1}^{M\_EAD} (D_j^R - EAD_j^E) \right)^2}{M\_EAD - 1} \, ,$$

where:

- $M\_EAD$ denotes the number of facilities that have defaulted during the relevant observation period which are covered by an EAD approach;

- $EAD_i^E$ is the estimated EAD of facility $i$;

- $D_i^R$ denotes the drawings (balance sheet exposure) at the time of default of facility $i$.

Calculate the p-value $1 - S_{M\_EAD-1}(T)$, where $S_{M\_EAD-1}(T)$ is the cumulative distribution function of the Student's t-distribution with $(M\_EAD - 1)$ degrees of freedom.

The test is performed at portfolio level.

The report should include results for:

- the number of facilities that have defaulted during the relevant observation period which are covered by an EAD approach ($M\_EAD$);

- the sum of the estimated exposure at default ($EAD^E$) and the sum of the observed drawings ($D^R$);

- the test statistic (T), the estimated variance ($s_{EAD}^2$) and p-value $1 - S_{M\_EAD-1}(T)$, where $S_{M\_EAD-1}(T)$ is the cumulative distribution function of the Student t-distribution with $(M\_EAD - 1)$ degrees of freedom.

## 2.9.4 Discriminatory power

The assessment of discriminatory power is aimed at ensuring that CCF models are able to discriminate between facilities with high and low CCF values. The measure used in this section to assess the discriminatory power of CCF models is the generalised AUC. That validation tool is based on a generalisation of the classical AUC that can be applied to multi-class problems. More information on the statistics referred to below can be found in the annex, Section 3.2.

Note that, for simplification purposes, the calculation is performed not at the level of individual facilities, but at the level of aggregated segments or grades/pools, as outlined below.

### 2.9.4.1 Current gAUC vs. gAUC at initial validation/development

The current discriminatory power of the CCF model is benchmarked against the discriminatory power measured (and deemed appropriate) at the time of the initial validation in the course of the model's development.

The gAUC for the relevant observation period is compared with the gAUC at the time of the initial validation[76] during development via hypothesis testing based on a normal approximation, assuming a deterministic gAUC at the time of development.

The data basis for this test consists of all facilities that have defaulted during the relevant observation period and are not covered by an EAD approach ($R = N - M\_EAD$). See Section 2.9.1 for definitions of the number of facilities ($N$) and the number of facilities covered by an EAD approach ($M\_EAD$). The treatment of outliers (see Section 2.9.1) is the same as it is in the institution's internal validation of CCF.

Calculate the test statistic:

$$S = \frac{gAUC_{init} - gAUC_{curr}}{s},$$

where $gAUC_{init}$ denotes the gAUC at the time of the initial validation, $gAUC_{curr}$ denotes the gAUC for the relevant observation period, and $s$ denotes the estimated standard deviation of $gAUC_{curr}$.

The gAUC is calculated on the basis of facility grades or pools as the ordinal segmentation of the CCF. Facility grades or pools are defined in the same way as in the institution's internal validation. If the model is based on more than 20 facility grades/pools or is a continuous CCF model, the test is performed using 12 predefined "CCF segments" on the basis of the following criteria:

> *Segment 1*: facilities i with 0% ≤ $CCF_i^E$ < 5%;
>
> *Segment 2*: facilities i with 5% ≤ $CCF_i^E$ < 10%;
>
> *Segment 3*: facilities i with 10% ≤ $CCF_i^E$ < 20%;
>
> ...
>
> (*10% CCF steps from Segment 3 to Segment 11*)
>
> …
>
> *Segment 12*: facilities i with 100% ≤ $CCF_i^E$.

The test is applied at portfolio level.

---

[76] See the definition of "initial validation/development" in point (h) of Section 2.3.

The report should include:

- results for $\text{AUC}_{\text{init}}$, $\text{AUC}_{\text{curr}}$ and the estimated variance ($s^2$);

- results for the test statistic ($S$) and the p-value $1 - \Phi(S)$, where $\Phi$ denotes the cumulative distribution function of the standard normal distribution;

- information on the sample used to calculate the initial gAUC: the time period of the validation sample (start date and end date), the number of facilities and the variance in the validation sample.

## 2.9.5 Qualitative validation tools (application portfolio)

The analysis of qualitative aspects (application portfolio) is aimed at assessing the distribution of the estimated CCF and its evolution over time. For those facilities in the application portfolio which are covered by an EAD approach, a simplified analysis is carried out.

The results of the following validation tools are reported:

1. CCF assignment process statistics regarding the relative frequency of facilities with missing CCF or EAD estimates in the application portfolio as defined in point (f) of Section 2.9.1;

2. CCF – application portfolio distribution at CCF facility grade/pool level or at the level of predefined CCF segments;

3. EAD – application portfolio statistics at portfolio level.

All summary statistics in this section are computed on the basis of the composition of the application portfolio at the beginning and end of the observation period.

### 2.9.5.1 CCF assignment process statistics

**Objectives of the tool**

The objective of this validation tool is to verify the appropriateness of important aspects of the CCF or EAD values assigned to the portfolio within the scope of the model – more specifically, facilities that present missing estimates.

**Description**

The statistic considered here is the relative frequency of a specific CCF or EAD model deficiency, namely the occurrence of facilities with missing CCF estimates (see point (f) of Section 2.9.1).

**Implementation**

Calculate the $\text{M\_miss}/\text{M}$ summary statistic, where $M$ denotes the number of facilities (application portfolio) as defined in Section 2.9.1 and $\text{M\_miss}$ denotes the number of facilities with missing CCF or EAD values at the beginning of the relevant observation period.

**Scope**

The summary statistic is calculated at portfolio level.

Results should be reported for the number of facilities (application portfolio) and the number of facilities with missing CCF or EAD values (M_miss).

## 2.9.5.2     CCF application portfolio distribution

Objectives of the tool

The objective of this validation tool is to analyse the distribution of the estimated CCF in application portfolios using key drivers of the CCF, such as average line usage, as defined in point (j) of Section 2.9.1, and the average undrawn amount.

Description

The statistics considered here are the number of facilities (application portfolio) without those covered by an EAD approach, the average estimated CCF (see point (a) of Section 2.9.1), the average line usage and the average undrawn amount of the application portfolio by facility grade/pool or by predefined CCF segment. If the model is based on 20 facility grades/pools or less, results are reported at the facility grade/pool level used in the institution's internal validation. Otherwise (including in the case of continuous CCF models), results are reported using the 12 predefined CCF segments (see Section 2.9.3.1).

Statistics on facilities are reported at both the beginning and the end of the relevant observation period.

Implementation

Calculate the number of facilities, number-weighted average line usage, the number-weighted average CCF, and the average undrawn amount at facility grade or pool level. The estimated CCF should be the CCF as used for own funds requirements on the reference date, as defined in point (a) of Section 2.9.1. All figures should be reported for both the beginning and the end of the relevant observation period.

In addition, calculate the Population Stability Index for the estimated CCF, which is defined as:

$$\text{PSI} = \sum_{i=1}^{K} (p_{i,2} - p_{i,1}) \ln\left(\frac{p_{i,2}}{p_{i,1}}\right).$$

Here, $p_{i,j}$ denotes the relative frequency of the observed value $i$ in sample $j$, where $j$ refers to the beginning of the relevant observation period ($j = 1$) and the end of the relevant observation period ($j = 2$) respectively, and where $K$ denotes the number of facility grades/pools or segments.

Scope

The statistics should be calculated as follows. If the model is based on 20 facility grades/pools or less, statistics are calculated at facility grade or pool level. Otherwise, they are calculated using the 12 predefined CCF segments. In each case, they are calculated using (i) the sample at the beginning of the relevant observation period and (ii) the sample at the end of that observation period. The Population Stability Index is calculated at portfolio level.

Reporting of the results

The report should include results for:

- the name of the facility grade/pool or segment;

- the number of facilities (application portfolio) without those covered by an EAD approach, the number-weighted average of the estimated CCF, average line usage and the average undrawn amount for the application portfolio by facility grade/pool or by predefined CCF segment.[77] Results should be reported both for the sample at the beginning of the relevant observation period and for the sample at the end of that observation period;

- the Population Stability Index for facilities covered by a CCF approach (calculated on the basis of the number of facilities (application portfolio) at the beginning and end of the observation period).

### 2.9.5.3     EAD application portfolio

**Objectives of the tool**

The objective of this validation tool is to analyse facilities in the application portfolio which are covered by an EAD approach.

**Description**

The statistics considered here are total estimated EAD, the sum of current drawings and total original exposure for facilities which are covered by an EAD approach (see point (g) of Section 2.9.1).

**Implementation**

For facilities in the application portfolio that are covered by an EAD approach, calculate the total number of facilities, total estimated EAD, the sum of current drawings and total original exposure at portfolio level at both the beginning and the end of the observation period.

**Scope**

Summary statistics are calculated at portfolio level at the beginning and end of the observation period.

**Reporting of the results**

The report should include results for:

- the number of facilities, total estimated EAD, the sum of current drawings and the total original exposure at portfolio level for facilities covered by an EAD approach (at both the beginning and the end of the relevant observation period).

## 2.10     Slotting approach for specialised lending exposures

The validation tools outlined in this section are all aimed at ensuring the adequacy of slot assignment, which determines both risk weights and the expected loss.

The requirements set out in the CRR are complemented by the Regulatory Technical Standards on Assigning Risk Weights to Specialised Lending Exposures under

---

[77] If the model is based on 20 facility grades/pools or less, results are reported at the facility grade/pool level used in the institution's internal validation. Otherwise (including in the case of continuous CCF models), results are reported using the 12 predefined CCF segments.

Article 153(9) of Regulation (EU) No 575/2013.[78] Those RTS introduce four classes of exposure: project finance, real estate, object finance and commodities financing. Credit institutions may use separate templates for each of those exposure classes in line with their internal validation process (i.e. regard them as separate "models"). However, this is not obligatory. It is up to each institution to decide whether such granular reporting adds significant value to the reported results (although whatever decision is made should be applied consistently over time).

The adequacy of slot assignment is assessed using the following areas of investigation:

   (a)   predictive ability;

   (b)   loan tenor check;

   (c)   stability.

### 2.10.1    Predictive ability – slot back-testing

Objectives of the tool

The objective of slot back-testing is to assess whether slot assignment is able to adequately reflect the possible loss. This is accomplished by comparing the expected loss and the mean realised loss rate using hypothesis testing based on the one-sample z-test.

Description

The one-sample z-test compares the expected loss (EL) with the mean realised loss rate under the null hypothesis that the expected loss is greater than the mean realised loss rate. Under the null hypothesis, the test statistic is asymptotically normally distributed. The p-value of the test serves as a measure of the adequacy of slot assignment.

The data basis for the z-test consists of all defaults whose recovery process has been closed within the relevant observation period, but also of the historical default rates observed in the last five consecutive years on the whole portfolio. While the relevant observation period in which the recovery process ends (due to curing, liquidation, etc.) is the same for each observation, the length of the recovery process will typically be different for each of the closed defaults (i.e. the defaults might have occurred in different years).

Implementation

Calculate the test statistic:

$$Z = \frac{\mu_1 \mu_2 - EL}{\sqrt{s^2}},$$

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{5} \sum_{j=0}^{4} \frac{D_{-j}}{N_{-j}} \\ \frac{1}{N_{def}} \sum_{i=1}^{N_{def}} L_i \end{pmatrix}, \; s_L^2 = \frac{\sum_{i=1}^{N_{def}} (L_i - \mu_2)^2}{N_{def}(N_{def}-1)}, \; s_D^2 = \frac{\mu_1(1-\mu_1)}{25} \sum_{j=0}^{4} \frac{1}{N_{-j}},$$

$$s^2 = s_D^2 s_L^2 + \mu_1^2 s_L^2 + \mu_2^2 s_D^2,$$

---

[78]   Final draft EBA/RTS/2016/02 published on 13 June 2016.

where:

- $N_{def}$ is the number of customers whose recovery process has been closed within the relevant observation period;

- $L_i$ is the realised loss rate for customer $i$;

- $D_{-j}$ is the number of customers that defaulted in the j-th year prior to the relevant observation period, where $D_0$ is the number of customers that have defaulted during the relevant observation period;

- $N_{-j}$ is the number of non-defaulted customers at the beginning of the j-th year prior to the relevant observation period, where $N_0$ is the number of non-defaulted customers at the beginning of the relevant observation period;[79]

- *EL* is the expected loss that would have been used to calculate own funds requirements at the beginning of the year in which the customer defaulted based on the slotting method in place at the end of the relevant observation period;

- $s_D^2$ and $s_L^2$ denote the variance of $\mu_1$ and $\mu_2$ respectively, and $s^2$ denotes the variance of $\mu_1\mu_2$, under the assumption that $\mu_1$ and $\mu_2$ are independent and $D_{-j}$ is independent over time.[80]

Calculate the p-value $1 - \Phi(Z)$, where $\Phi$ is the cumulative distribution function of the standard normal distribution.

The test is performed at two different levels:

1. Portfolio level: In this case, EL is the expected loss for the portfolio – i.e. the arithmetic average of the expected losses for all defaults considered pursuant to Article 158(6) of the CRR (customers whose recovery process has been closed within that observation period) – that would have been used to calculate own funds requirements at the beginning of the year in which the customer defaulted based on the slotting method in place at the end of the relevant observation period.

2. Slot level: In this case, EL for a given slot is the expected loss pursuant to Article 158(6) of the CRR for that slot. Here, the data basis is restricted to customers which would have been assigned to a specific slot at the beginning of the year in which the customer defaulted based on the slotting method in place at the end of the relevant observation period. The historical default rates that are calculated correspond to the specific slot in question.

---

[79] Note the fundamental difference between $N_{-j}$ and $N_{def}$ – i.e. $N_{def}$ denotes defaulted customers, while $N_{-j}$ refers to non-defaulted customers.

[80] $S_D^2$ is given by the variance of the binomial distribution, whereas the probability of default in each year is approximated by $\mu_1$.

The report should include results for:

- the expected loss (EL) at the relevant level;

- the number of non-defaulted customers at the beginning of the j-th year prior to the relevant observation period ($N_{-j}$), where $j = 0, \ldots, 4$;

- the number of defaults observed during the j-th year prior to the relevant observation period ($D_{-j}$), where $j = 0, \ldots, 4$;

- the number of customers whose recovery process has been closed within the relevant observation period ($N_{def}$);

- the means and variances of default rates and losses (i.e. $\mu_1$, $\mu_2$, $s_D^2$, $s_L^2$ and $s^2$);

- the test statistic ($Z$) and p-value $1 - \Phi(Z)$, where $\Phi$ is the cumulative distribution function of the standard normal distribution.

## 2.10.2   Loan tenor check

The objective of this validation tool is to verify the assignment of the loan tenor, which is a determining factor in slot assignment. This analysis should be based on all customers for which the institution assigns risk weights in accordance with Article 153(5) and Article 158(6) of the CRR that have remaining maturities of less than 2.5 years.

The statistics considered here relate to the relative frequency of maturity extensions (either directly or via refinancing) for customers[81] with remaining maturities of less than 2.5 years.

Calculate the $M^1/N$ and $M^2/N$ statistics, where $N$ denotes the number of customers for which the exposure-weighted average remaining maturity[82] is less than 2.5 years at the beginning of the relevant observation period. $M^1$ is the subset of those $N$ customers for which the exposure-weighted average remaining maturity at the beginning of the relevant observation period is smaller than the exposure-weighted average remaining maturity at the end of that observation period. Meanwhile, $M^2$ denotes the subset of those $M^1$ customers for which the exposure-weighted average remaining maturity at the end of the relevant observation period is less than 2.5 years. The exposure-weighted average remaining maturity for customer $i$ ($M_i$) is calculated as:

$$M_i = \frac{1}{\sum_{j=1}^{F_i} E_i^j} \sum_{j=1}^{F_i} M_i^j E_i^j,$$

---

81   In the context of slotting, the term "customers" refers to the unit used for assessment under Article 153(5) and Article 158(6) of the CRR.

82   The definition of "remaining maturity" is identical to that applied in Article 153(5) and Article 158(6) of the CRR.

where:

- $F_i$ is the number of facilities of customer $i$;

- $M_i^j$ is the remaining maturity associated with facility $j$ of customer $i$;

- $E_i^j$ is the original exposure[83] associated with facility $j$ of customer $i$.

Those summary statistics are calculated at portfolio level.

The report should include results for:

- the number of customers for which the exposure-weighted average remaining maturity is less than 2.5 years at the beginning of the relevant observation period ($N$);

- the subset of those $N$ customers for which the exposure-weighted average remaining maturity at the beginning of the relevant observation period is smaller than the exposure-weighted average remaining maturity at the end of that observation period ($M^1$), as well as the $M^1/N$ ratio;

- the subset of those $M^1$ customers for which the exposure-weighted average remaining maturity at the end of the relevant observation period is less than 2.5 years ($M^2$), as well as the $M^2/N$ ratio.

## 2.10.3 Customer migrations

The objective of this validation tool is to analyse the numbers of customers migrating across slots during the relevant observation period.

The statistics calculated describe the number of customers migrating over a certain number of slots.

These statistics show the number of customers that migrate from slot i at the beginning of the relevant observation period to slot j at the end of that observation period ($N_{ij}$).

Calculate $N_{ij}$ for customers that migrate between the following categories at the beginning and end of the relevant observation period:

Categories at the beginning of the relevant observation period:

- performing slots for non-defaulted customers;

- slot for defaulted customers.

---

[83] "Original exposure" is defined in point (g) of Section 2.3.

Categories at the end of the relevant observation period:

- performing slots for non-defaulted customers;

- slot for defaulted customers (including defaulted customers that have left the model during the observation period);

- non-defaulted customers for which a different method in accordance with the definition in point (e) of Section 2.3 is used to determine own funds requirements;

- non-defaulted customers that have terminated their business relationship with the credit institution during the relevant observation period (by analogy to the definition in point (h)(iv) of Section 2.5.1).

<table>
<tr><td>Scope</td><td>For each slot at the beginning of the relevant observation period, this test is performed for every possible migration destination at the end of that observation period.</td></tr>
<tr><td>Reporting of the results</td><td>The report should include the number of customers ($N_{ij}$) migrating from slot 1 to 5 at the beginning of the relevant observation period to slot 1 to 5 or to a different method at the end of the relevant observation period, or that have terminated their business relationship with the credit institution during the relevant observation period.</td></tr>
</table>

# 3 Annexes

## 3.1 Annex 1 – Statistics on the discriminatory power of PD models

**Basic notations**

Let $C$ be a portfolio of customers, $R$ an ordered set and $r$ a rating on $C$ with outcomes in $R$ (i.e. a random variable $r: C \to R$). $C$ may be thought of as the portfolio at the beginning of the one-year observation period, and $r(c)$ may be thought of as the rating assigned to customer $c \epsilon C$ at the beginning of the observation period. We assume that in the order in $R$, high ratings are always supposed to indicate good credit quality. For example, for two customers, $a$ and $b$, the following is true:

$$AA = r(a) > BBB+ = r(b) \Rightarrow PD^E(a) < PD^E(b).$$

Let $A$ be the subset of customers in $C$ which have defaulted during the observation period and $B$ the subset of customers in $C$ which have not defaulted during the observation period.

**Mann-Whitney U statistic and AUC**

For any pair $(a, b) \in A \times B$, let

$$u_{a,b} := \begin{cases} 1, & if\ r(a) < r(b), \\ 1/2, & if\ r(a) = r(b), \\ 0, & if\ r(a) > r(b). \end{cases}$$

If $|.|$ denotes the cardinality of a set, the Mann-Whitney U statistic is calculated as:

$$U := \sum_{(a,b) \in A \times B} u_{a,b} = |\{(a,b) \in A \times B | r(a) < r(b)\}| + \frac{|\{(a,b) \in A \times B | r(a) = r(b)\}|}{2},$$

and the area under the ROC curve is calculated as:

$$\text{AUC} := \frac{U}{|A| \cdot |B|} \in [0,1].$$

**Estimates of variance and covariance**

For the following tests, it is necessary to estimate the variance ($\sigma^2$) of the AUC or the covariance ($\sigma_{12}$) of the AUCs of two different rating assignments for portfolio $C$ for the observation period under consideration.

If the vector $V_{10} = \left(V_{10,a}\right)_{a \in A}$ is defined as:

$$V_{10,a} = \frac{1}{|B|} \sum_{b \in B} u_{a,b}\ , \forall a \in A,$$

and the vector $V_{01} = \left(V_{01,b}\right)_{b \in B}$ is defined as:

$$V_{01,b} = \frac{1}{|A|} \sum_{a \in A} u_{a,b}\ , \forall b \in B,$$

then an estimate of the variance $\sigma^2$ is given by:

$$s^2 = \frac{\widehat{\text{var}}(V_{10})}{|A|} + \frac{\widehat{\text{var}}(V_{01})}{|B|}.$$

Here, $\widehat{\text{var}}$ denotes the unbiased sample variance – i.e. for a vector $V = (V_i)_{i \in I}$,

$$\widehat{\text{var}}(V) = \frac{1}{|I| - 1} \sum_{i \in I} \left( V_i - \frac{1}{|I|} \sum_{j \in I} V_j \right)^2.$$

**Multiple observation periods**

For portfolios with a low number of defaults, it might be necessary to aggregate the rankings of multiple observation periods in order to obtain a stable result for the AUC. Consider $T$ different observation periods $t \epsilon \{1, \dots, T\}$ and the corresponding portfolios $C^{(t)}$ and rating assignments $r^{(t)} : C^{(t)} \to R$. Let $C$ be the disjoint union of $C^{(t)}$, which can be expressed as:

$$C := \bigcup_{t=1}^{T} C^{(t)} \times \{t\}$$

(into which each $C^{(t)}$ is embedded in a natural way). This definition is used in order to treat each occurrence of a customer in the portfolios for the various observation periods as a distinct data point. Now, for $c \epsilon C^{(t)}$, the following definitions apply:

$$\tilde{r}^{(t)}(c) := \frac{\left| \{ c' \epsilon C^{(t)} | r^{(t)}(c') \le r^{(t)}(c) \} \right|}{|C^{(t)}|} \epsilon [0,1]$$

and

$$\tilde{r}(c) := \tilde{r}^{(t)}(c) \quad \text{if} \quad c \epsilon C^{(t)}.$$

Thus, the rating assignments $\tilde{r}^{(t)}(c)$ for the various periods are merged to form one single assignment $\tilde{r}(c)$, leaving the ranking for each period unchanged, but replacing the rating outcomes with the cumulative probabilities of the distribution functions in order to make the ratings comparable across periods, while eliminating systematic differences between the absolute risk assignments for the various periods. Note that the values for $\tilde{r}$ must not be confused with default probabilities.

## 3.2 Annex 2 – Statistics on the discriminatory power of LGD/CCF models

Although this annex refers only to LGD models, all results are equally valid for CCF models.

**Basic notations**

Let $\text{LGD}_i^{\text{E}}$ and $\text{LGD}_i^{\text{R}}$ be estimated and realised LGD respectively for facility $i$.

If the model is based on 20 facility grades/pools or less, estimated LGD is clustered on the basis of the LGD estimates for facility grades or pools and ordered from low to

high. Realised LGD values are then discretised on the basis of the estimated LGD for those facility grades or pools and ordered from low to high as well.[84]

If the model is continuous or based on more than 20 facility grades/pools, an ordinal segmentation of LGD is applied using the LGD segments as defined in Section 2.6.2.1, so that the $LGD_i^E$ and $LGD_i^R$ values for all facilities $i$ are discretised.[85] The 12 segments are then ordered from low to high (i.e. from Segment 1 to Segment 12).

The following presentation relates to the case of the 12 LGD segments. The basis for the test is a two-way contingency table (12 times 12)[86] with all possible combinations of discretised $LGD_i^E$ (12 possible segments as rows) and $LGD_i^R$ (12 possible segments as columns) and the observed frequencies for each combination for all pairs of defaulted facilities within the sample (see the corresponding LGD section for details of the relevant data basis for this test).

Let $a_{ij}$ denote the observed frequency in cell $(i, j)$ (i.e. segment $i$ for $LGD^E$ and segment $j$ for $LGD^R$) in the 12x12 contingency table described above. Let $r_i = \sum_j a_{ij}$ be the total for row $i$, $c_j = \sum_i a_{ij}$ be the total for column $j$ and $F = \sum_i \sum_j a_{ij}$ be the total frequency. Let

$$A_{ij} = \sum_{k<i} \sum_{l<j} a_{kl} + \sum_{k>i} \sum_{l>j} a_{kl},$$

$$D_{ij} = \sum_{k>i} \sum_{l<j} a_{kl} + \sum_{k<i} \sum_{l>j} a_{kl},$$

and

$$P = \sum_i \sum_j a_{ij} A_{ij},$$

$$Q = \sum_i \sum_j a_{ij} D_{ij}.$$

$P$ can be understood as twice the number of agreements (i.e. for a given combination of estimated "$i$" and realised "$j$" discretised LGD, the total frequency of observations with both indices greater or smaller than the given combination) in the ordering of the cell indices when all pairs of observations are compared. Similarly, $Q$ is twice the number of disagreements (i.e. for a given combination of estimated "$i$" and realised "$j$" discretised LGD, the total frequency of observations with at least one index greater or smaller than the given combination).

---

[84] The first class of realised LGD consists of all facilities with realised LGD that is smaller than or equal to the smallest estimated LGD; the second class consists of all facilities with realised LGD that is smaller than or equal to the second smallest estimated LGD which are not already part of the first class, and so on. The last class of realised LGD consists of all facilities with realised LGD that is greater than the greatest estimated LGD.

[85] Note that the LGD segments defined in Section 2.6.2.1 apply only to estimated LGD. However, in this annex the same discretisation process will be applied to realised LGD.

[86] Where a model is based on 20 facility grades/pools or less, the size of the contingency table will be the number of facility grades/pools times the number of facility grades/pools plus one.

The following definition of Somers' D (C|R) assumes that the row variable $LGD^E$ is regarded as an independent variable, while the column variable $LGD^R$ is regarded as dependent. The gAUC or Somers' D (C|R) is estimated as:

$$gAUC = \frac{D\,(C|R) + 1}{2},$$

$$D(C|R) = \frac{P - Q}{w_r},$$

and the gAUC's standard deviation ($s$) can be estimated as:

$$s = \frac{1}{w_r^2} \sqrt{\sum_i \sum_j a_{ij} \left( w_r d_{ij} - (P - Q)(F - r_i) \right)^2},$$

where

$$w_r = F^2 - \sum_i r_i^2,$$

$$d_{ij} = A_{ij} - D_{ij}.$$

Somers' D (C|R) and its standard deviation ($s$) are computed on the basis of Brown and Benedetti (1977)[87] and Göktaş and İşçi (2011)[88].

---

[87] Brown, M. and Benedetti, J. (1977), "Sampling Behavior of Tests for Correlation in Two-Way Contingency Tables", *Journal of the American Statistical Association*, Vol. 72, pp. 309-315.

[88] Göktaş, A. and İşçi, Ö. (2011), "A Comparison of the Most Commonly Used Measures of Association for Doubly Ordered Square Contingency Tables via Simulation", *Metodološki zvezki*, Vol. 8(1), pp. 17-37.

## Abbreviations

| | | | |
|---|---|---|---|
| AUC | area under the ROC curve | LGD | loss given default |
| CA | competent authority | MWB | matrix weighted bandwidth |
| CCF | credit conversion factor | PD | probability of default |
| CRD | Capital Requirements Directive (Directive 2013/36/EU) | ROC | receiver operating characteristic |
| | | RTS | regulatory technical standards |
| CRR | Capital Requirements Regulation (Regulation (EU) No 575/2013) | RWEA | risk-weighted exposure amount |
| | | SSM | Single Supervisory Mechanism |
| CV | coefficient of variation | | |
| EAD | exposure at default | | |
| EBA | European Banking Authority | | |
| ECB | European Central Bank | | |
| EU | European Union | | |
| IRB | internal ratings-based | | |