|  |
| --- |
| 17 November 2015 | ESMA/2015/1736 |

|  |
| --- |
| Reply form for the Discussion Paper on the validation and review of Credit Rating Agencies’ methodologies |
|   |

|  |
| --- |
| Date: 17 November 2015 |

Responding to this paper

The European Securities and Markets Authority (ESMA) invites responses to the specific questions listed in the ESMA Discussion Paper on the validation and review of Credit Rating Agencies’ methodologies, published on the ESMA website.

*Instructions*

Please note that, in order to facilitate the analysis of the large number of responses expected, you are requested to use this file to send your response to ESMA so as to allow us to process it properly. Therefore, ESMA will only be able to consider responses which follow the instructions described below:

* use this form and send your responses in Word format (pdf documents will not be considered except for annexes);
* do not remove the tags of type <ESMA\_ QUESTION\_VR\_CRA\_1> - i.e. the response to one question has to be framed by the 2 tags corresponding to the question; and
* if you do not have a response to a question, do not delete it and leave the text “TYPE YOUR TEXT HERE” between the tags.

Responses are most helpful:

* if they respond to the question stated;
* contain a clear rationale, including on any related costs and benefits; and
* describe any alternatives that ESMA should consider

**Naming protocol**

In order to facilitate the handling of stakeholders responses please save your document using the following format:

ESMA\_VR\_CRA\_NAMEOFCOMPANY\_NAMEOFDOCUMENT.

E.g. if the respondent were ESMA, the name of the reply form would be:

ESMA\_VR\_CRA\_ESMA\_REPLYFORM or

ESMA\_VR\_CRA\_ESMA\_ANNEX1

To help you navigate this document more easily, bookmarks are available in “Navigation Pane” for Word 2010 and in “Document Map” for Word 2007.

***Deadline***

Responses must reach us by **19 February 2016.**

All contributions should be submitted online at [www.esma.europa.eu](http://www.esma.europa.eu) under the heading ‘Your input/Consultations’.

***Publication of responses***

All contributions received will be published following the end of the consultation period, unless otherwise requested. **Please clearly indicate by ticking the appropriate checkbox in the website submission form if you do not wish your contribution to be publicly disclosed. A standard confidentiality statement in an email message will not be treated as a request for non-disclosure.** Note also that a confidential response may be requested from us in accordance with ESMA’s rules on access to documents. We may consult you if we receive such a request. Any decision we make is reviewable by ESMA’s Board of Appeal and the European Ombudsman.

***Data protection***

Information on data protection can be found at [www.esma.europa.eu](http://www.esma.europa.eu) under the headings ‘Legal notice’ and ‘Data protection’.

# Introduction

Please make your introductory comments below, if any:

< ESMA\_COMMENT\_VR\_CRA\_1>

Fitch appreciates ESMA’s open dialogue on the validation and review of Credit Rating Agencies’ methodologies, a subject which both ESMA and Fitch agree is the cornerstone of high-quality ratings. Rightly so, it is one of ESMA’s priorities to enhance the criteria validation process across the industry and Fitch agrees with ESMA that some of the ideas suggested in this validation paper will enhance that process.

European and US regulation has led to important changes that have enhanced the robustness, independence, transparency, and governance of rating agencies. This has substantially strengthened policies, procedures, organizational structure, and the industry as a whole. Critically these regulator-led changes have not materially altered the rating product. ESMA’s definition of predictive power would change the credit rating product.

In its Executive Summary, ESMA quotes the regulation saying ‘Credit rating agencies should use rating methodologies that are rigorous, systematic, continuous and subject to validation including by appropriate historical experience and back-testing.’ Policymakers foresaw a possibility that regulators could encroach on CRA’s methodologies when implementing this legislation. To prevent this from happening, policymakers explicitly state in the remainder of the paragraph that ‘such a requirement should not, however, provide grounds for interference with the content of credit ratings and methodologies by the competent authorities and the Member States.’ ESMA’s definition of predictive power changes the definition of a rating and requires changes in methodologies. Fitch sees this as interfering with the content and meaning of credit ratings. Whilst Fitch thinks this is unlikely to be ESMA’s intention, it would be a consequence.

Fitch’s existing ratings definitions explicitly state that “*credit ratings express risk in relative rank order, which is to say they are ordinal measures of credit risk and are not predictive of a specific frequency of default or loss*”. Assigning ‘expectations (absolute numbers of ranges) per rating category’ would be in conflict with this definition and so would require a change to our product.

Fitch’s existing criteria assess the vulnerability of a transaction or entity to credit risk, but does not assign a probability of that event happening. To do this Fitch would need to switch to a cardinal scale of default, which would be inherently more volatile. Transactions and entities would have to be upgraded or downgraded based on the probability of an event happening rather than their vulnerability to that event.

For example, Fitch’s existing Residential Mortgage Backed Securities (RMBS) methodology assesses the vulnerability of RMBS securities to a housing crash. ‘Bsf’ rated RMBS bonds are more vulnerable to a housing crash than ‘AAAsf’ rated RMBS bonds. This relationship holds regardless of Fitch’s view on house prices and delinquencies in the next 12 months. If the ‘Bsf’ one-year default band was set at between 1% and 4% and there was not a housing crash then it is likely there would be zero ‘Bsf’ defaults. This would ‘fail’ the quantitative threshold because it is outside of the 1% to 4% range. In order to ‘pass’ the quantitative threshold Fitch would need to alter its criteria to assess the likelihood of a housing crash materialising. The ratings process would then need to upgrade bonds in rising house price environments in order for the ratings to produce ‘Bsf’ default rates of between 1% and 4%. Fitch doubts this is possible. Even if it was possible, it would lead to highly cyclical ratings.

If ESMA chooses to force on all CRAs its definition of ratings being predictors of a default rate – or even a range of default rates - then ESMA would create a perception of precision where Fitch never intended it to exist. The perception of precision in ratings has proven to create over-reliance on ratings in the past.

**The Role of Quantitative Metrics**

The criteria validation process, similar to the rating process, is inherently a qualitative judgement. In fact, this is one of the primary distinguishing features between a rating and a credit scoring process. The qualitative element of the criteria review process is reflected in the seniority of our Independent Review Function.

Notwithstanding the importance of qualitative factors in Fitch’s review process, Fitch agrees with ESMA that additional quantitative metrics could enhance the criteria validation process if the quantitative metrics measure the factors the CRA strives to achieve. The best way to achieve this is for the quantitative metrics to reflect the qualitative decision making process. This would then provide a form of quantitative framework for the qualitative decision making process.

The focus of the Criteria Officers’ analysis of ratings performance is on the historical rating transition. Fitch produced transition and default studies before it was required by regulation because it sees real value in these studies.

**Discriminatory Power, Predictive Power and Historical Robustness**

Fitch is aware of ESMA’s desire to analyse three separate performance indicators, one for each of discriminatory power, predictive power and historical robustness. In Fitch’s opinion these elements should not be looked at in isolation because of the large degree of overlap.

Individual tests may cover one, two or even all three of these factors. The aggregate of all the measurements and the final conclusion of the review will cover discriminatory, predictive power and historical robustness holistically. For example, it is clear that tests that measure the ordinal nature of defaults are primarily discriminatory tests. However, they are also historical robustness tests in that they look at defaults over time.

The focus of Fitch’s response below is on predictive power. We believe this term is not appropriate for ratings. Our understanding of ESMA’s definition leads us to conclude that the best way of demonstrating predictive power is via ratings transition studies.

< ESMA\_COMMENT\_VR\_CRA\_1>

1. Do you agree with ESMA’s view regarding the discriminatory power of methodologies?

<ESMA\_QUESTION\_VR\_CRA\_1>

Fitch agrees with ESMA’s definition that the discriminatory power of a methodology relates to its ability to rank order the rated entities in accordance with their future status (defaulted or not defaulted) within a predefined time horizon.

<ESMA\_QUESTION\_VR\_CRA\_1>

1. 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?

<ESMA\_QUESTION\_VR\_CRA\_2>

Fitch agrees that an Accuracy Ratio is a good minimum standard for demonstrating discriminatory power across a CRA’s ratings. It provides a starting point to the analysis rather than the final objective. If the Accuracy Ratio is not within the expected bounds then it provides grounds for additional investigation. It does not mean there is something wrong with the ratings.

Fitch is already working on introducing an Accuracy Ratio for its 2015 Transition and Default study. We believe it will be useful at an asset class level, for example non-financial corporates, financial institutions, or structured finance. The tool cannot be used to draw meaningful conclusions from populations of ratings with few default events.

<ESMA\_QUESTION\_VR\_CRA\_2>

1. 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.

<ESMA\_QUESTION\_VR\_CRA\_3>

We do not think complementary measures are always necessary. If the results from the Accuracy Ratio indicate that there could be a performance issue then at that point it may be necessary to apply additional measures. These could include the measures ESMA describes above, but it will depend on the asset class and the reason for the performance issue.

<ESMA\_QUESTION\_VR\_CRA\_3>

1. 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.

<ESMA\_QUESTION\_VR\_CRA\_4>

We think that an Accuracy Ratio that all of the CRAs apply combined with further measures specific to each CRA is the best solution. Fitch does not believe ESMA should prescribe additional measures.

<ESMA\_QUESTION\_VR\_CRA\_4>

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

<ESMA\_QUESTION\_VR\_CRA\_5>

Qualitative measures are an essential part of Fitch’s criteria review process. However, these qualitative measures are adapted for specific asset classes and criteria. It is not appropriate to standardise these because different asset classes lend themselves to analysis in different ways. We do not believe ESMA should prescribe additional measures that would be required for all sectors and all points of the credit cycle.

<ESMA\_QUESTION\_VR\_CRA\_5>

1. Do you agree with ESMA’s view regarding the predictive power of methodologies?

<ESMA\_QUESTION\_VR\_CRA\_6>

No. The proposal is flawed in that it does not take into account the nuances in the credit rating product that CRAs and investors have understood for generations.

The regulation demands that criteria are ‘continuous, systematic and robust’. The reference to credit rating criteria being ‘predictive’ was introduced by ESMA in an RTS. Fitch believes that the RTS should more accurately reflect the credit ratings product, or at a minimum recognize not all CRAs have leapt to a belief that they can accurately offer defined probabilities of default in their product.

If ESMA declines to change its RTS then Fitch believes ESMA should define ‘predictive’ to more accurately reflect credit ratings. A credit rating is an opinion about the ‘vulnerability’ of an entity to default. Fitch is ‘predicting’ that entities that are more vulnerable to default will default more frequently over the long-term. This definition is already captured in the discriminatory tests. Fitch is not – and has never – predicted specific default rates or ranges of default rates.

Fitch’s rating definitions and ratings process reflect the ordinal nature of our credit ratings. Fitch’s individual rating committees assess the relative vulnerability of entities against peers rather than against a forecast default rate.

ESMA states that users of credit ratings expect ratings ‘to meet creditworthiness-related expectations’… ‘derived from CRAs’ historical performance.’ ESMA’s interpretation is that investors are only interested in the default rate associated with each rating category. Fitch’s experience is that investors are more interested in the stability of ratings. For example, investors have come to expect stability from AAA ratings, whereas they would expect more ratings volatility from B ratings.

The stability of ratings often drives their investment decisions because downgrades – short of a default – can have a significant impact on their investment returns. For example, a rating downgrade could result in an entity being excluded from an index, or an investment guideline, both of which could mean the investor has to sell the bond at a loss.

There is so little data on defaults in most asset classes that if ESMA would like a band measurement that would trigger additional analysis of a criteria then this band should be based on ratings transitions. This more populous data set is less likely to be overly influenced by an idiosyncratic event. It will also enable standard tests to be applied to a far wider set of criteria, many of which have experienced very few defaults.

We believe it would be more useful for the CRA validation process if the predictive measures ESMA prescribes are focused on ratings transition rather than default.

<ESMA\_QUESTION\_VR\_CRA\_6>

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

<ESMA\_QUESTION\_VR\_CRA\_7>

Fitch disagrees with introducing any statistical measures to test rating performance against observed default rates because neither our criteria nor our rating process are designed to produce specific default rates.

Fitch does believe that a statistical measure that looked at the performance of ratings against the expected rating transition rates could enhance the validation process.

<ESMA\_QUESTION\_VR\_CRA\_7>

1. 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?

<ESMA\_QUESTION\_VR\_CRA\_8>

No. Fitch agrees that statistical measures should be used, but believes that each individual CRA should be able to conduct its own tests to reflect its methodology, rating definitions and rating performance.

It is appropriate for ESMA to pursue standardisation around the Accuracy Ratio for discriminatory power because it is a widely used measure that is understood by investors. The predictive power tests are for internal use only and therefore there is no market benefit to standardisation, especially considering that these tests were not designed for assessing credit ratings.

<ESMA\_QUESTION\_VR\_CRA\_8>

1. 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.

<ESMA\_QUESTION\_VR\_CRA\_9>

No. It should then be left to individual CRAs to determine what tests to use.

<ESMA\_QUESTION\_VR\_CRA\_9>

1. 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.

<ESMA\_QUESTION\_VR\_CRA\_10>

Fitch considers statistical measures that assess the ordinal nature of credit ratings and the performance of ratings across the rating spectrum to be more appropriate.

<ESMA\_QUESTION\_VR\_CRA\_10>

1. 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.

<ESMA\_QUESTION\_VR\_CRA\_11>

Qualitative measures are an essential part of Fitch’s criteria review process. However, these qualitative measures are adapted for specific asset classes and criteria. We do not believe ESMA should prescribe additional measures that would be required for all sectors and all points of the credit cycle.

<ESMA\_QUESTION\_VR\_CRA\_11>

1. 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.

<ESMA\_QUESTION\_VR\_CRA\_12>

TYPE YOUR TEXT HERE

<ESMA\_QUESTION\_VR\_CRA\_12>

1. 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?

<ESMA\_QUESTION\_VR\_CRA\_13>

TYPE YOUR TEXT HERE

<ESMA\_QUESTION\_VR\_CRA\_13>

1. Do you agree with ESMA’s view regarding the historical robustness of methodologies?

<ESMA\_QUESTION\_VR\_CRA\_14>

Fitch disagrees with ESMA’s definition of historical robustness as being ‘other dimensions that do not relate to [the methodology’s] discriminatory power or predictive power’. We consider historical robustness to cover all measurements that assess the performance of ratings based on historical data, especially longer time series of data. This definition creates a large degree of overlap in the meanings of discriminatory, predictive and historically robust. For example, Fitch considers ratings stability to offer information about the predictive power and historical robustness of ratings.

The requirement of an annual criteria review means that the focus of the discriminatory and predictive tests is the one-year performance, while considering longer term performance and experience. Fitch believes that historical robustness refers to the need to look further back into the historical performance. Tests (including discriminatory or predictive power tests) that assess the performance of ratings over multiple years and through different points in the economic and credit cycle would be historical robustness tests.

The purpose of these longer-dated tests is to make sure that there are not constant changes in criteria that allow for the annual review to be satisfied but miss a gradual deterioration of performance over a longer term horizon.

<ESMA\_QUESTION\_VR\_CRA\_14>

1. 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?

<ESMA\_QUESTION\_VR\_CRA\_15>

Fitch uses transition and default matrices to assess ratings performance and finds this the most useful measure to demonstrate criteria sufficiency. Transition matrices provide multi-dimensional information across the ratings spectrum. ESMA’s proposal of summarising the upgrade / downgrade frequency and magnitude at each rating level with one number produced by a stability index score will remove most of the relevance of this analysis.

<ESMA\_QUESTION\_VR\_CRA\_15>

1. 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.

<ESMA\_QUESTION\_VR\_CRA\_16>

No. We do not think these are appropriate. Users of credit ratings are familiar with transition and default matrices and we believe this is the preferred method for assessing ratings stability.

<ESMA\_QUESTION\_VR\_CRA\_16>

1. 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.

<ESMA\_QUESTION\_VR\_CRA\_17>

Qualitative measures are an essential part of Fitch’s criteria review process. However, these qualitative measures are adapted for specific asset classes and criteria. We do not believe ESMA should prescribe additional measures that would be required for all sectors and all points of the credit cycle.

<ESMA\_QUESTION\_VR\_CRA\_17>

1. Do you agree with ESMA’s view regarding the validation of methodologies with limited quantitative evidence?

<ESMA\_QUESTION\_VR\_CRA\_18>

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

It has been ESMA’s wish that criteria become increasingly specific using published assumptions for different asset types. The fine tuning of criteria to a finite degree often reduces the sample population, to which it can apply, reducing the reliability of statistical measures as proposed in this discussion paper.

<ESMA\_QUESTION\_VR\_CRA\_18>

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

<ESMA\_QUESTION\_VR\_CRA\_19>

Fitch agrees that combining asset classes or sub-asset classes with similar risk characteristics in order to perform joint validation assessments is a reasonable proxy. Fitch disagrees that creating model-driven ratings or hypothetical transactions is a good proxy. The performance of these hypothetical ratings will not be the same as an actual rating and can therefore not be relied upon.

<ESMA\_QUESTION\_VR\_CRA\_19>

1. 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?

<ESMA\_QUESTION\_VR\_CRA\_20>

Fitch agrees that combining rating categories or using an extended time frame may provide some additional information. It is more difficult to use a definition of default that is another rating category (eg BBB rather default (D). Default is not an opinion about the future state of a credit but rather an observable fact. Measuring the accuracy of an opinion against an observable fact (default) is a more objective test than assessing the performance against another rating category.

<ESMA\_QUESTION\_VR\_CRA\_20>

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

<ESMA\_QUESTION\_VR\_CRA\_21>

It may not be possible to perform these tests because of the limited quantitative evidence.

<ESMA\_QUESTION\_VR\_CRA\_21>

1. 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?

<ESMA\_QUESTION\_VR\_CRA\_22>

Fitch agrees that transition matrices are the minimum measures that should be used for validation of methodologies. Fitch does not agree that benchmarking to other instruments is consistently appropriate or necessary.

Fitch and investors expect ratings and market-based measures of credit performance, such as CDS spreads, to diverge because the market-based measures are typically more volatile than credit ratings. CDS spreads and bond yields can provide insight into directional trends and the health of the credit markets, which in some cases has served as a useful early warning signal of faltering credits. However, spreads have a tendency to ramp up during periods of market volatility, only to revert to prior levels once sentiment improves. In these instances, elevated spreads overestimated fundamental credit risk and would provide an incorrect calibration point for ratings.

<ESMA\_QUESTION\_VR\_CRA\_22>

1. 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.

<ESMA\_QUESTION\_VR\_CRA\_23>

In some cases complementary historical robustness measures may add further information, however, it is not possible to define what data or measures need to be used when there is limited quantitative evidence. The CRA should use the data and measures it thinks is required to produce ratings that are sensible predictors of creditworthiness.

<ESMA\_QUESTION\_VR\_CRA\_23>

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

<ESMA\_QUESTION\_VR\_CRA\_24>

TYPE YOUR TEXT HERE

<ESMA\_QUESTION\_VR\_CRA\_24>

1. Do you agree that thresholds should be set for the quantitative validation techniques?

<ESMA\_QUESTION\_VR\_CRA\_25>

Yes. Fitch believes quantitative metrics can play an important role in providing a framework for a qualitative debate. However, the qualitative debate must remain the most important factor in the review process. Thresholds should simply draw attention to ratings performance that appears out of line. In the vast majority of cases it is expected that the rating performance is explainable and no further action would be required.

<ESMA\_QUESTION\_VR\_CRA\_25>

1. Do you agree that the Internal Review Function should decide on these values?

<ESMA\_QUESTION\_VR\_CRA\_26>

Yes.

<ESMA\_QUESTION\_VR\_CRA\_26>

1. **Do you agree that predefined actions should be defined by the CRAs when the thresholds are met?**

<ESMA\_QUESTION\_VR\_CRA\_27>

A threshold breach should be reported and documented, but it is not possible to predefine what to do if one metric breaches before the reason for that breach is known. There are simply too many unforeseeable reasons that could trigger a breach.

Upon a breach a member of the Independent Review Function (Criteria Officer in the case of Fitch) should decide what course of action to take (including taking no action). The action (or no action) and the reason for the action should also be documented.

<ESMA\_QUESTION\_VR\_CRA\_27>