![ESMA — European Securities and Markets Authority]

Feedback Statement

On the Discussion Paper on the validation and review of Credit Rating Agencies' methodologies

13 July 2016 | ESMA/2016/1122

Table of Contents

Definitions, Legislative References and Acronyms

ESMA	European Securities and Markets Authority
CRAs	Registered Credit Rating Agencies
Discussion Paper	Discussion Paper on the validation and review of Credit Rating Agencies' methodologies
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
CAP	Cumulative Accuracy Profile
ROC	Receiver Operator Characteristic

1 Executive Summary

Reasons for publication

The European Securities and Markets Authority (ESMA) published on 17 November 2015 a Discussion Paper on the validation and review of Credit Rating Agencies (CRAs)' methodologies (Discussion Paper)[1]. ESMA held an Open Hearing on this topic on 25 January 2016.

In this feedback statement, ESMA considers the responses received to the Discussion Paper during Q1 2016. ESMA has decided to consult on guidelines on the validation and review of CRAs' methodologies based on its supervisory experience of CRAs' application of Articles 8(3) and 8(5) of the CRA Regulation and Articles 7 and 8 of the RTS on rating methodologies, and the views expressed to ESMA following the Discussion Paper. The proposed guidelines reflect discussions with various stakeholders, including industry participants (mainly through feedback given to the relevant Discussion Paper and Open Hearing). ESMA is of the view that guidelines on 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.

Contents

Section 2.1 summarizes the general responses ESMA has received, and ESMA's view.

Section 2.2 summarizes the responses to the section on discriminatory power, and ESMA's view.

Section 2.3 summarizes the responses to the section on predictive power, and ESMA's view.

Section 2.4 summarizes the responses to the section on historical robustness, and ESMA's view.

Section 2.5 summarizes the responses to the section on methodologies with limited quantitative evidence, and ESMA's view.

Section 2.6 summarizes the responses to the section on identifying and addressing anomalies, and ESMA's view.

The **Annex** provides the Opinion of the Securities and Markets Stakeholder Group.

Next Steps

ESMA issues, together with this feedback statement, a consultation paper on the proposed guidelines.

---

[1] ESMA/2015/1735 Discussion Paper on the validation and review of Credit Rating Agencies' methodologies
https://www.esma.europa.eu/sites/default/files/library/2015-1735_discussion_paper_on_validation_final.pdf

2 Feedback Statement

1. This section provides a summary of the responses to the consultation, identifying the main comments from respondents and ESMA's view of those responses, together with clarifications of ESMA's intention, where appropriate.

2.1 General responses received

2. ESMA received in total twenty six responses, of which sixteen were confidential.

3. Sixteen respondents provided general remarks to the Discussion Paper. The general remarks received provide a good overview of the general tone of the more detailed responses received to the questions put forward in the Discussion Paper.

4. Three respondents from the CRA industry supported the measures proposed in the Discussion Paper and recommended in their general remarks that further measures could be included.

5. The majority of respondents who were users of credit ratings or other participants supported the Discussion Paper and suggested that CRAs should also disclose their validation process, as well as summarize validation results and revisions made to methodologies as a result of their validation process. One respondent stated that CRAs should be subject to the same constraints and assessments as envisaged in the Basel Committee on Banking Supervision's framework and applied to banks using the Internal Ratings-Based approach under credit risk for calculating their minimum capital requirements, in order to ensure consistency across regulations.

6. **ESMA's response**: ESMA notes the views related to further disclosure. However these recommendations sit outside 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) and consequently ESMA does not require CRAs to publish their validation process or the results of their validation. Material changes to methodologies should however follow the disclosure requirements under Articles 8(5a) and 8(6) of the CRA Regulation.

7. Several respondents from the CRA industry stated that the quantitative/statistical measures and tests proposed as minimum techniques for validation, or noted as examples of complementary measures in the paper were not relevant to their credit ratings. These respondents stated that this was because they did not attach to their credit ratings a specific probability of default, and that the description of predictive power used by ESMA in the Discussion Paper[2] could potentially change the definition

---

[2]Page 16, para 40 of the Discussion paper: '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'

of a credit rating and the intended meaning of credit ratings. The majority of respondents who expressed this view were CRAs who stated that their credit ratings were an ordinal measure of credit risk, and that the approach adopted by ESMA, particularly in reference to predictive power, implied a level of precision not intended in their credit ratings. It was argued that a consequence of the proposed minimum techniques regarding quantitative/statistical tests for demonstrating predictive power, could be that CRAs focus on targeted default rates, irrespective of the economic cycle, and in consequence introduce more volatility and mechanistic reliance in credit ratings. Some CRAs further argued that in requesting this type of measures ESMA could be perceived as interfering with the meaning of the credit ratings.

8. **ESMA's response**: ESMA does not share the view that its proposals would change the product that CRAs issuing ordinal credit ratings offer, for a number of reasons:

    a. While ordinal credit ratings may be the primary objective for a number of CRAs, rank ordering is not the sole goal. For example, all CRAs take into account the default rates of their credit rating categories, even if the ordinal system is maintained, and the CRAs would be likely to consider these default rates under their validation process if they differ significantly from their expectations or past experience.

    b. Validation is an internal process and ESMA is not proposing that CRAs share any expectations of ratings behaviour (e.g. expected default probabilities) publicly.

    c. ESMA does not propose to require CRAs to establish specific expectations. ESMA has suggested that CRAs may establish expectations based on ranges per credit rating category for example, giving flexibility and allowing CRAs to implicitly recognize the impact of potential factors that could influence the expectations of the CRAs on rating behaviour.

    d. The suggestion that this approach would result in greater credit rating volatility over-simplifies ESMA's approach. ESMA is not suggesting that CRAs should automate their approach so that if a rating category exceeds or falls below their expectations, the CRAs should change their methodology/credit ratings mechanistically. ESMA believes however that with the use of predictive power measures, a CRA will enhance its validation process and have a more consistent and objective approach with which to identify and assess when a credit rating methodology is not performing as expected, and decide on the appropriate next steps, if any.

    e. ESMA does not intend to interfere with the content, product or rating philosophy of a CRA's credit ratings or credit rating methodologies as per Article 23 of the CRA Regulation.

9. ESMA therefore maintains in the proposed guidelines the same approach and measures on predictive power as proposed in the Discussion Paper.

10. One respondent suggested that ESMA should consider creating different validation frameworks for the credit ratings of entities and the credit ratings of financial obligations, stating that this would reflect the fundamental differences in credit rating subjects, data availability and performance expectations.

11. **ESMA's response**: ESMA is of the view that the expectations set out in the Discussion Paper, and the discretion given to the CRAs to establish for themselves an appropriate internal validation framework are sufficient to enable CRAs to reflect potential differences in the characteristics of various asset classes.

12. Two respondents expressed concern that an outcome of the Discussion Paper was the introduction of purely quantitative parameters in the validation and review of credit rating methodologies. The respondents were concerned that this would exclude the benefits brought to the validation process from the use of qualitative measures and expert judgement applied by staff.

13. **ESMA's response**: It is not the intention of ESMA that the measures proposed in the Discussion Paper or the proposed guidelines should result in CRAs mechanistically focusing on quantitative measures in the validation process, to the exclusion of qualitative measures and expert judgment.

14. ESMA recognises that good quality validation is the outcome of the processes, governance, measures, and equally importantly, the expert judgment used by CRAs. ESMA is of the view that good quality validation strikes a balance 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. ESMA's view is that the validation of the credit rating methodologies should include both qualitative and quantitative techniques.

15. ESMA focused the Discussion Paper and the proposed guidelines on quantitative measures, as this is where the industry appears least clear on ESMA's expectations. A benefit of quantitative measures is that they provide further objectivity to the validation process, particularly as it can be harder to recognise and articulate the inherent assumptions used in interpreting qualitative measures. This does not however mean that ESMA believes that quantitative measures should solely drive a validation process and ESMA does not expect that validation outcomes should mechanistically rely on quantitative measures.

16. ESMA has clarified this intent in the proposed guidelines.

17. Four CRA respondents noted that the statistical tests/approaches proposed in the Discussion Paper were limited by the assumptions used in them, the dependency on large data samples and the impact of the time horizon chosen. One respondent stated that focusing on statistical measures would lead to certain business models of CRAs

to be precluded because the statistical measures gained significance with larger numbers of credit ratings.

18. **ESMA's response**: ESMA recognises that all statistical tests/measures have certain assumptions and that these assumptions may not always fully reflect the environment in which CRAs issue credit ratings. ESMA notes, however, that the statistical tests/measures have the benefit that the assumptions can be clearly and transparently articulated when interpreted. ESMA expects that CRAs will apply expert judgment when interpreting the results of statistical tests/measures, including considering how the assumptions used and limitations in tests may affect results.

19. In addition, there are other more complicated statistical tests/measures that address most of the limitations brought up by the respondents. ESMA invites the CRAs to explore these tests/measures in order to address their concerns.

20. With regard to the dependency on large data samples, ESMA has included quantitative tests/measures only in the cases of sufficient quantitative evidence.

21. The time horizon is one of the parameters that the CRAs should decide upon when performing their validation process.

22. Five respondents raised a concern that there are certain risks in elevating any particular measure to the position of a *de facto* standard measure, and that one measure should not be mandated over another.

23. **ESMA's response**: ESMA notes that the Discussion Paper and the proposed guidelines are a result of its supervisory experience and the need to raise the practices of the industry in this particular area.

24. ESMA has reflected the concern in the use of *de facto* measures in the proposed guidelines by mirroring current good industry practice when setting out the measures it 'typically expects', and giving CRAs discretion in the further complementary measures they may choose to apply. Where a CRA chooses to diverge from the measures ESMA 'typically expects', it should document its rationale, explaining how it meets the regulatory requirements (Articles 8(3) and 8(5) of the CRA Regulation and Articles 7 and 8 of the RTS on rating methodologies), as clarified in the Discussion Paper and the proposed guidelines. The examples of complementary measures are also not intended to be an exhaustive list. ESMA believes that the suite of measures the CRAs will use, including qualitative analysis and expert judgement, along side the quantitative measures that ESMA expects, will ensure that no measure is over-depended on.

25. One respondent argued that the description of 'back testing' proposed by ESMA[3], should not be used. The respondent argued that where there is limited available credit

---

[3] Page 9, paragraph 11 of the Discussion Paper: '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'.

rating data to support predictive power, it is likely to provide limited evidence of discriminatory power also.

26. **ESMA's response**: ESMA acknowledges this challenge, but is of the view that the exemption provided under Article 8 of 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)[4] only applies to predictive power, and therefore CRAs should still assess the meaningfulness of performing the discriminatory power measures for methodologies with limited quantitative evidence. In recognition of this challenge, ESMA proposes that CRAs should first consider to enhance data and consider measures enabling them to perform statistical tests for demonstrating the discriminatory power of their credit rating methodologies. To aid this process, ESMA has provided examples based on the "good practice" it has seen through its supervision in the proposed guidelines.

27. Two respondents noted that ESMA treated discriminatory power, predictive power, and historical robustness as different dimensions, but argued that these dimensions have a large degree of overlap and that individual tests may cover one, two or all of these dimensions.

28. **ESMA's reponse**: ESMA agrees that some tests/measures may cover more than one dimension, depending on the structure and implementation of these tests/ measures.

29. One respondent noted that the Discussion Paper urges respondents to think in categories of sufficient quantitative evidence versus lack of quantitative evidence, and the border between the two is not necessarily clear cut.

30. **ESMA's response:** ESMA acknowledges this element of the Discussion Paper's approach, which reflects the RTS on rating methodologies. ESMA has given CRAs discretion in how they interpret when a methodology has limited quantitative evidence in the proposed guidelines.

31. One respondent stated that the proposed approach will produce a disproportionate increase of workload for smaller CRAs.

32. **ESMA's response**: This concern does not appear to be widely shared across the industry. An increased workload was identified by only one respondent. No respondent provided data on the potential impact and costs that could emerge as a consequence of ESMA's proposals. Equally, ESMA found that it was respondents representing

---

[4] 'In cases where there is limited quantitative evidence to support the predictive power of a credit rating methodology, a credit rating agency shall be exempt from complying with Article 7 of this Regulation if it: (a) ensures that credit rating methodologies are sensible predictors of credit worthiness; (b) applies internal procedures in a consistent way and over time and across different market segments; (c) has processes in place to ensure that systemic credit rating anomalies highlighted by back-testing are identified and are appropriately addressed'.

smaller CRAs who most frequently reported more advanced quantitative techniques in validating their credit rating methodologies, and who recommended the use of further complementary measures as part of validation. ESMA regards the measures it typically expects in the proposed guidelines as providing a good baseline for CRAs' validation of methodologies, with sufficient flexibility, irrespective of their size. The requirement for CRAs to consider further complementary measures provides adequate flexibility to reflect the nature, scale and complexity of a CRA's business. ESMA therefore does not intend to change its approach based on this feedback.

2.2 Discriminatory Power

Q1: Do you agree with ESMA's view regarding the discriminatory power of methodologies?

33. The majority of respondents supported ESMA's view regarding the discriminatory power of methodologies.

34. Two respondents stated that in some sectors, the low default rate (for example, sovereigns) makes it difficult to demonstrate discriminatory power. Another respondent noted that in general, statistical measures only became useful when used with an extremely large number of observations.

35. One respondent stated that ESMA's use for the future status of creditworthiness of 'defaulted/non defaulted' was too narrow, given that there may be other relevant credit events when assessing discriminatory power (for example, breach of collateral or covenant, or skipped payment).

36. One respondent stated that the measure of discriminatory power should apply to credit ratings rather than credit rating methodologies, as credit ratings can be based on more than one methodology. The same respondent also stated that it is appropriate to consider all credit rating transitions, not only transition to default.

37. **ESMA's response**: ESMA acknowledges the challenge of demonstrating discriminatory power in low probability of default methodologies or credit rating methodologies with limited quantitative evidence. In the proposed guidelines, ESMA provides some measures which CRAs may consider when assessing discriminatory power in these circumstances.

38. ESMA recognizes that other types of credit events may be useful indicators as part of a CRA's validation, but is of the view that a general description linked to non default/default event is most relevant and appropriate for the CRA sector. ESMA notes also that each CRA has its own default definition.

39. ESMA recognizes that in some instances CRAs assign ratings by considering more than one credit rating methodology. However, the RTS on rating methodologies requires CRAs to demonstrate discriminatory power of a methodology and ESMA's

approach is consistent with the RTS on rating methodologies. ESMA allows CRAs flexibility in how they will validate their different credit rating methodologies, either primary, or secondary, or across asset classes. Regarding the credit rating transitions, ESMA has included under the historical robustness section the transition (migration) matrices and the analysis of the movement of the credit ratings.

Q2: 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?

40. The majority of respondents agreed with the statement.

41. Some CRA respondents provided caveats in their agreement, including:

    a. Two CRA respondents noting the challenges in using statistical measures in methodologies that are naturally low default.

    b. Two CRA respondents argued that the accuracy ratio should only be applicable to credit ratings that have pre-set default probabilities.

    c. One non-CRA respondent disagreed with ESMA's approach, stating that the Cumulative Accuracy Profile (CAP) curve was not a good reflection of CRA ratings, and that in most cases a default was the result of a stochastic process.

42. **ESMA's response**: ESMA recognizes and mentions in the proposed guidelines measures to address demonstration of discriminatory power in credit rating methodologies which have limited quantitative evidence. ESMA does not share the view that the use of the accuracy ratio requires an expected default probability. Based on its own supervisory experience and feedback it has received from the Discussion Paper respondents, ESMA does not share the view with respondent's criticism of the CAP curve.

43. In the proposed guidelines, ESMA will describe the accuracy ratio (or related measures), as derived from the CAP or the Receiver Operating Characteristic (ROC) curve, as a measure that ESMA typically expects CRAs to use as part of their validation.

Q3: 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.

44. The majority of respondents thought that the Kolmogorov-Smirnov statistic and ROC curve (along with a confusion matrix) added value in limited circumstances and that in the majority of cases the accuracy ratio derived from the CAP or the ROC curve is a sufficient measure of discriminatory power. Some respondents also stated that CRAs should be able to use alternative complementary measures.

45. **ESMA's response**: ESMA agrees that it should be a CRA's decision which additional complementary measures it chooses to apply in its validation, as was reflected in the original proposed text of the Discussion Paper ('complementary measures *such as* the Kolmogorov-Smirnov statistic and the ROC curve'). ESMA agrees that the calculation method of the accuracy ratio (through the CAP or the ROC curves) should be decided by the CRAs. Given that respondents have not demonstrated a consensus on preferred alternative examples to the Kolmogorov-Smirnov statistic, ESMA will continue to propose this measure as an example of possible complementary measures. This does not however prevent CRAs from opting to use any of the alternative examples they cite in their feedback as complementary, provided that these complementary measures meet the regulatory requirements (Articles 8(3) and 8(5) of the CRA Regulation and Articles 7 and 8 of the RTS on rating methodologies), as clarified in the Discussion Paper and the proposed guidelines. The Kolmogorov-Smirnov statistic is used as an example only.

Q4: 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.

46. The majority of respondents did not propose additional measures. Measures proposed by CRA respondents included a review of distribution, multiyear stability analysis, cluster analysis, the mean rating of default companies, investment grade default rate, and confidence interval calculations. One non-CRA respondent suggested that a study of peer groups of different industries for a 3-5 year time horizon may be useful.

47. **ESMA's response**: ESMA recognizes the measures suggested as mainly qualitative in nature and so does not include them in the proposed guidelines as additional quantitative measures. ESMA supports the use of confidence interval calculations for the quantitative discriminatory power measures (e.g. accuracy ratio) as more advanced quantitative techniques that could further enhance the validation process of the CRAs.

Q5: Are there qualitative measures that are appropriate for demonstrating the discriminatory power of methodologies? If yes, please explain the measures and your rationale.

48. Respondents suggested a diverse range of qualitative measures including:

   a. Distribution of observed default rates per rating grades.

   b. Visual checks of CAP/ROC Curve.

   c. Qualitative market feedback.

   d. Median/mean rating before default.

   e. Qualitative assessment of transition matrix.

f. Number of defaulted entities associated with the low risk credit rating categories.

g. Peer group and industry analysis.

h. Scenario analysis.

i. Root cause analysis of large and unexpected credit rating changes as well as of default cases.

49. **ESMA's response**: As stated earlier in the document, ESMA supports the use of qualitative measures as part of the validation and review of credit rating methodologies. The focus of the Discussion Paper and proposed guidelines does however remain on the quantitative measures that should be used as part of validation. Given the wide range of potential measures recommended, and the fact that there was no consensus around one measure, ESMA will continue to include in the proposed guidelines, as an example of qualitative measures related to discriminatory power, the distribution of the observed default rates. ESMA leaves to the discretion of the CRAs the identification of the qualitative measures that are most appropriate for their validation.

2.3 Predictive Power

Q6: Do you agree with ESMA's view regarding the predictive power of methodologies?

50. One respondent suggested that ESMA considers creating different validation frameworks for credit ratings of entities and credit ratings of financial obligations, stating that this would reflect the fundamental differences in credit rating subjects, data availability and performance expectations.

51. **ESMA's response**: As noted in paragraph 11, ESMA is of the view that the expectations set out in the Discussion Paper, and the discretion given to the CRAs to establish for themselves an appropriate internal validation framework are sufficient to enable CRAs to reflect potential differences in the characteristics of various asset classes.

52. Four respondents from the CRA industry stated that the quantitative/statistical measures and tests proposed as minimum techniques or noted as examples of complementary measures in the paper were not relevant to their credit ratings because they did not attach to their credit ratings a specific probability of default. The respondents argued further that the description of predictive power used by ESMA in the Discussion Paper [5] could potentially change the definition of a credit rating and the intended meaning of credit ratings. The majority of respondees who expressed this view were CRAs who stated that their credit ratings were an ordinal measure of credit

---

[5] Page 16, para 40 of the Discussion Paper: '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'.

risk, and that the approach adopted by ESMA implied a level of precision not intended in their credit ratings. These respondents argued that a consequence of the proposed minimum techniques regarding quantitative/statistical tests for demonstrating predictive power, could be that CRAs focus on targeted default rates, irrespective of the economic cycle, and in consequence introduce more volatility and mechanistic reliance in credit ratings. Some CRAs further argued that in requesting this type of measures ESMA could be perceived as interfering with the meaning of the credit ratings.

53. **ESMA's response**: ESMA does not share the view that its proposals would change the product that CRAs issuing ordinal credit ratings offer, for a number of reasons:

   a. While ordinal credit ratings may be the primary objective for a number of CRAs, rank ordering is not the sole goal. For example, all CRAs take into account the default rates of their credit rating categories, even if the ordinal system is maintained, and the CRAs would be likely to consider these default rates under their validation process if they differ significantly from their expectations or past experience.

   b. Validation is an internal process and ESMA is not proposing that CRAs share any expectations of credit ratings behaviour (e.g. expected default probabilities) publicly.

   c. ESMA does not propose to require CRAs to establish specific expectations. ESMA has suggested that CRAs may establish expectations based on ranges per credit rating category, for example, giving flexibility to the approach and allowing CRAs to implicitly recognise the impact of potential factors that could influence the expectations of the CRAs on credit rating behaviour.

   d. The suggestion that this approach would result in greater credit rating volatility over-simplifies ESMA's approach. ESMA is not suggesting that CRAs should automate their approach so that if a credit rating category exceeds or falls below their expectations, the CRAs should change their credit rating methodology/credit ratings mechanistically.ESMA believes however that with the use of predictive power measures, a CRA will enhance its validation process and have a more consistent and objective approach with which to identify and assess when a methodology is not performing as expected, and decide on the appropriate next steps, if any.

   e. ESMA does not intend to interfere with the content, product or credit rating philosophy of a CRA's credit ratings or methodologies as per Article 23 of the CRA Regulation.

54. ESMA therefore maintains in the proposed guidelines the same approach and measures on predictive power as proposed in the Discussion Paper.

Q7: Do you agree that statistical measures of predictive power increase the quality of validation of CRAs' methodologies and should be performed by the CRAs?

55. The majority of respondents from the CRA industry, users of credit ratings and other stakeholders agreed with ESMA's view that statistical measures of predictive power increase the quality of validation of CRAs methodologies and should be performed by the CRAs.

56. Some CRA respondents challenged ESMA's view, arguing that the use of statistical measures of predictive power did not work well with rank-order credit ratings, or credit ratings that did not have an expected probability of default.

57. One respondent also stated that statistical tests were of limited value because they were only valid under certain sets of assumptions.

58. **ESMA's response**: Per ESMA's response to question 6, ESMA believes that demonstrating the predictive power of a credit rating methodology does strengthen the quality of validation.

59. ESMA recognises that all statistical tests/measures have certain assumptions and that these assumptions may not always fully reflect the environment in which CRAs issue credit ratings. ESMA notes, however, that the statistical tests/measures have the benefit that the assumptions can be clearly and transparently articulated when interpreted. ESMA expects that CRAs will apply expert judgment when interpreting the results of statistical tests/measures, including considering how the assumptions used and limitations in tests may affect results.

60. In addition, there are other more complicated statistical tests/measures that address most of the limitations brought up by the respondents. ESMA invites the CRAs to explore these tests/measures in order to address their concerns.

61. Given the reasons set out under questions 6 and 7, ESMA will propose statistical measures that a CRA should typically use in demonstrating the predictive power of a credit rating methodology.

Q8: Do you agree that the binomial and the chi-square test 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?

62. Some respondents, representing users of credit ratings and CRAs, agreed with ESMA's view.

63. One CRA respondent recommended also the Spiegelhalter test, while another stated that the Area Under Curve (AUC), CAP and ROC were more useful measures.

64. Three respondents cited challenges in using these quantitative measures, including that assumptions of independence of default events which underlie these tests do not hold in financial and credit performance data.

65. Three respondents stated that these minimum statistical measures were appropriate only when ratings referred to default probabilities. Another respondent stated that while statistical measures should be used, predictive power measures were for internal use only and therefore there is no market benefit to standardisation. One non-CRA argued that CRAs could meet the intentions of the regulator without the use of the statistical measures proposed.

66. **ESMA's response**: ESMA will not include the Spiegelhalter test as it was proposed by only one respondent, but thanks the respondent for the proposal. ESMA does not share the view that discriminatory measures can be used to demonstrate predictive power.

67. Although CRAs are not required to disclose results of predictive power measures, the benefit from these proposed guidelines is intended to be an improvement in industry practices in this particular area (since previous regulatory approaches that did not provide any guidance resulted in validation approaches below ESMA's supervisory expectations).

68. Please refer to feedback under question 11 for ESMA's response on non-statistical measures of predictive power.

Q9: 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.

69. Feedback to question 9 was similar to that received to question 8.

70. The majority of the CRAs' respondents and the users of credit ratings' respondents supported complementary measures and the examples given. One CRA respondent stated that it should be left to individual CRAs to determine what tests to use.

71. Other CRA respondents gave feedback on the Brier Score and Vasicek one factor model test. One CRA respondent stated that these measures were not useful where credit ratings are not defined by default probabilities. One CRA stated that the measures did not add information to the predictive power of methodologies while another argued that AUC, CAP and ROC were preferable measures. Two CRA respondents noted challenges in the use of these measures. One respondent noted that the Brier Score can be easily influenced by low default probabilities and the Vasicek one factor model test simplified the correlation analysis to just one correlation factor, and only works for large portfolios. Another noted that the Vasicek one factor model test requires a Gaussian correlation parameter which has to be assumed.

72. **ESMA's response**: ESMA does not share the view that discriminatory measures can be used to demonstrate predictive power. ESMA recognizes that both the Brier Score and Vasicek one factor test use assumptions which need to be considered when interpreting results, but notes that these assumptions can be clearly and objectively assessed and understood by CRAs (please refer to paragraphs 18-19 for further clarifications on ESMA's response regarding the assumptions of the measures/tests). Further, ESMA suggests these two measures as complementary ones; if a CRA is of the opinion that these specific measures do not add any valuable information in their validation process, they may use other statistical measures/tests that provide better insight into the predictive power of their methodologies. This gives CRAs flexibility to use other measures if these would be more appropriate, provided that these complementary measures meet the regulatory requirements (Articles 8(3) and 8(5) of the CRA Regulation and Articles 7 and 8 of the RTS on rating methodologies), as clarified in the Discussion Paper and the proposed guidelines. For CRAs whose credit ratings refer to a creditworthiness measure other than default probabilities, please refer to questions 12 and 13.

73. Given that the majority of respondents supported these measures, ESMA will use them as non-exhaustative examples of complementary measures in the proposed guidelines.

Q10: 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.

74. The majority of respondents did not recommend further measures. Suggestions that were made included:

   a. Sign test.

   b. Test of monotonicity.

   c. Calibration diagram to test under/over-estimation.

   d. Normal test per probability of default bucket over multiple years.

   e. Evaluation of (cumulative) default rates.

   f. Investment grade default rate.

   g. Median rating before default.

   h. Summary statistics, describing location, spread, shape and dependence.

   i. Spiegelhalter test.

   j. Transition matrices and mobility index.

k. Presentation of CRA's expectations per rating notch, segment and sub-segment of expected loss based on long term through the cycle assessment of average loss rates and unexpected losses based on the maximum loss rates (estimated at a high confidence level) in a period of significant economic distress.

75. **ESMA's response**: The majority of the proposed measures could be used either for demonstrating the discriminatory power or historical robustness of credit rating methodologies, or are qualitative measures. Given the lack of consensus around other potential additional measures, and general support for examples of additional measures given in the Discussion Paper, ESMA will use these same examples of complementary measures from the Discussion Paper in the proposed guidelines.

Q11: 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.

76. The majority of respondents did not recommend qualitative measures appropriate for demonstrating the predictive power of methodologies. The following measures were recommended by a minority of respondents:

a. Histogram of default rates overall buckets (established probabilities of defaults vs realised default rates and analysis of the nature/direction of difference).

b. Qualitative measures of environmental factors including industry, market, company information and management of credits.

c. Qualitative analysis of actual loss rates versus the CRA's expectations.

77. **ESMA's response**: Given the lack of consensus around potential qualitative measures, ESMA will not include qualitative measures in the proposed guidelines. ESMA leaves to the discretion of the CRAs the identification of potential qualitative measures that are most appropriate for their validation.

Q12: 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.

78. The majority of respondents agreed that CRAs using methodologies related to creditworthiness measures other than default probabilities should use statistical measures to demonstrate the predictive power of their methodologies. However they did not recommend measures. Three CRA respondents did not agree that CRAs using credit rating methodologies related to creditworthiness measures other than default probabilities should use statistical measures to demonstrate the predictive power of their methodologies.

79. One respondent recommended for expected loss credit ratings that tests may be performed comparing assumed loss given default with the actual realised Loss Given Default (independent of the related default probability), while another respondent stated that for expected loss credit ratings, loss given default is either flat, or it increases down the rating scale, so default rates remain meaningful and consequently the respondent argued there is no need to demonstrate the predictive power of the loss given default.

80. One respondent recommended that the supervisor should provide values for the remaining parameter in Vasicek based formulas for unexpected and expected losses.

81. **ESMA's response**: Given the limited feedback and recommendations provided to ESMA on statistical measures for creditworthiness measures other than default probabilities, ESMA will not include in the proposed guidelines measures it typically expects CRAs to use where their credit rating methodologies are related to creditworthiness measures other than default probabilities. ESMA also notes that any parameters or assumptions used in the validation process should be chosen by the CRAs.

Q13: 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?

82. Respondents' views were mixed. Some respondents agreed that if ESMA establishes that there is a need for further guidance to the industry, this guidance should also cover the demonstration of predictive power of credit rating methodologies related to creditworthiness measures other than default probabilities, while others disagreed. Three respondents stated that if ESMA issues guidance, it should ensure that the guidance is sufficiently flexible and allows industry to explore alternatives for demonstrating credit ratings' predictive power or those measures to be applied on an *ad hoc* basis.

83. **ESMA's response**: While most respondents thought that any further guidance should include the demonstration of predictive power for methodologies with creditworthiness measures other than default probabilities, there was no clear consensus on which other creditworthiness measures credit ratings could refer to or what the appropriate typical measures should be in such cases. ESMA will continue to investigate and analyze this as an area of potential future guidance, but will not include measures in these proposed guidelines. For credit ratings which refer to creditworthiness measures other than default probabilities, ESMA will set out in the proposed guidelines that it typically expects a CRA to compare the expected outcome to the observed results using relevant quantitative measures and to document the rationale for its choices.

2.4 Historical Robustness

Q14: Do you agree with ESMA's view regarding the historical robustness of methodologies?

84. The majority of respondents from the CRA industry, users of credit ratings and other stakeholders agreed with ESMA's view regarding the historical robustness of methodologies.

85. Alternative views offered included two respondents who suggested that historical robustness covers all measurements that assess the performance of credit ratings based on historical data, creating overlap and incorporation with the assessment of discriminatory and predictive power.

86. Two respondents stated concern that historical robustness was harder to apply for methodologies with a low number of observations, particularly in the case of smaller CRAs.

87. One of these respondents stated there should not be excessive disclosure requirements on historical robustness.

88. One CRA respondent noted that when implementing statistical tests along all three dimensions, this created a high chance that if a test does not fail in one dimension, it will fail in one or both of the other two dimensions.

89. **ESMA's response**: ESMA will continue to use the description of historical robustness it proposed, given the general agreement for this description. While ESMA agrees that a CRA will use the historical data available to it in order to demonstrate not only historical robustness but also discriminatory and predictive power (since back-testing, which is comprised of these three dimensions, is based on historical data), ESMA believes that its description has the added benefit of encompassing tests that add value to validation but do not relate to discriminatory or predictive power.

90. ESMA acknowledges that there are challenges to validating methodologies with limited quantitative evidence, and addresses these challenges in the relevant section of the proposed guidelines and through its responses to feedback received under questions 18 – 24 in this feedback statement.

91. ESMA will not include expectations on the disclosure of validation results as it is not required by the CRA Regulation.

92. ESMA also acknowledges that a potential outcome of validation with thresholds may be that a methodology triggers a threshold along one dimension but not other dimensions. ESMA will leave it to the discretion of the CRA on what are the appropriate actions in these cases, following the application of expert judgement. Please refer to the relevant section of the proposed guidelines and questions 25 – 27 in this feedback statement.

Q15: 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?

93. The majority of respondents representing users of credit ratings, other participants and CRAs agreed with ESMA's view.

94. Where CRA respondents did not fully agree with ESMA, they supported the use of transition and default matrices to assess credit ratings performance as they provide multi-dimensional information, but did not support statistical measures of stability. The challenges raised against the measures of stability included:

   a. That the measure requires an 'original population' which CRAs may not have.

   b. Removes the relevance of other analysis of upgrade and downgrade frequency and magnitude at each credit rating level.

   c. Does not add value, as analysis of the transistion matrices allows one to understand credit rating stability at a more granular level.

95. One respondent did not support minimum measures for historical robustness, stating that there was no scientific consensus on what is appropriate. Another respondent stated that as part of assessing historical robustness, modifier data (usually represented as '+' and '–' on rating categories) should be included, and that validation should consider linkages between credit ratings, specifically the link between short term and long term credit ratings.

96. **ESMA's response**: In response to feedback, ESMA has changed statistical measures of stability (such as Population/System Stability Index) from measures that ESMA would 'typically expect' to examples of 'complementary measures'.

97. ESMA does not however share the view that the statistical measures of stability can only be performed on an 'original population', as it can also be performed on current year and previous year data, although CRAs typically develop a methodology on a development population as well.

98. ESMA notes also that statistical measures of stability are intended as measures that could be calculated not only at the credit ratings level but they may also refer to characteristics of population and their stability, thus they should not be viewed as a replacement to the analysis of transition matrices.

99. ESMA leaves it to the discretion of CRAs whether to perform their validation on a modifier level, should available data permit so.

Q16: 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.

100. The majority of respondents from the CRA industry and users of credit ratings supported ESMA's proposal without further comments. One respondent suggested

that ESMA expands its proposal to include that ESMA should develop and maintain a web-based credit rating platform that would allow users of credit ratings to compare relevant validation results of CRAs.

101. One CRA respondent stated that the use of transition matrices was a better tool for validation than the examples of complementary measures included in the Discussion Paper, because users of credit ratings were familiar with these tools, and the CRA believed it is the preferred method for assessing credit ratings stability.

102. Other CRA respondents and other market participants made caveats to the complementary measures given as examples. The majority of these comments were on benchmarking. Three respondents noted that the benefits of benchmarking between CRAs was limited because the population of rated entities and credit rating criteria are different between CRAs, and that while this type of benchmarking may be established if two CRAs have common methodologies and rated populations, it will not necessarily conclude if historical robustness has been demonstrated.

103. On benchmarking to market indicators, four respondents noted limitations due to issues such as that market indicators are normally significantly more volatile than credit ratings. One of these respondents suggested that they are better used as a point of reference that may prompt additional discussion and qualitative analysis.

104. On univariate analysis, four CRA respondents noted that univariate analysis had the limitation that it did not provide analysis of actual realised performance of the credit ratings. This included one respondent who noted that because credit ratings are not mechanical outputs from a model, it is unclear how the univariate analysis of credit rating determinants translates to an inference about the credit ratings themselves.

105. On the use of distribution analysis, two respondents noted that this would be affected by whether credit ratings are point in time or through the cycle.

106. **ESMA's response**: ESMA will continue to include these measures as examples in the proposed guidelines given the broad support for them. ESMA notes that these measures will be included in the proposed guidelines only as examples of complementary measures. ESMA does not share the view that validation by a CRA should be driven by what users are familiar with, and agrees that the use of benchmarking is a useful point of reference that may prompt additional discussion and qualitative analysis.

107. ESMA acknowledges all caveats by respondents to these examples of complementary measures and expects that where a CRA choses to use these as complementary measures, it will recognise and incorporate limitations to measures when it interprets them.

108. Per ESMA's response to other questions, the disclosure of validation results is not a regulatory requirement. Accordingly, ESMA will not develop and maintain a web-

based rating platform that would allow users of credit ratings to compare relevant validation results of CRAs.

Q17: 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.

109. The majority of CRA respondents did not recommend additional measures. One respondent who represented other market participants argued against excessive requirements on this point, as these might be an unnecessary barrier to competition.

110. Different CRA respondents suggested the following additional measures:

   a. Histogram of attributes distribution.

   b. Upgrade and downgrade rates, large rating action rates, reversible rates, rating drift and rating volatility, average transition matrices.

   c. A heat map derived from a bivariate density estimation to help visualise the patterns within a transition matrix.

   d. A measure proposed in an academic paper which summarizes rating changes and rating volatility, the frequency of upgrades and downgrades, as well as the size of the changes in ratings, in a single number.

   e. Ad hoc macroeconomic metrics or sector specific metrics.

111. Users of ratings and other market participants suggested:

   a. Monte carlo simulation for comparing performance of methodologies between different populations.

   b. Multivariate techniques such as PCA and Factor Analysis.

   c. A proprietary model.

112. **ESMA's response**: ESMA does not share the view that requirements in this area would create an unnecessary barrier to competition, given the feedback received from the CRA industry, including smaller CRAs. ESMA notes that some of the proposals are further elaborating what ESMA described as analysis of the movement of the credit ratings while proposals coming from non-CRA respondents could be technically and computionally challenging. Given the lack of consensus around potential additional measures, and the general approval for the examples of complementary measures that ESMA included in the Discussion Paper, ESMA will not include further measures in the proposed guidelines, but notes the measures suggested as measures CRAs may choose to use as additional measures and thanks respondents for their suggestions. In addition, ESMA has included some of the examples provided by the

respondents with regard to the movement analysis of the credit ratings stemming from the transition (migration) matrices.

2.5 Limited Quantitative Evidence

Q18: Do you agree with ESMA's view regarding the validation of methodologies with limited quantitative evidence?

113. A majority of CRA respondents, users of credit ratings, and also a trade association agreed with ESMA's view. One CRA respondent stated that for limited quantitative evidence, the CRA should perform and present a deep analysis of the quantitative score (used by the credit rating process) and could integrate the quantitative score by a random factor (normal distributed) of the same quantity of the qualitative contribution in the credit rating. It also suggested that where a CRA uses a mathematical score model as part of the credit rating, it could be useful to apply the quantitative measures to a benchmarking population based on the score model results, allowing a more rigorous validation approach to the quantitative component of the methodology where there is limited quantitative evidence.

114. Other CRA respondents offered caveats and observations to ESMA's view:

a. Two CRA respondents noted that where there is limited quantitative evidence to support predictive power, there was likely to be limited evidence to demonstrate discriminatory power also. Another CRA respondent stated that it would be very difficult to circumvent the challenges posed by limited quantitative evidence by any of the techniques mentioned in the Discussion Paper.

b. Another CRA respondent stated that the decision about whether there is limited quantitative evidence cannot be based on pre-defined metrics because it depends on the sufficiency of attribution, such as whether there are similar sectors with more data and the volatility of a particular sector.

c. Another CRA stated that for discriminatory power, it thought the grouping of credit rating categories and the use of an extended time period were the most useful approaches in the Discussion Paper.

d. Another CRA stated that it has limited value to generate transition matrices where there are not enough historical credit ratings to do this.

e. One CRA stated that additional analysis should be performed only in cases where a CRA identifies alternative solutions, and that additional measures should be discretionary.

f. One CRA did not agree that for sectors with few or no defaults it would be appropriate to consider credit ratings that are assigned the lowest credit rating category as instances of 'default' as it would be a circular process.

115. Other CRA respondents disagreed with ESMA's view:

   a. One CRA stated it did not believe in the utilty of enhanced data or similar measures to enable the performance of statistical tests for the validation of methodologies with limited quantitative evidence and was concerned that it could increase risks to credit ratings performance.

   b. Another CRA respondent disagreed with ESMA's approach, because it had no defaults and limited data.

   c. Another CRA suggested that qualitative elements in validation should be sufficient in limited quantitative evidence methodologies.

116. **ESMA's response**: ESMA recognizes that the validation of credit rating methodologies with limited quantitative evidence can be a difficult and challenging process. This may require a CRA to use more resourceful approaches and apply more interpretation and expert judgment than it would have to in the validation of methodologies with sufficient quantitative evidence.

117. ESMA does nonetheless expect that CRAs will make appropriate assessment and take appropriate action in order to ensure that they have made sufficient effort to perform a robust validation, particularly in the demonstration of historical robustness and discriminatory power which is a requirement under Article 8 of the RTS on rating methodologies. Based on the good practice ESMA has seen in the industry, ESMA is of the view that the techniques it is putting forward have been shown not to need significant historical data, and meaningfully address the validation challenges posed by credit rating methodologies with limited quantitative evidence.

118. Further to this, ESMA stresses that the proposed guidelines will require that the CRAs set their own criteria for when to consider a methodology as having limited quantitative evidence.

119. The measures that ESMA is putting forward are a list of non-exhaustive examples, and while ESMA believes that these examples for enhancing data are generally low risk, it is ultimately a decision for a CRA to assess whether data enhancements (including ones not noted in the proposed guidelines) are an appropriate tool to aid in the validation of their limited quantitative evidence methodologies.

120. ESMA thanks respondents for proposing potential additional measures. Per ESMA's approach towards the validation of portfolios with sufficient quantitative evidence, ESMA does not expect thresholds or pre-defined metrics to mechanistically define the next steps of validation, but rather that CRAs wil use predefined metrics, in conjunction with expert judgment to put them in their appropriate context and interpretation (please refer to paragraphs 13 – 15 of this feedback statement).

121. Given this feedback, ESMA will not change its general approach to the validation of limited quantitative evidence methodologies as proposed in the Discussion Paper.

Q19: Do you agree that CRAs should, as a first step, investigate data enhancement in validating methodologies with limited quantitative evidence?

122. There was broad agreement from all respondents to ESMA's suggestions, but with caveats on the use of some of them:

   a. Limitations around the ability of CRAs to verify third party data quality.

   b. Hypothetical transactions may behave differently than actual aged transaction data.

   c. Combining asset classes to perform joint validation assessment leaves out the specificities identified for each asset class.

   d. Most CRAs methodologies are not based on mathematical models but rather analyst-driven, so models applied to enhance data need to replicate analyst driven credit rating opinions.

   e. Two CRAs suggested further data enhancement measures: a deeper score test and applying a reduced credit rating method that only uses easily accessible and public information for the purposes of validation.

123. Where respondents disagreed with the approach put forward by ESMA, reasons given included:

   a. One CRA stated that the risks of a particular asset class/sub-asset class are unlikely to be identical, and data enhancement methods could lead to unreliable results;

   b. One CRA stated that expanding data sets and creating hypothetical transactions will not increase the number of actual credit ratings so will not enhance the predictive power of credit ratings issued by a CRA;

   c. One CRA stated that data enhancement for largely qualitively driven credit ratings, would not lead to being able to infer scores retrospectively;

   d. One CRA stated that due to its coverage of credit ratings, data enhancement was not possible now, but would be possible in 2017 when it could use the Solvency Capital Requirement to determine near defaults which would enable it to enhance data.

124. **ESMA's response**: ESMA acknowledges the limitations to potential investigative measures, but given the broad consensus for proposed investigative measures and that the requirement is only to consider investigative measures, ESMA will maintain

its proposed approach. ESMA expects that the CRAs will tackle or take into account these uncertainties when considering/applying these approaches.

Q20: 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?

125. The majority of respondents from the CRA and users of credit ratings agreed with ESMA's proposals without qualification.

126. Challenge to ESMA's proposal was provided by three CRAs to the use of a 'relaxed' default definition. One of these respondents argued it was more akin to studying the migration rates of credit ratings rather than discriminatory power of a methodology.

127. Another CRA argued that statistical tests alone were insufficient to assess discriminatory power, while a second CRA argued that for all methodologies consideration should be given to the possibility of performing statistical testing on quantitative data, provided it can be sure that the results of such testing are reliable.

128. One CRA respondent fundamentally disagreed with ESMA's proposal, arguing that Article 8 of the RTS on rating methodologies requires evidence of the discriminatory power of actual credit ratings. One CRA respondent gave strong push back to ESMA's proposal, arguing that statistical tests in the case of limited quantitative evidence were not useful, and that the results would be strongly affected by the assumptions for the data expansion and therefore not comparable with statistical results based on real credit rating actions data. The same respondent further argued that such measures have a cost burden that is not proportionate for smaller CRAs.

129. **ESMA's response**: ESMA acknowledges the limitations to potential investigative measures, but given the broad consensus for proposed investigative measures and that the requirement is only to consider investigative measures, ESMA will maintain its proposed approach. ESMA expects that the CRAs will tackle or take into account these uncertainties when considering/applying these approaches.

Q21: Do you agree that historical robustness measures should be performed when validating methodologies with limited quantitative evidence?

130. The majority of CRAs and users of credit ratings agreed without qualification that historical robustness measures should be performed when validating methodologies with limited quantitative evidence.

131. One CRA agreed but caveated that statistical tests to demonstrate historical robustnesss should only be given consideration when the results of such testing are reliable, and that where a CRA takes the view that it is not possible to perform statistical tests, the decision and rationale could be documented.

132. One CRA caveated that due to the low significance of observations, only expert judgement should be performed, and another CRA that the tests may not be possible to perform due to limited quantitative evidence.

133. Another CRA noted that while these assessments may reveal methodological deficits, this is not typically the case, and another CRA stated that for limited quantitative evidence methodologies, it will likely indicate an inability to determine the methodology's robustness due to a lack of historical experience.

134. One market participant argued against excessive requirements in this area as it may put an unnecessary barrier to competition, although this comment is made under the assumption that ESMA expected CRAs to publicly share its validation, which is not the case.

135. Another market participant was against the performance of historical robustness measures when validating methodologies with limited quantitative evidence, arguing that this should not be expected because there is no single set of historical robustness measures which are applicable for the validation of all kinds of methodologies

136. **ESMA's response**: While respondents in general agreed that historical robustness measures should be performed when validating methodologies with limited quantitative evidence, there was a noted concern about how effectively these measures could be applied and yield outcomes as part of the validation of a limited quantitative evidence methodology. ESMA will maintain its proposed approach because the caveats to the use of historical robustness measures do not detract from the benefit that consistently performing and interpreting measures for historical robustness adds to the quality and objectivity of validation.

Q22: 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?

137. The majority of respondents agreed, especially with regard to the use of transition matrices, but with caveats dependent on the type of benchmarking (similar to responses to question 16).

138. For external benchmarking to other CRAs, some CRAs and market participants stated that this should not be applicable to all CRAs because the population of rated entities and credit rating criteria varies between CRAs.

139. For external benchmaking to other market indicators, some CRAs and one market participant stated that market indicators may take into account risks other than creditworthiness, such as market risk.

140. Other CRAs noted that market indicators were typically more volatile than credit ratings and may over estimate fundamental credit risk. Another CRA noted that there

are occasionally market constraints, such as market dislocations, and at these times benchmarking to market indicators was not appropriate.

141. One CRA disagreed 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. The CRA stated that measures should not be the minimum but that CRAs should instead verify if these measures are applicable to the specific business and characteristics of the rated entities and asset classes.

142. One market participant did not agree because having minimum measures in place would suggest to users that these measures were a reliable source to evaluate the quality of the CRA, although ESMA does not suggest that these measures, or other complementary measures, are disclosed to users and recognizes that the measures will have limitations in their use.

143. Another market participant suggested a number of measures for public disclosure. This is outside the boundary of the CRA Regulation.

144. **ESMA's response**: Similarly to where ESMA has proposed the use of benchmarks for validation of methodologies with sufficient quantitative evidence, ESMA recognizes that benchmarking, whether to other CRAs or market indicators, has limitations, such as those noted by respondents, which a CRA should consider as part of the expert judgement it applies in interpreting the measures.

145. ESMA does not share the view that a CRA should verify if these measures are applicable to the specific business and characteristics of the rated entities and asset classes, given that these measures can be applied very broadly, and the consistent and systematic use of them would add rigour to the validation process.

146. ESMA will therefore maintain the approach it proposed in the Discussion Paper.

Q23: 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.

147. A number of CRAs agreed that complementary historical robustness measures add further information to the validation processes for methodologies with limited quantitative evidence without qualification. One of these CRAs suggested that distribution analysis is a useful complementary measure, and that univariate analysis should be considered with caution when the methodology has a very low number of cases.

148. Users of credit ratings and other participants agreed also. One user of credit ratings stated that measures should be disclosed, but this is not required by the CRA Regulation or proposed by ESMA.

149. Other CRA respondents agreed, particularly for the use of transition matrices and analysis. Two CRAs also noted benchmarking as a useful tool and one of these also noted the benefits of a 'logic test' where the soundess of the analytical approach is examined by comparing how a key credit rating factor is analysed in a methodology with other alternatives.

150. One CRA agreed but said the complementary measures should be a CRA's decision. Another CRA noted that backtesting was very difficult for a new methodology due to lack of available data.

151. One CRA disagreed because it did not know of any complementary historical robustness measures.

152. **ESMA's response**: There was a general consensus that complementary historical robustness measures do add further information to the validation processes for methodologies with limited quantitative evidence, although there was a divergence of views about the relative benefits of different measures. Given this feedback, ESMA will maintain its current approach that CRAs should consider complementary historical robustness measures such as those noted under the 'Historical Robustness' section.

Q24: Are there additional measures that CRAs should use when validating methodologies with limited quantitative evidence? If yes, please explain the measures and your rationale.

153. The measures proposed by this question's respondents where as follows:

    a. One CRA argued that the binomial and HL test can also be run on methodologies with limited quantitative evidence .

    b. One CRA suggested the idea of taking credit scores and a random factor.

    c. One CRA suggested the use of the Solvency Capital Ratio, which determines the solvency of insurance companies, as an additional measure to validate a methodology.

    d. One CRA and one user of credit ratings suggested the use of transition matrices.

    e. One CRA suggested that it had found stochastic sensitivity analysis for its limited quantitative evidence methodologies helpful in understanding the sensitivity of a methodology.

    f. One other participant suggested that for limited quantitative evidence the key to validation was analysing the evidence of individual credit ratings and suggested a proprietary model to analyse this.

154. **ESMA's response**: ESMA appreciates the input and recommendations of respondents in proposing additional measures. Given i) the wide range of

recommendations, ii) the support for measures already proposed by ESMA as examples, and iii) that there was no consensus around proposed additional measures, some of which were specific to a CRA's own methodologies and credit rating philosophy, ESMA will not include further examples of additional measures in the proposed guidelines. ESMA however notes the measures suggested as measures CRAs may choose to use as additional measures and thanks respondents for their suggestions.

2.6 Identifying and addressing anomalies

Q25: Do you agree that thresholds should be set for the quantitative validation techniques?

155. Most CRA respondents agreed without qualification, as did the users of credit ratings who replied.

156. Some respondents agreed but with caveats about how thresholds should be applied while others disagreed:

   a. One trade association and one market participant said thresholds should not be harmonized across CRAs.

   b. Another CRA said that ESMA should fix the database on which tests are run and another CRA that thesholds should be consistent across CRAs.

   c. Two CRAs stated that thresholds should not lead to mechanical evaluations and actions, but be used to inform validation.

   d. Two CRAs disagreed that thresholds should be applied, arguing that results across CRAs would not be comparable and one of these CRAs argued that quantitative measures would be better interpreted ad hoc.

   e. One other participant disagreed with this position, stating that there should be no obligation until there is an agreement among scientists and other experts about the meaningfulness of the validation technique.

157. **ESMA's response**: While a number of respondents made caveats to the use of thresholds, there was general support for this approach and a recognition that it adds to the objectivity of validation. ESMA agrees that thresholds should not lead to mechanical evaluations and actions, but should be used to inform validation, and the documentation of how thresholds will be applied is an important element in maintaining the objective use of thresholds.

158. ESMA does not intend to harmonise the use of thresholds across CRAs, as thresholds should reflect the asset classes, methodologies and type of credit ratings that a CRA issues. The use of a fixed database would detract from this approach. Under the

approach proposed by ESMA, while validation results would not be comparable across CRAs (which is not the objective of these proposed guidelines), they would add a challenging, consistent and objective component to a CRA's internal validation.

159. ESMA does not share the view that no benefit can be derived from the use of thresholds until there is complete consensus agreement among experts.

160. Given this feedback, ESMA will maintain its current approach that CRAs should internally set thresholds for their quantitative validation techniques in order to identify and address potential anomalies highlighted by back-testing.

Q26: Do you agree that the Internal Review Function should decide on these values?

161. Most CRA respondents agreed, and two users and one other market participants agreed without qualification that the Review Function should decide on these values.

162. One trade association stated that thresholds should be set according to each CRAs' procedures. One CRA suggested that all parties involved in the methodology/criteria development process should be involved in the assessment.

163. One other participant did not see the added value of ESMA giving guidance on this point while another another participant disagreed that the Internal Review Function should decide on these values, but did not offer an alternative approach.

164. **ESMA's response**: Given the general support for ESMA's proposal, ESMA will maintain its approach that the Review Function should decide on these values.

165. ESMA is of the view that the setting of thresholds by persons involved in credit rating activities may impair the independence of this process.

Q27: Do you agree that predefined actions should be documented by CRAs for when the thresholds are met?

166. Most CRA respondents and users of credit ratings agreed that predefined actions should be documented by CRAs for when the thresholds are met.

167. Other CRA respondents agreed but with caveats and observations:

   a. Five CRAs and one trade association stated that actions should not be mechanistic, or automatically trigger the modification of the methodology/model.

   b. One CRA stated that predefined actions should be tailored to each specific CRA.

168. Two other market participants disagreed; one without further elaborating, and another because it assumed predefined actions would be publicly disclosed and doubted the value added by disclosures.

169. **ESMA's response**: Given the general support for ESMA's proposal, ESMA will maintain its approach. ESMA stresses that the intent for predefined actions is on the one hand to be tailored to each specific CRA and on the other hand not to create mechanistic changes to methodologies but rather document actions in advance for increasing the objectivity and consistency of validation.

3 Annex: Opinion of the Securities and Markets Stakeholder Group

The opinion of the Securities and Markets Stakeholder Group[6] is as follows:

Executive Summary

The Securities and Markets Stakeholder Group is an advisory group, and we do not have the technical expertise to answer the consultation in detail. However we feel it is important that when considering the consultation responses that ESMA keeps to mind the wider context of maintaining market integrity and protecting investors and therefore make the following high level response to its Discussion Paper on the validation and review of Credit Rating Agencies' methodologies.

The validation of credit ratings cannot be considered in isolation. Any assessment must be set in context of the circumstances under which it was applied and take on-board any influence and/or bias which may have occurred as a result of the fees paid to the ratings agency for the specific rating or indeed ancillary services.

As the SMSG has stated previously we believe that the arrival of the European Ratings Platform (ERP) will greatly assist not only EMSA but interested third parties including academics and journalists in identifying possible anomalies in methodologies as well as in their application.

Checking that a methodology is valid will not in and of itself protect investors. The transparency that the ERP will provide not only on the performance of individual ratings but also on fees and fee arrangements will help highlight where and when there are problems with the application of any specific methodology.

Background

Many academic (1), regulatory (2) and government (3) investigations into the origins of 2008 Global Financial Crisis have identified that it was the improper application of credit rating methodologies and not the methodologies themselves which precipitated the US sub-prime mortgage backed securities debacle which set off the wider crisis.

We believe that it is important to learn this important lesson from recent financial history when seeking to ensure that the ratings issued by credit rating agencies regulated and authorised within the European Union are sound and fit for purpose.

1) Lawrence J. White, Credit Rating Agencies and the Financial Crisis: Less Regulation of CRAs Is a Better Response www.stern.nyu.edu/sites/default/files/assets/documents/con_039549.pdf

2) The Turner Review A regulatory response to the global banking crisis www.fsa.gov.uk/pubs/other/turner_review.pdf

---

[6] https://www.esma.europa.eu/sites/default/files/library/2016-smsg-011_smsg_advice_on_validation_of_cras_methodologies.pdf

Council on Foreign Relations: The Credit Rating Controversy www.cfr.org/financial-crises/creditrating-controversy/p22328

3) Final Report of the National Commission on the Causes of the Financial and Economic Crisis in the United States

www.gpo.gov/fdsys/pkg/GPO-FCIC/pdf/GPO-FCIC.pdf

Monitoring fees, incentives and accuracy

ESMA by necessity a risk-based regulator and cannot be expected to sign off every single rating methodology created by a CRA. Rather it should target its resources to examining methodologies which have been brought into question by either a low accuracy ratio or where the financial relationship between CRA and company seeking the rating warrant further investigation.

As stated above in the executive summary the SMSG believes that the arrival of the European Ratings Platform in July 2016 and the transparency it will bring to the practices, remuneration and performance of CRAs will be an invaluable tool in helping ESMA identify which methodologies it needs to target with its limited resources.