



European Securities and
Markets Authority

Discussion Paper

**On the validation and review of Credit Rating Agencies'
methodologies**



Responding to this paper

The European Securities and Markets Authority (ESMA) invites responses to the questions listed in this Discussion Paper on the validation and review of Credit Rating Agencies' methodologies.

All contributions should be submitted online at www.esma.europa.eu under the heading 'Your input - Consultations'.

Please follow the instructions given in the document "Reply form for the Discussion Paper on the validation and review of Credit Rating Agencies' methodologies" also published on the ESMA website.

Comments are most helpful if they:

- respond to the question stated;
- indicate the specific question to which the comment relates;
- contain a clear rationale; and
- describe any alternatives ESMA should consider.

ESMA will consider all comments received by **February 19 2016**.

Publication of responses

All contributions received will be published following the close of the consultation, unless you request otherwise. Please clearly and prominently indicate in your submission any part you do not wish to be publically disclosed. A standard confidentiality statement in an email message will not be treated as a request for non-disclosure.

The collection of confidential responses is without prejudice to the scope of Regulation (EC) No 1049/2001. Possible requests for access to documents will be dealt in compliance with the requirements and obligations laid down in Regulation (EC) No 1049/2001.

Data protection

Information on data protection can be found at <http://www.esma.europa.eu> under the heading [Legal Notice](#).

Who should read this paper

All interested stakeholders are invited to respond to this Discussion Paper. In particular, responses are sought from Credit Rating Agencies registered in accordance with Regulation (EC) No 1060/2009, users of credit ratings and experts in the field of credit risk methodology validation and review.

Table of Contents

Definitions / Acronyms used	4
1 Executive Summary	5
2 Introduction	7
3 Background to Validation.....	9
4 Current Industry Status.....	11
4.1 Sufficient Quantitative Evidence	11
4.2 Limited Quantitative Evidence	13
4.3 Identifying and Addressing Anomalies.....	13
5 Validation of Methodologies with Sufficient Quantitative Evidence	15
5.1 Discriminatory Power	15
5.2 Predictive Power	16
5.3 Historical Robustness.....	18
6 Validation of Methodologies with Limited Quantitative Evidence.....	20
7 Identifying and addressing anomalies.....	23
Annex I: Summary of Consultation Questions	25

Definitions / Acronyms used

ESMA	European Securities and Markets Authority
CRAs	Registered Credit Rating Agencies
CRA Regulation	Regulation (EC) No 1060/2009 of the European Parliament and of the Council of 16 September 2009 on credit ratings agencies (as last amended by Regulation (EU) No 462/2013)
RTS on rating methodologies	Commission Delegated Regulation (EU) No 447/2012 of 21 March 2012 supplementing Regulation (EC) No 1060/2009 of the European Parliament and of the Council on credit rating agencies by laying down regulatory technical standards for the assessment of compliance of credit rating methodologies

1 Executive Summary

Reasons for publication

This discussion paper seeks stakeholders' views on the validation and review of Credit Rating Agencies' (CRAs) methodologies. This discussion paper will help the European Securities and Markets Authority (ESMA) to develop further its views on the quantitative and qualitative techniques used as part of the validation of methodologies required under Article 8(3) of the Regulation (EC) No 1060/2009 of the European Parliament and of the Council of 16 September 2009 on credit ratings agencies (as last amended by Regulation (EU) No 462/2013) (CRA Regulation). This Article states that *'a credit rating agency shall use rating methodologies that are rigorous, systematic, continuous and subject to validation based on historical experience, including back testing'*. In particular, this discussion paper focuses on the last part of Article 8(3), i.e. *'subject to validation based on historical experience, including back testing'*. This discussion paper also asks for views on the quantitative and qualitative techniques that should be used as part of the review of methodologies required under Article 8(5) of the CRA Regulation. This Article states, *inter alia*, that a CRA shall *'review its credit ratings and methodologies on an ongoing basis and at least annually'*.

This discussion paper requests views on how CRAs should demonstrate rating methodologies' *'discriminatory power'*, *'historical robustness'*, *'predictive power'* or that the methodologies are *'sensible predictors of credit worthiness'*. This is as part of meeting requirements set out in Article 8(3) of the CRA Regulation, Article 8(5) of the CRA Regulation and the Commission Delegated Regulation (EU) No 447/2012 of 21 March 2012 supplementing Regulation (EC) No 1060/2009 of the European Parliament and of the Council on credit rating agencies by laying down regulatory technical standards for the assessment of compliance of credit rating methodologies (RTS on rating methodologies). In addition, the discussion paper seeks views on how CRAs should meet the requirement in both Articles 7 and 8 of the RTS on rating methodologies that the CRAs shall have *'processes in place to ensure that systemic credit rating anomalies highlighted by back-testing are identified and are appropriately addressed'*.

ESMA has decided to issue a discussion paper on the validation and review of credit rating methodologies based on its supervisory experience of CRAs' application of Article 8(3) and Article 8(5) of the CRA Regulation. The discussion paper reflects discussions with competent authorities who supervise the validation and review of credit risk / rating methodologies in the financial services industry, and experts / academics in the field. ESMA is of the view that a discussion of how CRAs should meet Articles 8(3) and 8(5) of the CRA Regulation will help to ensure the consistent application of validation and review measures for demonstrating the discriminatory power, predictive power and historical robustness of methodologies, as well as to identify measures that CRAs should implement when validating and reviewing methodologies with limited quantitative evidence.

Input from stakeholders will help ESMA in providing further guidance to the industry. ESMA would appreciate if input is provided with supporting data. Input will be kept confidential where required and requested.

Contents

The remainder of the paper is structured as follows:

Section 2 introduces the topic discussed in this paper;

Section 3 provides a background to validation;

Section 4 provides a summary of current industry practice;

Section 5 requests views in relation to validation and review of methodologies where there is sufficient quantitative evidence;

Section 6 requests views in relation to validation and review of methodologies where there is limited quantitative evidence;

Section 7 requests views in relation to the identification and addressing of anomalies;

Annex I is a summary of all questions.

Next Steps

Respondents are asked to provide responses by February 19 2016. ESMA will review responses to the paper with the intention of issuing a Feedback Statement in Q1 2016 and establishing whether there is a need for further guidance to the industry.

2 Introduction

1. Article 8(3) of the CRA Regulation requires CRAs to use rating methodologies that are rigorous, systematic, continuous and subject to validation based on historical experience, including back testing. In this respect, the RTS on rating methodologies sets out the rules to be used in the assessment of compliance of credit rating methodologies with the requirements laid down in Article 8(3) of the CRA Regulation. In particular the RTS on rating methodologies sets out high level requirements for credit rating methodology validation, including that CRAs have to demonstrate the discriminatory power, predictive power and historical robustness of their methodologies.
2. Article 8(5) of the CRA Regulation states that a CRA shall, inter alia, review its credit ratings and methodologies on an ongoing basis and at least annually.
3. Article 22a of the CRA Regulation states that ESMA, in the exercise of its ongoing supervision of CRAs, shall examine regularly CRAs' compliance with Article 8(3) of the CRA Regulation. ESMA has decided to issue a discussion paper on the quantitative and qualitative techniques used as part of the validation and review of methodologies based on its supervisory experience of CRAs' application of Articles 8(3) and 8(5) of the CRA Regulation. ESMA recognises that CRAs have a number of individual challenges in the validation and review of their methodologies. For example, recently established CRAs tend to have a short run of historical data that can be used for the validation and review of methodologies and consequently these CRAs tend to use more qualitative techniques. Equally, there are particular asset classes which will normally have limited data available for validating and reviewing the respective methodologies through quantitative techniques, such as low default asset classes.
4. Given these challenges, ESMA is of the view that a discussion paper on the use of quantitative and qualitative techniques by CRAs in implementing Article 8(3) and Article 8(5) of the CRA Regulation will help to ensure a consistent and appropriate standard in the use of quantitative and qualitative techniques used as part of validation and review, particularly in demonstrating the discriminatory power, predictive power and historical robustness of methodologies. ESMA also asks views on the measures CRAs should consider when validating and reviewing methodologies with limited quantitative evidence. The measures included in the discussion paper reflect ESMA's supervisory experience of CRAs, discussions with competent authorities who supervise credit risk / rating methodologies validation and review in the financial services industry, and expert / academics in the field.

5. This discussion paper refers to both the validation and review of a CRA's methodologies. In the remainder of this document both the words 'validation' and 'review' are used interchangeably instead of 'validation and review' for ease of reading .
6. The word 'methodology' is used in this document as to mean all components that a credit rating methodology may consist of, including the models and the key rating assumptions.
7. Per article 23 of the CRA Regulation, this discussion paper raises questions which do not imply or suggest interference with the content of credit ratings or methodologies.

3 Background to Validation

8. ESMA recognises that validation is a broad and evolving discipline. This section outlines the main characteristics of the validation approach with respect to CRAs' methodologies.
9. The validation approach of CRAs can be divided into two parts. The first part refers to the validation processes and governance developed by CRAs. This includes, among other components, i) the gathering of necessary information from reliable sources and the assessment of data quality (e.g. accuracy, completeness, timeliness and appropriateness of the data used for validation purposes), ii) the documentation of the relevant policies and procedures, including the end products of a validation exercise and the content of them, iii) the governance structure employed and iv) the record keeping arrangements. These aspects of validation are addressed in the CRA Regulation outside of Article 8(3) and Article 8(5) and the RTS on rating methodologies¹.
10. The second part, which this discussion paper focuses on, refers to the quantitative and qualitative techniques applied by the CRAs for validating their methodologies. Validation techniques typically consist of several components which involve both quantitative and qualitative assessments. ESMA is of the view that the right balance should be struck by CRAs between the application of quantitative and qualitative techniques. ESMA understands that both kinds of techniques can provide valuable insight into the performance of methodologies, and that, dependent on the circumstances (e.g. asset class or data availability), the degree to which quantitative and qualitative techniques are applied may differ. However, ESMA is of the view that in most cases the validation of the methodologies should include both qualitative and quantitative techniques.
11. It is commonly understood that validation techniques are usually divided into two categories: back-testing and benchmarking. The back-testing of the methodologies refers to the comparison of the expected to the observed outcome of the credit ratings assigned by these methodologies and it consists of techniques demonstrating the methodologies' discriminatory power, predictive power and historical robustness. There is a large variety of techniques for each of the 3 components (discrimination, prediction, robustness) of back-testing. The benchmarking of methodologies, on the other hand, refers usually to the comparison of methodologies' credit ratings to external credit risk measures.

¹ For example, data quality of ratings is addressed in Article 8(2), policies and procedures is addressed in point 3 of Section A of Annex I of the CRA Regulation, governance is addressed in points 2 and 9 of Section A of Annex I of the CRA Regulation and recording keeping in point 7 of Section B of Annex I of the CRA Regulation

12. Articles 7 and 8 of the RTS on rating methodologies also set out further rules to be used in the assessment of compliance of credit rating methodologies regarding their validation for both the validation process and the validation techniques used. These rules concern the following:

- a. assessment of the assumptions used in rating models and their deviation from the actual default and loss rates;
- b. examination of a methodology's sensitivity to changes;
- c. assessment of historic credit ratings produced by the validated methodology;
- d. assessment of the validation inputs, including the size of the data samples;
- e. taking into account the main geographic areas of the rated entities or financial instruments;
- f. in-sample and out-of-sample testing;
- g. analysis of previous validation results; and
- h. assessment of the consistent application of internal procedures.

13. This discussion paper does not address these further rules in Articles 7 and 8 of the RTS on rating methodologies as ESMA has not currently identified through its supervisory experience significant inconsistencies in the application of these rules.

4 Current Industry Status

14. This section sets out ESMA's understanding of the current industry status based on its general supervisory experience and ESMA's investigation into validation conducted for the four largest CRAs during 2014 and 2015.

4.1 Sufficient Quantitative Evidence

Discriminatory Power

15. ESMA found broad variation in the extent to which CRAs demonstrated the discriminatory power of their methodologies in a quantitative manner.

16. One or more CRA demonstrated the discriminatory power of their methodologies by using the Accuracy Ratio (AR)², a metric which estimates the rank ordering power of methodologies. AR was the most commonly used discriminatory measure. A number of CRAs also used the Kolmogorov-Smirnov statistic for demonstrating the discriminatory power of their methodologies. In more than one instance, CRAs assessed their methodologies' discriminatory power by reviewing the (average) ratings in certain time periods (e.g. 1 year) before the observation of a default event. Qualitative measures were used as well, such as the assessment of the distribution of the observed default rates.

17. These measures were in some instances compared to either the average performance of a similar or over-arching asset class (e.g. corporates or aggregated structured finance performance) or other internal threshold values.

18. One or more CRA found it challenging to perform tests to confirm the discriminatory power of their methodologies, even if there was what ESMA considered to be sufficient quantitative evidence to do so.

19. One or more CRA used at least two of the techniques identified above. In addition, one or more CRA used confidence intervals / levels for their statistical tests, including the bootstrapping technique. ESMA identifies these approaches as examples of good practice in the industry.

² In this paper, the term Accuracy Ratio also encompasses the Gini Coefficient or other similar measures.

Predictive Power

20. CRAs cited challenges in measuring predictive power, including the argument that their ratings were based on an ordinal rather than cardinal scale and that the expected behaviour of the rating categories could not be fixed since it relates to various factors, including the economic or business cycle of an industry / economy.
21. CRAs mainly used the Binomial test in the assessment of the predictive power of their methodologies. Other tests that were used by the CRAs included the Hosmer-Lemeshow Chi-Square test, the Normal test, the Brier Score and a test comparing 3-year cumulative default rates to the ECAs (External Credit Assessment Institutions) monitoring level benchmarks of Basel II³. For applying these tests, CRAs usually set confidence intervals / levels. In some instances, CRAs compared qualitatively the difference between the expected and the observed behaviour of their ratings.
22. One or more CRA used internal thresholds in order to assess the performance of its methodologies with regards to their predictive power.
23. One or more CRA used at least two of the techniques identified above. ESMA identifies this approach as an example of good practice in the industry.

Historical Robustness

24. The vast majority of the CRAs referred to transition (migration) studies in demonstrating the historical robustness of their methodologies. The majority of CRAs also performed additional measures, such as reviews of large movements, reviews of ratings which were downgraded from investment to non-investment grade, stability measures regarding the ratings' distribution or the characteristics of the underlying population of the methodologies (e.g. System / Population Stability Index) and qualitative assessment of frequency distributions.
25. Since these measures are more qualitative than quantitative in nature, CRAs typically did not have any thresholds in place. However, some CRAs installed thresholds for the stability measures as well as for certain statistics calculated from the transition matrices (e.g. the percentage observed in the diagonal or related statistical tests).
26. One or more CRA used at least two of the techniques identified above. ESMA identifies this approach as an example of good practice in the industry.

³ International Convergence of Capital Measurement and Capital Standards, Basel Committee on Banking Supervision, June 2006

4.2 Limited Quantitative Evidence

27. Where CRAs validated methodologies that had limited quantitative evidence, in the majority of cases they used qualitative measures to demonstrate that credit rating methodologies were sensible predictors of credit worthiness (Article 8 of the RTS on rating methodologies).
28. These measures included the benchmarking of the credit ratings assigned by the methodologies to other credit risk measures (or of the actual methodologies to similar methodologies of other CRAs) as well as some of the robustness techniques outlined above.
29. ESMA did observe one or more CRA enhancing the available data so that they could apply more quantitative measures as part of the validation. For example, one or more CRA developed hypothetical transactions (by randomly distributing an existing portfolio of underlying assets) in order to increase the number of relevant observations for the purposes of validation. In another instance, one or more CRA developed a model that could predict ratings based on its methodology and subsequently applied this model to a larger population than its rating universe. An additional approach applied by one or more CRA was the use of third party data.
30. ESMA recognises the challenges for CRAs in validating methodologies with limited quantitative evidence. Nonetheless, ESMA is of the opinion that CRAs should enhance the validation techniques they apply in such cases and put in place more qualitative measures in order to perform a more robust validation, based on historical experience of their methodologies, per Article 8(3). ESMA is also of the opinion that CRAs should further consider whether statistical measures may be used in such cases.
31. One or more CRA benchmarked their ratings to other credit risk measures and used at least two of the robustness techniques identified above. ESMA identifies this approach as an example of good practice in the industry.

4.3 Identifying and Addressing Anomalies

32. As described in the above paragraphs, there were cases where one or more CRA used internal thresholds for assessing the performance of their methodologies.
33. However, in most of these cases, the thresholds in place did not have a link to a specified action that the CRA should take in the case of a breach of these thresholds.

34. One or more CRA has established internal thresholds for all their validation quantitative techniques and these thresholds are linked to predetermined actions based on a traffic light approach (for example, a breach may result in the predetermined action to undertake a further review). ESMA identifies this approach as an example of good practice in the industry.

5 Validation of Methodologies with Sufficient Quantitative Evidence

5.1 Discriminatory Power

35. Article 7(1) of the RTS on rating methodologies requires a CRA to ‘use credit ratings methodologies that are supported by quantitative evidence of the discriminatory power of the credit rating methodology’.
36. ESMA is of the view that the discriminatory power of a methodology relates to its ability to rank order the rated entities in accordance to their future status (defaulted or not defaulted) at some predefined time horizon.
37. ESMA has found that the validation techniques used to demonstrate the discriminatory power of methodologies differ among CRAs. ESMA is of the view that it would raise standards in the industry if CRAs consistently use a minimum set of statistical measures in demonstrating the discriminatory power of their methodologies.
38. ESMA is of the view that a CRA should demonstrate the discriminatory power of its methodologies using a range of statistical measures. A CRA should use the cumulative accuracy profile (CAP) curve in conjunction with the accuracy ratio. In addition, a CRA should consider complementing these measures with additional statistical measures, such as the Kolmogorov-Smirnov statistic or the receiver operator characteristics (ROC) curve (along with a confusion matrix), and qualitative measures, such as the distribution of the observed default rates.

Questions

1. Do you agree with ESMA’s view regarding the discriminatory power of methodologies?
2. Do you agree that the Accuracy Ratio, as derived from the CAP curve, is the minimum statistical measure that a CRA should use as part of its validation processes for demonstrating the discriminatory power of its methodologies?
3. Do you agree that complementary measures such as the Kolmogorov-Smirnov statistic and the ROC curve (along with a confusion matrix) add further information to the discriminatory power of methodologies? If not, please explain why.
4. Are there additional quantitative measures that CRAs should use and which would add further insight into the discriminatory power of methodologies? If yes, please explain the measures and your rationale.

5. Are there qualitative measures that are appropriate for demonstrating the discriminatory power of methodologies? If yes, please explain the measures and your rationale.

5.2 Predictive Power

39. Article 7(2)(a) of the RTS on rating methodologies requires a CRA to ‘use credit rating methodologies that describe the historical robustness and predictive power of credit ratings issued using the relevant methodology over appropriate time horizons and across different asset classes’.
40. ESMA is of the view that the predictive power of a methodology can be demonstrated by comparing the expected behaviour of the ratings assigned from this methodology to the observed results. For performing this comparison, a CRA should define internally its expectations (absolute numbers or ranges) per rating category with regards to the measure of creditworthiness its ratings refer to.
41. ESMA has observed that the majority of the CRAs find assessing the predictive power of their methodologies challenging. In certain cases, CRAs state that their ratings are based on an ordinal rather than a cardinal ranking which limits the extent to which internal expectations are relevant to the validation of the predictive power of a methodology, given the volatility of these expectations across the economic cycle.
42. These CRAs do nonetheless have an expectation of an acceptable range of a creditworthiness measure associated to their rating categories (e.g. a large observed default rate for low risk credit rating categories would in fact be taken into account during the validation of a methodology) which in ESMA’s opinion should be further articulated and related to demonstrating the predictive power of a methodology.
43. On the other hand, a number of CRAs do have internal expectations of what they expect from their ratings and some of them perform relevant statistical tests for demonstrating the predictive power of their methodologies.
44. ESMA is of the opinion that the users of credit ratings expect ratings not only to be an accurate opinion on the rank ordering of the rated entities, but also to meet creditworthiness-related expectations. These expectations are derived from CRAs’ historical performance as well as other information, such as the exercises relating rating categories to specific values or ranges of creditworthiness measures (e.g. ECAIs’ mapping or ECB’s ECAF mapping).
45. ESMA has found that the validation techniques used to demonstrate the predictive power of methodologies differ among CRAs, including cases where CRAs have cited challenges in demonstrating the predictive power of methodologies. ESMA is of the

view that it would raise standards in the industry if CRAs consistently use a minimum standard of statistical measures in demonstrating the predictive power of their methodologies. ESMA is of the view that a CRA may use different approaches for defining their internal expectations (e.g. by statistical calculation or by reference to the historical performance of their ratings).

46. For ratings which refer to default probabilities, ESMA is of the view that a CRA should compare the expected probabilities of default to the observed default rates using the binomial and the chi-square tests. In addition, a CRA should consider complementing the above mentioned measures with further statistical measures, such as the Brier Score or the Vasicek one-factor model test.
47. ESMA recognises that credit ratings can act as opinions not only on default but also on other creditworthiness measures (e.g. loss severity). ESMA includes in this discussion paper statistical measures regarding the predictive power of a methodology that refer only to the comparison between expected probabilities of default and observed default rates. This is because the vast majority of the registered CRAs which state that their ratings do predict the creditworthiness of the rated entities relate their credit ratings to this creditworthiness measure (i.e. default). ESMA is of the view that a CRA whose ratings refer to a creditworthiness measure other than default probabilities should develop and employ relevant statistical tests for comparing the expected behaviour of the ratings to the observed results.

Questions

6. Do you agree with ESMA's view regarding the predictive power of methodologies?
7. Do you agree that statistical measures of predictive power increase the quality of validation of CRAs methodologies and should be performed by the CRAs?
8. Do you agree that the binomial and the chi-square tests are the minimum statistical measures that a CRA (when its ratings refer to default probabilities) should use as part of its validation processes for demonstrating the predictive power of its methodologies?
9. Do you agree that complementary measures such as the Brier score and the Vasicek one-factor model test add further information to the predictive power of methodologies (when the CRAs' ratings refer to default probabilities)? If not, please explain why.
10. Are there additional measures that CRAs should use and which would add further insight into the predictive power of methodologies when the CRAs' ratings refer to default probabilities? If yes, please explain the measures and your rationale.

11. Are there qualitative measures that are appropriate for demonstrating the predictive power of methodologies when the CRAs' ratings refer to default probabilities? If yes, please explain the measures and your rationale.
12. Do you agree that CRAs using methodologies related to creditworthiness measures other than default probabilities should use statistical measures to demonstrate the predictive power of their methodologies? If yes, please state the potential creditworthiness measures that methodologies could relate to and the corresponding statistical measures as well as any appropriate qualitative measures.
13. If ESMA establishes that there is a need for further guidance to the industry, should this guidance also cover the demonstration of predictive power of methodologies related to creditworthiness measures other than default probabilities?

5.3 Historical Robustness

48. As stated above, Article 7(2)(a) of the RTS on rating methodologies requires a CRA to 'use credit rating methodologies that describe the historical robustness and predictive power of credit ratings issued using the relevant methodology over appropriate time horizons and across different asset classes'.
49. ESMA is of the view that the historical robustness of a methodology can be demonstrated by assessing other dimensions that do not relate to its discriminatory or predictive power. Examples of these dimensions are the stability of the ratings assigned by the methodology, the stability of the characteristics of the rated entities covered by the methodology (e.g. in relation to previous years, as well as to the development population of the methodology) and the distribution of the assigned ratings. In addition, further validation techniques could be considered under this broad category such as the univariate analysis of key drivers of the methodology (ESMA notes that if the univariate analysis includes statistics such as the AR or the Information Value then this assessment could also be categorised as a measure demonstrating discrimination) and the comparison of the ratings to external credit risk measures (benchmarking).
50. ESMA has found a wide range of measures undertaken by CRAs in demonstrating the historical robustness of their methodologies. These included common measures which assessed the stability of ratings through transition / migration matrices, such as the assessment of upgrade and downgrade rates (direction and magnitude of changes). Some CRAs also used a range of other measures, including comparing performance against the historical performance of other CRAs.

51. ESMA is of the view that it would raise standards in the industry if CRAs consistently use a number of minimum measures in demonstrating the historical robustness of their methodologies. ESMA is of the view that a CRA should demonstrate the stability of its methodologies using statistical measures, such as the Population / System Stability Index. In addition, a CRA should consider producing transition (migration) matrices and analysing the movement of the ratings.
52. A CRA should also consider complementing the above mentioned measures with further qualitative analysis, such as the analysis of the ratings' distributions, univariate analysis of key determinants of ratings, or the benchmarking of the ratings to external credit risk measures (e.g. ratings of other CRAs, credit default swaps spreads, bond yields).

Questions

14. Do you agree with ESMA's view regarding the historical robustness of methodologies?
15. Do you agree that stability statistical measures and the transition (migration) matrices are the minimum measures that a CRA should use as part of its validation processes for demonstrating the historical robustness of its methodologies?
16. Do you agree that complementary measures such as distribution analysis, the univariate analysis of rating determinants and benchmarking add further information to the historical robustness of methodologies? If not, please explain why.
17. Are there additional measures (qualitative or quantitative) that CRAs should use and which would add further insight into the historical robustness of methodologies? If yes, please explain the measures and your rationale.

6 Validation of Methodologies with Limited Quantitative Evidence

53. Article 8 of the RTS on rating methodologies provides an exemption to CRAs from complying with Article 7 of the RTS on rating methodologies in cases where there is *'limited quantitative evidence to support the predictive power of a credit rating methodology'*. Article 8 states that in these instances, a credit rating agency should:

- a) ensure that credit rating methodologies are sensible predictors of credit worthiness;
- b) apply internal procedures in a consistent way and over time and across different market segments;
- c) have processes in place to ensure that systemic credit rating anomalies highlighted by back-testing are identified and are appropriately addressed.

54. ESMA is of the view that a CRA should establish itself the minimum number of ratings and / or defaults that a methodology should have in order to be validated in accordance to Article 7 of the RTS on rating methodologies since this could differ per asset class. CRAs should internally establish the relevant policies and procedures for deciding if there is limited quantitative evidence to support the predictive power of a methodology. These policies and procedures should at a minimum define the responsible persons / parties for taking this decision as well as the relevant criteria that this decision will be based on.

55. In the validation of methodologies with limited quantitative evidence, ESMA is of the view that a CRA should, as part of the process of validating its methodologies, seek to enhance the data sample in order to, if possible, apply Article 7 of the RTS on rating methodologies. A CRA should consider data enhancement solutions such as:

- expanding the data sample with the use of third party data (if available and subject to verifying data quality);
- combining (if meaningful) asset classes or sub-asset classes with similar risk characteristics in order to perform joint validation assessments; or
- creating, if possible, hypothetical transactions that can be used to expand the available data.

ESMA is of the view that a CRA should document its decision making process for determining whether or not to use data enhancement techniques.

56. If CRAs are unable to enhance their data samples, ESMA is of the view that the main tasks of the CRAs are to i) ensure that credit rating methodologies are sensible predictors of credit worthiness and ii) perform back-testing.
57. As stated previously, ESMA understands back-testing as consisting of measures applied in assessing the discriminatory power, the predictive power and the historical robustness of methodologies. Under Article 8, measures of predictive power do not need to be performed if there is limited quantitative evidence. ESMA is of the view that the historical robustness measures described above could be applied also in cases of limited quantitative evidence (in order to capture both requirements of sensible predictors of credit worthiness and back-testing). As such, CRAs should consider measures that may enable them to perform statistical tests to demonstrate the discriminatory power of their methodologies, as described above too.
58. More specifically, ESMA is of the view that a CRA should consider measures enabling it to perform statistical tests for demonstrating the discriminatory power of its methodologies. A CRA should consider measures such as:
- the use of a 'relaxed' default definition for the purposes of validation (e.g. if an asset class is a low default one, then use, for the purposes of validation, the ratings of the highest credit risk, non-default, rating category / categories as default observations);
 - combining rating categories; or
 - using an extended time period;

ESMA is of the view that a CRA should document its decision making process and set out the rationale for the methods it uses to enhance its ability to perform statistical tests for demonstrating the discriminatory power of its methodologies, including whether it has rejected the use of a method.

59. ESMA has found that the majority of CRAs, when validating their methodologies, considered that there was limited quantitative evidence to support their predictive power and used qualitative measures to demonstrate that the methodologies were sensible predictors of credit worthiness.
60. ESMA recognises the challenges for CRAs in validating methodologies with limited quantitative evidence. Nonetheless, ESMA is of the opinion that CRAs should enhance the validation techniques they apply in such cases and put in place more qualitative measures in order to perform a more robust validation of their methodologies. ESMA is of the view that it would raise standards in the industry if

CRAs consistently use a number of minimum measures when validating methodologies with limited quantitative evidence.

61. More specifically, ESMA is of the view that a CRA should consider producing transition (migration) matrices and analysing the movement of the ratings as well as benchmarking the ratings to external credit risk measures (e.g. ratings of other CRAs, credit default swaps spreads, bond yields).
62. In addition, a CRA should consider complementing the above mentioned measures with the other historical robustness measures mentioned above, if applicable, and assess the meaningfulness of performing the discriminatory power measures mentioned above as well.

Questions

18. Do you agree with ESMA's view regarding the validation of methodologies with limited quantitative evidence?
19. Do you agree that CRAs should, as a first step, investigate data enhancement in validating methodologies with limited quantitative evidence?
20. Do you agree that CRAs should, as a second step, investigate measures that may enable them to perform statistical tests to demonstrate the discriminatory power of their methodologies?
21. Do you agree that historical robustness measures should be performed when validating methodologies with limited quantitative evidence?
22. Do you agree that the transition (migration) matrices and benchmarking are the minimum measures that a CRA should use as part of its validation processes for methodologies with limited quantitative evidence?
23. Do you agree that complementary historical robustness measures add further information to the validation processes for methodologies with limited quantitative evidence? If not, please explain why.
24. Are there additional measures that CRAs should use when validating methodologies with limited quantitative evidence? If yes, please explain the measures and your rationale.

7 Identifying and addressing anomalies

63. Articles 7 and 8 of the RTS on rating methodologies make specific reference to the systemic anomalies. More specifically, both Articles mention that *'systemic credit rating anomalies highlighted by back-testing are identified and are appropriately addressed'*.
64. ESMA is of the view that the CRAs should internally set thresholds for their quantitative validation techniques in order to identify and address potential anomalies highlighted by back-testing.
65. These thresholds should be appropriately documented and recorded. ESMA is of the view that the Internal Review Function of CRAs should be responsible for deciding these thresholds, by making sure that they are i) relevant to the methodology being validated, ii) a challenging and consistently applied component of the validation process by being set at appropriate levels (i.e. methodologies should not always pass all validation techniques nor should methodologies always not pass all validation techniques) and iii) adequately justified (i.e. clearly explaining the rationale for choosing them).
66. ESMA understands that the thresholds may differ per asset class; however, this should be appropriately justified by CRAs, especially in cases where the rating categories have the same characteristics across asset classes.
67. ESMA also understands a breach of a threshold will not always lead to methodology changes. Deviations from the thresholds could be justified by various factors such as the economic cycle. A CRA should predefine and justify the actions that deviations from the thresholds will result in.
68. ESMA's understanding is that both systemic and non-systemic anomalies should be identified and appropriately addressed. A CRA should distinguish systemic deviations from non-systemic ones and explain how the predefined actions would differ in such a case.
69. ESMA has found that a number of CRAs are already using specific thresholds for their validation techniques. In addition, some CRAs have already established predetermined actions when these thresholds are reached.
70. ESMA is of the view that it would raise standards in the industry if CRAs consistently use specific thresholds for their quantitative validation techniques in order to identify and address anomalies highlighted by back-testing.

Questions

25. Do you agree that thresholds should be set for the quantitative validation techniques?
26. Do you agree that the Internal Review Function should decide on these values?
27. Do you agree that predefined actions should be documented by CRAs for when the thresholds are met?

Annex I: Summary of Consultation Questions

1. Do you agree with ESMA's view regarding the discriminatory power of methodologies?
2. Do you agree that the Accuracy Ratio, as derived from the CAP curve, is the minimum statistical measure that a CRA should use as part of its validation processes for demonstrating the discriminatory power of its methodologies?
3. Do you agree that complementary measures such as the Kolmogorov-Smirnov statistic and the ROC curve (along with a confusion matrix) add further information to the discriminatory power of methodologies? If not, please explain why.
4. Are there additional quantitative measures that CRAs should use and which would add further insight into the discriminatory power of methodologies? If yes, please explain the measures and your rationale.
5. Are there qualitative measures that are appropriate for demonstrating the discriminatory power of methodologies? If yes, please explain the measures and your rationale.
6. Do you agree with ESMA's view regarding the predictive power of methodologies?
7. Do you agree that statistical measures of predictive power increase the quality of validation of CRAs methodologies and should be performed by the CRAs?
8. Do you agree that the binomial and the chi-square tests are the minimum statistical measures that a CRA (when its ratings refer to default probabilities) should use as part of its validation processes for demonstrating the predictive power of its methodologies?
9. Do you agree that complementary measures such as the Brier score and the Vasicek one-factor model test add further information to the predictive power of methodologies (when the CRAs' ratings refer to default probabilities)? If not, please explain why.
10. Are there additional measures that CRAs should use and which would add further insight into the predictive power of methodologies when the CRAs' ratings refer to default probabilities? If yes, please explain the measures and your rationale.
11. Are there qualitative measures that are appropriate for demonstrating the predictive power of methodologies when the CRAs' ratings refer to default probabilities? If yes, please explain the measures and your rationale.

12. Do you agree that CRAs using methodologies related to creditworthiness measures other than default probabilities should use statistical measures to demonstrate the predictive power of their methodologies? If yes, please state the potential creditworthiness measures that methodologies could relate to and the corresponding statistical measures as well as any appropriate qualitative measures.
13. If ESMA establishes that there is a need for further guidance to the industry, should this guidance also cover the demonstration of predictive power of methodologies related to creditworthiness measures other than default probabilities?
14. Do you agree with ESMA's view regarding the historical robustness of methodologies?
15. Do you agree that stability statistical measures and the transition (migration) matrices are the minimum measures that a CRA should use as part of its validation processes for demonstrating the historical robustness of its methodologies?
16. Do you agree that complementary measures such as distribution analysis, the univariate analysis of rating determinants and benchmarking add further information to the historical robustness of methodologies? If not, please explain why.
17. Are there additional measures (qualitative or quantitative) that CRAs should use and which would add further insight into the historical robustness of methodologies? If yes, please explain the measures and your rationale.
18. Do you agree with ESMA's view regarding the validation of methodologies with limited quantitative evidence?
19. Do you agree that CRAs should, as a first step, investigate data enhancement in validating methodologies with limited quantitative evidence?
20. Do you agree that CRAs should, as a second step, investigate measures that may enable them to perform statistical tests to demonstrate the discriminatory power of their methodologies?
21. Do you agree that historical robustness measures should be performed when validating methodologies with limited quantitative evidence?
22. Do you agree that the transition (migration) matrices and benchmarking are the minimum measures that a CRA should use as part of its validation processes for methodologies with limited quantitative evidence?

23. Do you agree that complementary historical robustness measures add further information to the validation processes for methodologies with limited quantitative evidence? If not, please explain why.
24. Are there additional measures that CRAs should use when validating methodologies with limited quantitative evidence? If yes, please explain the measures and your rationale.
25. Do you agree that thresholds should be set for the quantitative validation techniques?
26. Do you agree that the Internal Review Function should decide on these values?
27. Do you agree that predefined actions should be defined by the CRAs when the thresholds are met?