EvD Annual Evaluation Review (AER) 2012
Management Comments

Summary

- Management welcomes the detailed analysis of performance in Annex 3. The performance analysis is valued as very important for learning and focusing management efforts on its improvement. Management would caution however about the ability to draw conclusions based on limited and sometimes non-representative data as also confirmed in the report (e.g. in the discussion of environmental performance). Management proposes ways of lessening the shortcomings of limited data and improving the methodology of analysing the most important factors that determine performance in a more rigorous way.

- The analysis on the declining performance trends on additionality suggest that it could be best to focus on a narrow definition of financial additionality only due to methodological issues related to verification of non-financial additionality under the current approach. Management has strong reservations about ignoring other aspects of EBRD value-added that form part of the Bank's non-financial additionality. Further improvements in understanding and application of the current broader concept of additionality (including non-financial additionality) are however desirable, through clear guidance on the specific criteria on how to assess additionality.

- Management appreciates the thematic discussion of the main findings and recommendations from evaluations in 2011. Management remarks on the specific findings of individual studies have been already expressed in the relevant comments when the studies were completed, and in the "Follow-up of EvD recommendations". The main theme running through most of the findings from evaluation in 2011, the Results presentation was the focus of the recent "EvD Special Study –Performance Metrics-How well do EBRD projects specify expected results?" Management has provided extensive comments to this study (SGS13-113)and looks forward to the joint actions and discussions envisaged in the comments

- Management confirms that it is still too early to make any meaningful observations about the results of the new approach, the self-evaluation based on Operation Performance Assessments (OPA) format. Management notes that Banking is asked to adopt retrospectively the EvD proposed project results presentation under the "Performance Metrics" study which is not yet full agreed, and queries to what extent the gap in rating may have been adversely affected by this retrofitting.

- Management has discussed with EvD the shortcomings of the comparative analysis for project performance with TIMS information in AER 2011. The focus of the report is on the predictive power of on TI potential and risk rating on the ex-post EvD ratings. Using the same EvD dataset, the analysis done under the Results task force TIMS working group reinforces some points made tentatively by the EvD in the AER report, while in others the findings are different. This analysis confirms that the transition risk ratings at project inceptions are important. That justifies the TIMS paper recommendation that the scorecard target for the new projects should be based

on both TI potential and risks and not only on TI potential as it is currently the case. On the other hand, contrary to the key message of AER, the TIMS WG analysis concludes that the power of the ex-ante transition impact potential ratings to predict EvD ex-post ratings is very strong when controlling for risks.

1. **Performance analysis-Chapter 2 and Chapter 4**

Management values the overall performance discussion in Chapter 2 and welcomes the more detailed analysis in Chapter 4 and Annex3 on the trends identified for further review in the Management Comments on AER 2011. This analysis is important for understanding, learning and focusing Bank's efforts on the most important factors that affect performance.

Management would caution however that special care is warranted and appropriate qualifiers needed when generalising results of a sample analysis into far reaching conclusions about the performance at the institutional level or for different regions, dimensions etc. Special attention is needed to the sample representation across many dimensions and varieties of the Bank activities (country, sector, etc.)[1], and to the methodology used to analyse the most important factors that affect performance.

Data limitations: Given the limited number of annual evaluations, the statistics based on the three year moving average help to see the trend in total portfolio performance, both overall and on specific sub-dimensions, although its smoothing effect has its limitations in "detecting" cases of significant change, such as that of the financial crisis.[2] The limited data observations becomes more important when the data is used to analyse the year on year changes and even trends in a sample break-down, e.g. by regions or sector groups.

These shortcomings are revealed in the relevant and helpful analysis of the performance of projects in CA, where some conclusions are based on the results of an insignificant number of evaluations. Also, the report confirms and Management shares the reflection in the document that year on year changes in environmental and social performance, (and their change) is not considered significant due to the very small sample of projects under evaluation in any one year. One cannot generalise or speculate on why this may (or may not) be the case on the basis of limited data that do not reflect the full portfolio (there were no evaluations of Category A projects, for example, for several years).

Given these data limitations, Management believes that the performance analysis could be focused on changes over longer time periods. For example, an important question such as the impact of the financial crisis on performance could be better analysed in a rigorous and comprehensive way by comparing the evaluation results of two larger samples containing five year data, for each period before and after the crisis. The larger data size would enable to analyse simultaneously the effect of different factors, controlling for sector and regional characteristics. The comparative analysis by sector/region for a five year period /before and after the crisis could also be presented graphically.

---

[1] In Annex 1, Table 6 shows EvD sample representation by business volume which is interesting in pointing out areas of under or over representation. A similar table indicating the sample representation by number of projects may presents a different picture as project size maybe typically higher for some countries, such as Russia, or types of projects, such as sovereign projects.

[2] It would be helpful to show the moving average by year of completion as well as year of approval. This might help to identify the underlying reasons for the change, if any, in the overall performance and different sub-dimensions of portfolio of projects.

For instance, are the changes in performance between CA and other regions statistically significant prior and post crisis period? Has the performance of projects in CA changed since the creation of the ETCI? The pre-2004 sample includes 30 projects and the post-2004 sample includes 17 projects? The Bank has been aware of the project underperformance and has reacted to improve results. The Bank has materially improved its ability to deliver transition and project success since the creation of the ETC Initiative in 2004. The Evaluation Department ratings for EBRD projects in the ETC Initiative Central Asia countries (e.g., Kyrgyz, Mongolia, Tajikistan, Turkmenistan and Uzbekistan) have improved substantially during the ETC Initiative, both on the overall performance and the sub-dimensions.

Analysis Methodology: while the detailed analysis is important and contains interesting findings, some results are not well supported by the underlying analysis. Management believes that a more rigorous and comprehensive analysis is necessary to analyse the factors that explain the decline in overall performance indicators as well as performance across different dimensions or for a particular region.

The assessment of the performance trends in CA is important and relevant. It highlights the special conditions in Central Asia which makes these countries a higher risk environment in all respects. By going into depth, the analysis correctly points out that "Central Asia countries have been unfavourably compared with those of EEC...” It also points out the external factors "such as government interference or lack of an enabling business environment..." and concludes that the Central Asia region compares reasonably within the ETC Initiative countries. However, these factors are taken separately based on findings of specific evaluations. The effect of different factors such as country and sector specific, including structural, institutional and political could be analysed together using for example the EBRD transition and political indicators. Perhaps the list of external factors identified by EvD should also include the extremely low institutional capacity and poor understanding of corporate governance, which prevails both in the private and public sectors, and country specific political risk factors, including revolutions, wars, government changes, etc… that have been prevalent over the past 20 years.

Similarly, while the identified impact of the financial crisis on the decline in overall performance of the 2005 and 2007 cohorts is reasonable, it is not firmly based on a rigorous analysis. The analysis also suggests that other factors were responsible for those reports cited as “unsuccessful” (e.g. poor management etc.). But the crisis may have also influenced these factors. For instance, latent management weaknesses which might not prove fatal in a more benign atmosphere may be exposed, tested, and found wanting in crisis conditions. Table 1 in Chapter 2, with correlations between different performances indicators, shows that indeed achievements of operation objectives, transition impact and project financial performance are the most significant predictors of the overall performance of a project. However, all of these three performance dimensions are also very highly correlated with each other. A rigourous analysis is necessary to try to determine the main factor that affects project performance. Is it the project financial performance that determines success in transition and operational objectives? Has this changed before and after the crisis? These questions are important for learning and to help focus our efforts in project structure and during project implementation.

Additionality: The analysis on the performance trends on additionality, in particular the falling rate of “fully verified” projects by EvD, suggests it could be best to focus on a narrow definition of financial additionality due to methodological issues related to verification under the current approach. Management has strong reservations about ignoring other aspects of

EBRD value-added that form part of the Bank's non-financial additonality (please refer to the Annex for a more detailed discussion).

Indeed, financial additionality that includes Bank's unique terms of financing and the degree of resource mobilization is less contestable and "easier" to validate. Non-financial additionality, however, relates to other important elements of value–added the Bank can contribute as an international organisation, such as:

- reducing credit and financial risks and enhancing project sustainability through project structure (including through its influence on the regulatory, legal and political environment), and providing comfort against other investor's perception of risks and encouraging them to invest –the catalytic effect
- Introducing/ensuring improvements in standards, such corporate governance, environment, procurement etc.

Further improvement in understanding and application of the concept of non-financial additionality may be helpful, through clear guidance on the specific criteria on how to assess additionality.

In addition, it seems that the argument about a declining trend or a significant decrease in the ratio of projects where additionality is "fully verified" is not strongly backed by the statistical data presented. Chart 3 in Chapter 2 indicates that apart from the period 92-95 the share of project where additionality is "fully verified" is largely stable in the range 50%-60% (the slight year on year difference may be due to change in sample size).

2. Findings from EVD evaluation in 2011 - Chapter 3

Management appreciates the overview of the findings from evaluations in 2011. In particular, the report makes some interesting and valid recommendations, including on the need to sharpen our leverage on policy dialogue and relying more on Integrated Approaches. A thematic discussion of the main findings and recommendations is valuable.

Management remarks on the specific findings of individual studies have been already expressed in the relevant comments when the studies were completed, and in the "Follow-up of EvD recommendations". Management is therefore focusing on the main theme running through most of the findings from evaluation in 2011. Results presentation or "lack of clarity regarding the expected outcomes and the links between physical objectives and wider aims" are a recurring theme in the findings of the studies including for projects, TC, LTT programme, initiatives and country strategies. This issue was the focus of the recent "EvD Special Study –Performance Metrics-How well do EBRD projects specify expected results?" Management has provided extensive comments to this study (SGS13-113) and looks forward to the joint actions and discussions envisaged in the comments

3. Review of self -evaluation and monitoring in EBRD - Chapter 5

The quality of investments self-evaluation

Management welcomes the initial findings about positive signs showing a narrowing gap between ratings produced by self-evaluation and those produced independently by EvD. However, the report concludes that a wide gap still remains. Management would nevertheless stress that the new Operation Performance Assessments (OPA) format, which is the heart of the EvD's new 'self-evaluation' approach, is still an experimental new tool for the Banking OLs whose projects are subject to review. It is still too early to make any meaningful observations about the results of the new approach as it may take a few years for the OPA statistics and the results show the full extent of the change.

In addition, Management notes that the OPA format requires Banking to re-frame the project according to the results presentation structure recommended in the EvD 'Performance Metrics' (discussed above), that is still being discussed and its adaptation is not fully agreed. Banking is now required to adopt the EvD proposed project results presentation retrospectively as if the project objectives and the results indicators had been incorporated in the project in that format at the outset. Management has recently observed that certain projects were rated lower because of this and wonder to what extent the rating has been affected by the retrospective application. Overall, Management expects a longer adaptation time than originally anticipated and welcomes the continued engagement by EvD for training and feedback.

Impact monitoring of investment operations
Management has discussed with EvD the shortcomings of the comparative analysis for project performance with TIMS information in the AER 2011. Indeed, given the difference between EvD ex post ratings and the TIMS ratings (detailed below), a comparison of EvD rating could be done only for expected transition impact ex-ante or projects with "negligible" or "low" risk ex-post according to the TIMS.

The final report of the Results task force TIMS working group clearly presents the main differences between EvD ex post ratings and the TIMS ratings at the time of EvD evaluations, as cited below:

- While the TIMS only considers transition impact through channels identified and benchmarked ex ante, EvD evaluations looks at all seven potential sources of transition impact. Therefore, an EvD rating can be higher than the ex-ante TI potential.
- The EvD looks for broader evidence for transition impact, beyond the attainment of pre-set benchmarks, and at the context in which benchmarks were missed or attained.
- While the TIMS distinguishes between TI potential and risk ratings, EvD's ex-post transition impact ratings blend both concepts – i.e., they effectively represent expected transition impact (although they use the same terminology as TI potential, rating projects "excellent", "good", etc.).

The report focuses on the predictive power of ex-ante transition impact and risk ratings on the EvD ex-post rating. Using the same EvD dataset, the analysis done under the Results task force TIMS working group reinforces some points made tentatively by the EvD in the AER report, while in others the findings are different.

This analysis confirms that the transition risk ratings at project inception are important, and that justifies the TIMS paper recommendation that the scorecard target for the new projects should be based on both TI potential and risks and not only on TI potential as it is currently the case. The magnitude of the probability of getting a particular EvD rating ex-post varies greatly for different combinations of transition impact potential and risk ratings. For the same transition potential, the risk rating is a strong predictor of the EVD (and TIMS) ex-post rating, i.e. the riskier the project the lower the chance of getting the envisaged rating. For example, if a project is rated Good the probability of it also being rated Good by EvD is about 0.63 for a medium risk project and 0.5 for a high-risk project. For satisfactory rated projects, the probability of the project being again rated Satisfactory by EvD is 0.66 for a medium risk project and 0.7 for a high risk projects.

Contrary to the key message of AER, the TIMS WG analysis concludes that the power of the ex-ante transition impact potential ratings to predict EvD ex-post ratings is very strong when controlling for risks. For the same risk category, Excellent rated projects at final review tend to have better ex-post EvD (and TIMS) rating than Good projects, which in turn have better outcomes than Satisfactory projects. The "odd" results in the AER, such as "projects assessed ex ante as having "satisfactory" transition potential achieved the same rating (or better) at evaluation in 90% of cases" are likely to be explained due to the difference between EvD ex post ratings and the TIMS ratings, listed above.

4. Specific and technical comments

The selection of indicators in Table 40 (Annex 3 page 7/16) is somewhat confusing - the indicators broadly reflect those in the World Bank's Doing Business report. The ranking as presented suggests that Tajikistan is doing remarkably better than the Kyrgyz Republic, and Kazakhstan, while the overall rankings in Doing Business are quite different (Kazakhstan 49, Kyrgyz Republic 70 and Tajikistan 141). Interestingly, EBRD and donors have supported Investment Councils in Tajikistan and the Kyrgyz Republic since 2007. The 2012 Doing Business Report ranks all the 185 countries in terms of most improved business environment since 2005, with both countries in the list of Top 10 global performers (Tajikistan 9 and Kyrgyz Republic 10).

The EvD comparison with IFC, which reported better evaluation of projects (before 2008 financial crisis) – is interesting. However, it should be noted that the large majority of IFC projects in the Kyrgyz Republic are within the Financial Institution sector, which has been the best performing sector in the region. IFC has never ventured beyond the sovereign and financial institutions sectors to sectors such as SMEs – where EBRD has increasingly ventured during the ETC Initiative to meet the needs of the economy.

On page 3/38, abbreviations, perhaps it would be good to annotate country groups with countries, i.e. under "CA", "EEC" all countries which come under these acronyms would be listed for clarity.

ANNEX

Additionality

The AER suggests that “consideration should be given to assessing a more focused definition of financial additionality as a relevance criterion”, given certain methodological issues surrounding the current approach to evaluating the additionality of projects. It may be helpful to place this suggestion in the context of previous discussions between Management and the Board on the concept of additionality.

The starting point for these discussions was the definition of additionality in the Graduation Paper which discussed additionality in terms of the Bank’s contribution to the “existence, design, or functioning” of a project. This approach recognized that clients receive a bundle of services from EBRD including, but not limited to, finance. Therefore, even if a client enjoyed access to financial markets and if the Bank’s financing terms were not the key reason for that client’s engagement with the Bank, EBRD could still be additional. This approach reflected the reference in the Chairman’s report attached to the agreement establishing the Bank (designed to be used for future reference in interpreting the Articles) that the Bank should “not compete with other organizations; rather, it should complement or supplement existing financing possibilities. Delegates also understood that “financing” and “facilities” were broad terms involving the whole range of Bank operations…”

Management’s view is that that additionality is best viewed as a yes/no judgement made at the time of approval. It is possible that the intentions which lead the Bank and the client to engage in an operation will not be subsequently realized, but the issue is whether these are good faith intentions, reflecting reasonable expectations of results at the time of approval. If so, then the Bank may be deemed additional. If however EvD’s evaluation judgement is that the assertion of additionality was founded on unreasonable expectations of the Bank’s value-added, then indeed it would be right for the evaluation to put in question the additionality. This approach enables us to retain the concept of additionality as a judgement made at the time of approval, in contrast to transition impact which is monitored in light of project results.