

SPECIAL STUDY

Performance Metrics: how well do EBRD projects specify expected results?

July 2013

EBRD EVALUATION DEPARTMENT



European Bank
for Reconstruction and Development

EvD produces thematic or sector level evaluation reports in the form of Special Studies. These provide valuable insights to strengthen operational outcomes and institutional performance; they focus on larger issues for which a transactions lens is unsuitable and generate more widely applicable findings for a wider audience. The larger scope of these studies facilitates the use of innovative and robust evaluation methods. This study forms one of 12 such studies scheduled for EvD's 2012 work programme.

This report was prepared by: Amélie zu Eulenburg, Senior Economist (EvD) with support from two external consultants, Peter Robertson and Anders Grette. All authors including the EvD peer reviewers Keith Leonard and Christopher Olson are collectively referred to as the evaluation team.

Contents

Abbreviations	iv
Defined terms.....	iv
Executive summary.....	v
1. Introduction	1
2. Methodology	1
2.1 Elaboration of the evaluability checklist.....	1
2.2 The project sample and assessment process	4
3. Testing the evaluability checklist and related findings	5
3.1 Expression of expected results.....	5
3.1.1 <i>Statistics from the sample</i>	5
3.1.2 <i>Issues to be considered</i>	6
3.2 Specification of performance benchmarks	7
3.2.1 <i>Statistics from the sample</i>	7
3.2.2 <i>Issues to be considered</i>	8
3.3 Risks and mitigating factors.....	10
3.3.1 <i>Statistics from the sample</i>	10
3.3.2 <i>Issues to be considered</i>	10
4. Conclusions.....	11
4.1 How useful was the evaluability checklist in the test assessment?	11
4.2 How evaluable are EBRD investment operations?.....	11
4.3 Recommendations	12
4.3.1 <i>With a view to the evaluability checklist itself</i>	12
4.3.2 <i>With a view to the results shown by the sample</i>	13
Sources.....	15
Annex 1: Study methodology	1
Annex 2: List of sample projects and related features	1
Annex 3: Statistics from the assessment	1
Annex 4: Deviations in consultants ratings	1

Abbreviations

ADB	Asian Development Bank
CSWG	Country Sector Working Group (in EBRD)
ECG	Evaluation Cooperation Group
EvD	Evaluation Department
FRM	Final Review Memorandum
MDB	Multilateral Development Bank
IDB	Inter-American Development Bank
OCE	Office of the Chief Economist
OPID	Operation's Identity Number
OM	Operational Manual
OR	Operation Report (Board document)
QaE	Quality-at-entry
TC	Technical cooperation
TI	Transition impact
TIMS	Transition Impact Monitoring System

Defined terms

Baseline data	States the initial level of a situation/indicator so that change from that level can be measured later on (EvD definition)
Evaluability	The extent to which the expected results of a project are verifiable in a reliable and credible fashion (adapted from ECG definition)
Impacts	The positive or negative long-term effects to which an operation contributes, directly or indirectly, intended or unintended (adapted from OECD-DAC definition)
Indicator	Quantitative or qualitative factor or variable that provides a simple and reliable means to measure achievement and to reflect the changes connected to an operation (adapted from OECD-DAC definition)
Metrics	The targets, indicators and baseline data for key indicators, which are a means to monitor and report on progress during implementation and after completion (EvD definition)
Performance	The degree to which an intervention operates according to specific criteria/standards/guidelines or achieves results in accordance with stated goals or plans (adapted from OECD-DAC definition)
Results	Results are usually expressed as output, outcome or impact and describe a changed situation when objectives and transition impacts are achieved (EvD definition)
Results framework	Also referred to as "hierarchy of results" - shows the cause and effect relationship between project inputs and activities and the different levels of expected results (EvD definition)
Targets	A particular value attached to a result to be accomplished within a given timeframe - requires baseline values (EvD definition)

Executive summary

The term evaluability is understood as “the extent to which the value generated or the expected results of a project are verifiable in a reliable and credible manner”. International experience shows that projects with high evaluability scores are more likely to be evaluated as successful. Evaluability assessments done prior to project approval not only provide an early indicator of likely success but also make it possible to fine-tune the project design to ensure it will effectively measure progress during implementation and the achievements on completion. Finally, evaluability-related knowledge serves the wider aim of institutional learning as it allows insights to be extracted from experience and lessons to be generated for other projects

This purpose of this study was to develop and test an evaluability assessment checklist that could have broader relevance as a quality assurance and learning tool at the EBRD. Further, the results from pilot testing of the checklist on a statistically valid sample of recent Operation Reports (ORs) approved by the Bank between 1 July 2011 and 30 June 2012 provides immediate feedback to Operations staff on the feasibility of applying the checklist as well as an initial baseline of experience for future assessments.

The study has two main findings. The first relates to the utility of the checklist and the second to the results of applying the checklist to the sample projects. Overall, the evaluability checklist of 10 criteria assessing three aspects – expected results, performance benchmarks and indicators, and risks and mitigating factors – proved to be an effective tool. However, the two risk criteria were the most difficult to apply consistently and may require further refinement.

The picture is mixed in terms of the results from applying the checklist to the sample projects. EBRD operations score reasonably well for the definition of operational objectives and their results, and for transition impact results and their benchmarks. This is offset by relatively low scores for the quality of indicators/benchmarks (such as target values, timelines and baselines). These are necessary conditions to measure performance.

Key information on results (other than transition) is often not presented clearly and consistently within the OR but is scattered throughout the document. Few ORs effectively followed advice in the Operation Manual (OM) regarding section 3.5 (measuring/monitoring success). If performance-related information including baselines is not coherently stated and readily available in the OR it is unlikely to be used to monitor and evaluate performance. Consequently, the monitoring system's ability to provide an early warning that projects may be going off track in delivering expected results could be compromised.

The presentation of results in ORs often lacks a coherent framework wherein the operation's achievement of operational and transition-related results are explicitly linked across levels, including outputs and outcomes and their contribution to higher-level impacts. Each operation is required to describe its links with country and sector strategies and appropriate citations are typically provided, but no measures to assess how these results were to be achieved were provided.

Recommendations of the study are:

- I. With regard to the application of the evaluability concept and the checklist:
 - management to decide on the utility of the tested and adjusted checklist and what role it might play in the project appraisal process at the EBRD
 - EvD to develop practical guidance and tools for Operation staff as needed (for example, templates and good practice examples for developing a coherent hierarchy of results for individual projects).
- II. With regard to the results shown by the sample projects:
 - improve the presentation and quality assurance of section 3.5, Measuring/monitoring success, consistent with the advice in the OM
 - improve the quality of operational objectives performance metrics
 - ensure adequate baseline data is presented for all key results and related indicators/TIMS benchmarks at one central place in the OR, for instance section 3.5
 - develop an EBRD-specific results framework to improve reporting of institutional achievements through a more structured description of operational results and their links to Bank strategies, policies and corporate scorecard. Critical components of such a framework would be:
 - operational results presented with the standard terminology of outputs, outcomes and impacts
 - upward links between the expected project results and their potential contributions to transition impact at country/sector level clearly presented
 - country strategies and sector policies incorporate an aligned results framework
 - enabling institutional learning and matching the requirements from the Bank's wider modernisation agenda.

1. Introduction

Evaluability is defined as “the extent to which the value generated or the expected results of a project are verifiable in a reliable and credible fashion” (adapted from the Evaluation Cooperation Group [ECG] definition). The 2011 Briefing Note on Evaluability (PE11-540S) reviewed the approach and good practices of other multilateral development banks (MDBs) to evaluability and concluded that it would be relevant to the EBRD as well.

Experience shows that projects with high evaluability scores (in other words, they have sound *ex-ante* monitoring and evaluation systems) are more likely to be successful. Most importantly, evaluability assessments conducted before project approval provide an opportunity to fine-tune the project design to ensure it measures progress during implementation and achievements on completion. Further, these assessments can provide an early indicator of likely success or potential problems.

Consequently, the Performance Metrics study was included as a thematic evaluation in the Evaluation Department's (EvD) final work programme for 2012. The purpose of this study, the first evaluability assessment conducted at the Bank, was to:

- develop a checklist approach to assessing the evaluability of the EBRD's operations
- conduct an evaluability assessment of a random sample of recently approved operations thereby piloting the approach, testing the applicability of the checklist and producing a baseline figure for potential future assessments, and identifying examples of good practice
- raise the evaluability-awareness of operational management, who might find the results of the study, as well as the evaluability checklist itself, helpful as a quality assurance tool during project preparation.

2. Methodology

2.1 Elaboration of the evaluability checklist

By their very nature, evaluability checks are a quality-at-entry tool, performed before the formal approval of operations. The principal guidance on the project approval process is provided in the Operations Manual (OM) available on the EBRD intranet and frequently updated by assigned staff members. The OM provides guidance on the scope and content of the Final Review Memorandum (FRM) which is the basis of the Operation Report (OR) that is submitted to the Board for approval. In practice the OR is also referred to as the Board document. An OR's typical contents are outlined in Annex 1, Table 1.

The evaluability checklist mirrors the key principles applied at other MDBs but was especially designed for the EBRD and this study in particular. The checklist comprises 10 questions, or criteria, divided into three sections (See Table 2.1.1 below):

Table 2.1-1 Evaluability checklist

Evaluability criteria	Definitions applied by the evaluation team
1. Expression of expected results	
1.1 Operational objectives consistently stated	The project's operational objectives describe what the client will use the financing for (the business purpose or use of proceeds)
1.2 Operational objectives results stated	These describe the expected changes at project level after operational objectives are achieved
1.3 Transition impact results stated	These describe the expected systemic changes for each transition impact criteria at sector/economy level
1.4 Results showing linkages across levels	Describes both how results within different areas of a project are interdependent and how project results contribute to higher-level country, sector and institutional results
2. Specification of performance benchmarks/indicators	
2.1 Operational objective's indicators specified	The variables used to show operational objectives have been achieved
2.2 Transition impact's benchmarks specified	The variables used to show transition impacts have been achieved
2.3 Benchmarks/indicators specify target value and time period	Quantitative indicators require targets and an agreed timeframe; qualitative indicators require an agreed timeframe
2.4 Benchmarks/indicators specify baselines	Baseline describes the pre-operation approval status of each quantitative benchmark/indicator. This enables a comparison with the target value during implementation and at completion
3. Identification of risks and mitigating factors	
3.1 Mitigating factors stated for key risks including the EBRD's role in design and conditions	The review should focus on whether risks are specified as being in or out of the EBRD's control. The description of mitigating factors should be related to implementation management and transition risks
3.2 Why unmitigated risks are acceptable to the Bank explained	Where mitigating factors are not defined, has the acceptability of the risk been explained?

Section I asks the whether the operational objectives and related results are stated, whether the transition impact results are specified and whether the results are grouped within a hierarchy showing interdependencies. In general, an EBRD OR describes a range of objectives:

- the client's purpose in requesting the financing (referred to in this study as operational objectives)
- the Bank's purpose in supporting the client in line with its mandate: the expected transition impact
- the financial performance objectives
- the environmental objectives (depending on the project's categorisation).

It is important to note that this study and the checklist focus on the first two sets of objectives. The EBRD delivers transition impact through the projects it funds. There is a direct link between the operational and transition-related objectives of its investments, as project outputs lead to the achievement of the project outcome, which in turn contributes to the achievement of (transition) impact.

In principle, the scope did not include financial performance indicators relevant to sound banking and that are subject to credit monitoring by the Bank.¹ Another exception is the environment. The scope of the study excluded any review of environmental and social assessments, impact studies or action plans as these were often pending when the OR was presented for approval, and also because this would have required both a separate in-depth review with other expertise involved and an earlier OR approvals calendar cohort.

Further, as the Bank's projects are expected to support higher-level country or sector priorities, the checklist includes a criterion on the coherence and quality of these linkages, assessing whether the interdependencies between levels of results are clearly defined in the form of a hierarchy (question 1.4). Although a presentation of linkages between project results across different levels is not explicitly required, it is inferred by the OM guidance for the OR section 3.5, Measuring/monitoring of success, where it asks the OL to "provide a table summarising key objectives which emerge from the above sections, and indicating what specific parameters or benchmarks will be used to evaluate the Bank's success in meeting these. ... there will usually be other indicators to confirm whether a "checkable story" relevant to transition has been fulfilled ... tell the whole story on how we will measure/monitor achievement of objectives, particularly transition impact."²

The four criteria in **Section II** assess the status of performance measures as the basis for the project's monitoring arrangements including specification of target values, timeframes and baselines. The presence of clear and concise performance benchmarks and indicators is essential for effective monitoring and evaluation. These are required in two specific tables in section 3 of the OR., Firstly, selected benchmarks for transition impact and their timeframes are recorded in the table in section 3.1 to monitor achievement of each TI criterion. Secondly, in section 3.5 it is expected that key project objectives and their performance indicators will describe how the broader project storyline will be monitored and evaluated. The checklist criterion 2.3 assesses whether benchmarks/indicators specify target values and a time period. In addition, it is essential to specify the baseline status of key benchmarks and indicators at the start of the project in order to assess the magnitude of change.

Lastly, the two criteria in **Section III** assess the status of risks with a specific emphasis on mitigation factors and the EBRD's role in this. This section covers two main aspects: the key risks and issues associated with the operation (such as completion, environmental, management capacity risks) and the mitigating factors. Risks are mainly presented in section 5 of the OR, with risks to transition impact described separately in section 3. In general, the assessment reviewed whether risks were distinguished according to whether they are inside or outside of the EBRD's control and whether the Bank's role in mitigating the risks should be stated. For risks that can't be mitigated, criterion 3.2 in the checklist assesses whether the acceptability of the risk has been explained.

Although they are not performance metrics in themselves, risks are an important aspect of evaluability because they can have such a dramatic effect on the attainment of results by affecting or disrupting the assumed linkages. Consequently, project design must take into account risk occurrence and management by specifying them in the appraisal document. Identified risks should be accompanied by plausible mitigation measures and the responsibility and mechanisms for monitoring the occurrence of risks should be clear, specifically when the EBRD has a role in this.

The proposed scoring system accommodates a criterion where there are multiple instances to assess a criterion as well as those where the criterion is assessed by a "yes/no" or "all/none" answer. This is

¹ As the sample review shows, such indicators overlapped with TIMS benchmarks at times.

² http://intranet.ebrd.com/ebdnet/operations/opsman/ch1/1_5.shtml

elaborated in a four point scoring system with two positive and two negative scores similar to the overall performance rating applied by EvD, ranging from “highly successful” to “unsuccessful”.

Table 2.1-2 Rating system for the evaluability criteria

All	= criterion was true in every instance – 100%
Mostly	= criterion was true in >50% and <99% of instances
Few	= criterion was true in <49% and >1% of instances
None	= criterion was not true in any instance – 0%

It is important to note that the range on the top and the bottom of this rating scale is very narrowly defined by 100 and 0 per cent respectively, whereas the two medium ranges encompass everything between 50 and 99 per cent or 1 and 49 per cent. Taking an analogy with academic exams, one could pass with either 55 per cent or 90 per cent, both of which would be rated “mostly”. However, this rating does not reflect the very different quality across these two results. Evaluability results from other institutions suggest that a 70 per cent rating for the “all” category should be aimed for and that the two positive categories together should account for at least 90 per cent of each criterion. Annex 1 details the application of the rating system to the evaluability checklist with some practical examples.

2.2 The project sample and assessment process

To ensure that the study would mirror the current stage of project design at the Bank, a stratified random sample was selected from all EBRD operations submitted to the Board within a 12 month period over the last two years to test the evaluability checklist. The sample did not include project extensions that are rooted in OR from former years.

The target population deliberately concentrated on stand-alone debt and equity investment projects and excluded other matters approved through the Bank’s Board of Directors such as country strategies or a certain group of technical cooperation (TC) projects.³ Consistent with EvD’s annual work programme, multiple samples were taken randomly and tested against the population to ensure the best fit. Annex 2 provides a list of the 59 sample projects and illustrates other features such as the country, sector, size and private/state nature of the sample in greater detail.

Table 2.2.1: Required sample size

Period of project approval	Size of population	Size of sample	Coverage
July 2011 – June 2012	154	59	38.3%

The practical testing of the evaluability checklist at the EBRD was implemented by two external consultants who spent some time together at the Bank’s Headquarters to familiarise themselves with the EBRD’s current *modus operandi* and handling of project design and related documents. An extensive document review was supplemented with interviews and meetings throughout the Bank.

³ For instance, financed from the Bank’s Shareholders’ Special Fund and exceeding a contract value of €300,000.

Whilst each consultant was responsible for about half of the project sample, the evaluability checklist and suggested rating system were tested by both consultants simultaneously on a number of joint projects beforehand. This has helped to clarify the interpretation of criteria and created a common understanding to ensure consistency in rating throughout the study and is described in greater detail below.

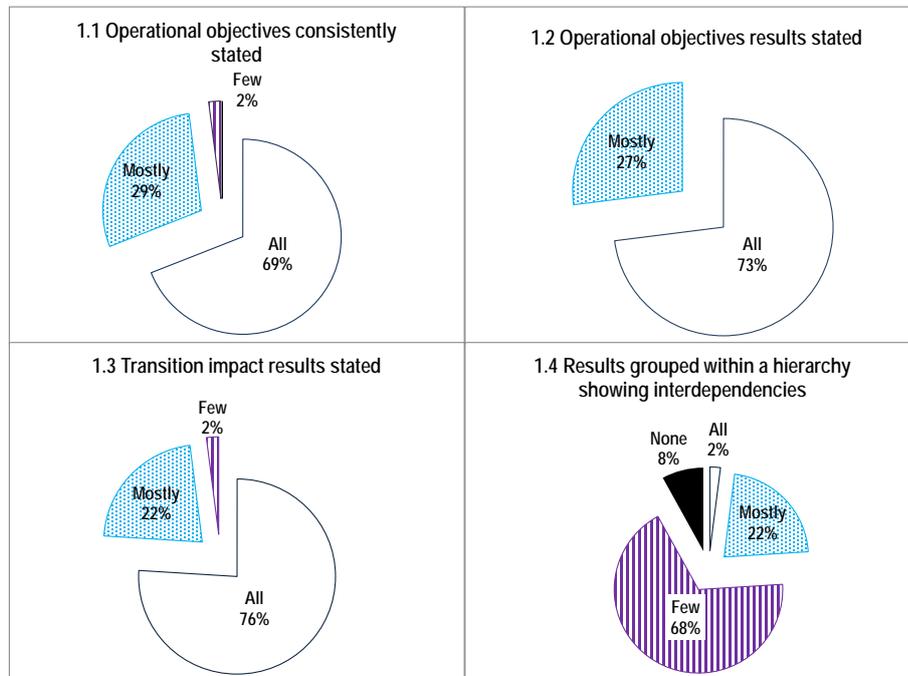
3. Testing the evaluability checklist and related findings

Section 3 gives an overview of the sample results by summarising the ratings of each individual evaluability criteria and describing general features and issues. The statistics in Annex 3 provide full details of all scoring. The utility of the checklist and related findings are discussed in section 4.

3.1 Expression of expected results

3.1.1 Statistics from the sample

Table 3.1.1: Statistics from the sample



Operational objectives were consistently stated in 69 per cent of the ORs and for two out of three statements in 29 per cent of the ORs; only in 2 per cent of the ORs were none of the statements consistent. Although each of the three statements has a specific purpose, and therefore the language and details included may vary, the checklist criteria assessed whether the essence of the project's purpose was consistently stated in each relevant section. Given that these statements describe the purpose of EBRD financing from a legal and fiduciary perspective, it would be reasonable to expect that the OR presentations would always be consistent.

Results for "all" the projects operational objectives were stated in 73 per cent of the ORs. However, these result statements were not found in one location but scattered across the narrative in sections 1 and 3, occasionally in the annexes.

Transition impact results for “all” TI criteria were stated in 76 per cent of ORs, and for “mostly” in 22 per cent with only 2 per cent of ORs having “few” results stated and no cases with zero results statements. These results were normally presented in the narrative in section 3.1, providing an indication of the type of change that could be expected under each TI criteria as a result of successful project implementation.

The linkages and interdependencies of expected results across different levels were made explicit in only a quarter of the sample projects. Positive examples tended to use section 3.1 to amplify the results of each of the proposed TI criteria, with the country and sector analysis taking the results to the next level of aggregation. This is discussed further in section 3.1.2 below.

3.1.2 Issues to be considered

As currently structured in an OR, operational objectives (including use of proceeds) and transition-related objectives and impact seem to be disconnected and yet the operational objectives need to be achieved in order to maximise transition impact. For example, while TI is clearly at the core of the EBRD’s mandate, it is not generally mentioned in the President’s recommendation at the beginning of the OR. On the other hand, the TI benchmark table in section 3 concentrates solely on measuring TI and does not refer to related operational objectives or their indicators. The guidance provided in the OM on section 3.5 suggests that this is where interrelationships between these two aspects of an operation should be shown but this is seldom effectively done.

Further, section 3.5 would also be the place to present the project in its strategic and institutional setting, including how the project will contribute to country or sector level results. This would explicitly show the links and interdependencies of expected results across different levels and enable the Bank to aggregate results from individual projects to a corporate level. However, the absence of an agreed methodology to define expected results in EBRD terminology makes it difficult to define the different levels of anticipated objectives and results. Consequently, section 3.5 in the OR is handled very diversely by operation leaders, ranging from a minimalist approach providing a sentence referring to reports to providing a structured framework elaborating additional indicators including baselines, target values and timeframe.

For good examples of project-specific objectives see the PSSF SG Banka Srbija SME project (OPID 43155) and in particular the process outlined to monitor the Partnership for Growth loan including Annexes 1 and 6. The SVL (OPID 42876) and Baring Vostok Private Equity Fund V (OPID 43703) projects both identify a small number of project-specific objectives with key indicators which would be more complete if specific performance measures were provided in the table or an annex.

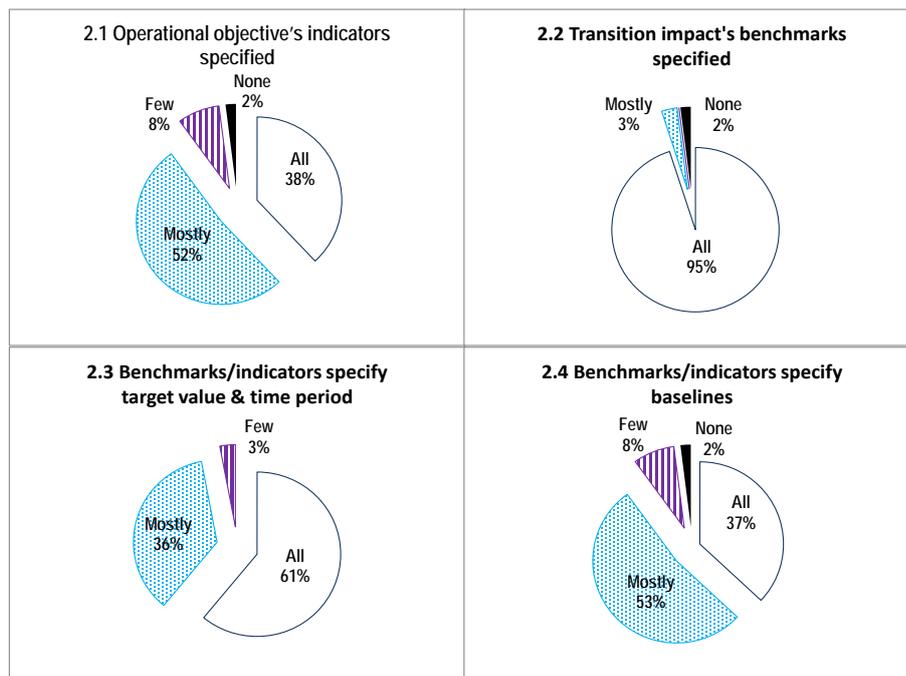
In the context of criterion 1.4, the study also reviewed country and sector strategies relevant to the sample projects. These strategies typically seemed to address key transition challenges and refer to past and current operations before going on to identify broad areas of potential operations while taking into account the flexibility necessary for an institution focused on the private sector in responding to demand. Similarly, in the reviewed OR the context of the individual operation was mostly confined to short statements intended to anchor the operation in the relevant EBRD country and sector strategies rather than elaborating on the strategic and operational justifications. This is another reason for the below-average scoring of criterion 1.4.

Within the sample, the Tiryaki Agribusiness project (OPID 43069) is a good example of how to link the project-level results to the expected impacts at country and sector level. The logical linkages between micro- and macro-level results, by way of the first contributing to the second, is supported by indicators throughout. More generally, a **results framework** describes interdependencies between different layers of results, specifying which is leading to which. The internationally accepted terminology used was introduced by OECD-DAC referring to outputs (immediate results of activities undertaken); outcomes (more comprehensive changes brought about by a number of outputs, such as a change of behaviour or organisational performance); and impact (the longer-lasting change contributed through achievement of one or more outcomes). Whilst outputs and outcomes are usually attributable to project activities, impacts pertain to the higher level sector, country or regional results to which a project contributes.

3.2 Specification of performance benchmarks

3.2.1 Statistics from the sample

Table 3.2.1: Statistics from the sample



In only 38 per cent of the ORs were indicators specified for "all" the operational objectives. Another 52 per cent were rated as "mostly" for this criterion, thus 90 per cent of the sample rated positively. However, the difference from the TI benchmarks (see Table 3.2.1) is striking, again reflecting the emphasis on TI over operational objectives. However, it was noted that project monitoring and disbursement procedures are linked to loan agreement conditions and covenants that are not always detailed in the OR main text.

In 95 per cent of the ORs, TI benchmarks (criterion 2.2) were specified in "all" instances. This finding was expected given the strategic focus under the EBRD mandate and the importance of benchmarks as

project implementation monitoring tools under TIMS. However, a significant variation in number, type and timing of benchmarks was notable. While three to five benchmarks were identified for most TI criteria, in a small number of projects the OR presentations included numerous (eight or more) and sometimes extremely detailed activity-type benchmarks.

In 61 per cent of the ORs, benchmarks and indicators specified both target values and time period in “all” cases. The balance of 39 per cent lacked one of these elements, most often a target value. The time period is mandated for each benchmark and was almost always provided in the monitoring tables in section 3 of the OR. Conversely, the indicators found in section 3.5 on measuring/monitoring of success seldom specified either target values or the time period.

Baselines for TI benchmarks and operational objective indicators were provided in “all” instances in only 37 per cent of the ORs. In a further 53 per cent of the ORs, baselines were provided for most of the benchmarks/indicators. In the remaining 10 per cent of the OR either “few” or no baselines were stated. A similar finding was drawn in EvD’s 2011 Special Study on the Early Transition Country Initiative (ETCI).⁴ Given the importance of baselines to effective evaluation of the achievement of results, this repeated observation is of concern and should be addressed by OCE and Operation Teams in the future.

The study did identify a number of good practice cases with regard to TI baseline data, benchmarks and target dates as presented below:

Box 3.2.1: Good practice examples for TI benchmarks with baseline data

Cersanit Syzran Sanitary Ware Expansion Project (OPID 42677)	ProCredit Bank Serbia – Senior Loan project (OPID 42496)
<ul style="list-style-type: none"> • Output/employee reaching 2000 pcs. p.a. by 2015 vs. 1000 in 2010 • Gross profit/employee reaching RUB 1 million. p.a. by 2015 vs. 0.1 million in 2010 • 1-2 new FDIs in Syzran to materialise by end 2013 following the Cersanit pioneer example • New industry technical standards beyond GOST adopted by Cersanit by end 2015 	<ul style="list-style-type: none"> • Gross portfolio growth: year-end value increase from €461-564 million in 2010-13 • Outreach: portfolio share of loans outside Belgrade to grow from 65-70% in same period • Local currency loans: year-end portfolio share reaching minimum 30% in 2014, from 26% in 2010 • Over 30 days PAR value to remain below 5 % at all times vs. 3.7% at end 2010

3.2.2 *Issues to be considered*

Apart from the rather low rating of less than 40 per cent of the ORs specifying indicators for “all” operational results, these indicators were scattered across OR sections 1 and 3 and/or in the annexes. Again, it would be expected from the guidance in the OM that these indicators would be reported in section 3.5. This was not the case in any of the reviewed ORs. Rather, it appears that the monitoring and evaluation focus in the OR is on transition impact criteria and not operational objectives. This finding is of concern and should be the focus of quality assurance reviews during project preparation.

⁴ Where about three projects out of 10 had baseline data specified for the key transition impact benchmarks, see PE11-510S

A project's operational objectives are central to the EBRD's fiduciary responsibility. Similar to other MDBs, the Agreement Establishing the Bank states "the Bank shall take the necessary measures to ensure that the proceeds of any loan made, guaranteed or participated in by the Bank, or any equity investment, are used only for the purposes for which the loan or the equity investment was granted and with due attention to considerations of economy and efficiency" (Article 13.xiii). (See basic documents of the European Bank for Reconstruction and Development (EBRD) 15.04.1991)

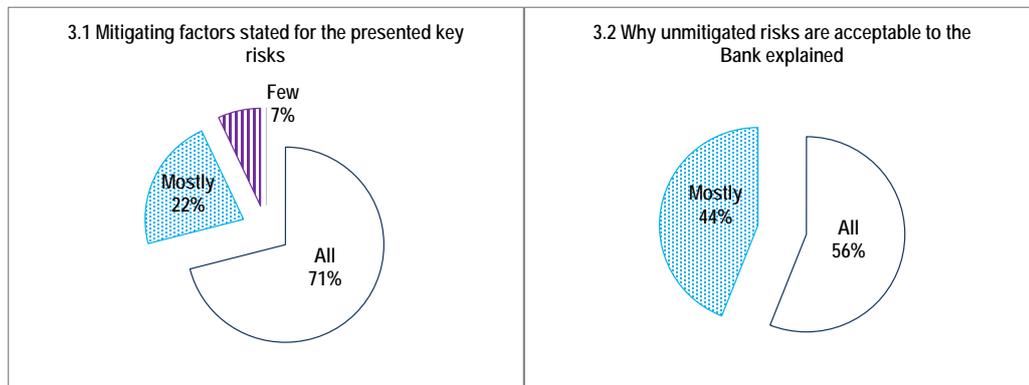
Typically, a project is clearly defined in the legal agreements and such definition conforms with the Use of Proceeds description contained in the OR Section 2 Key terms and conditions. This describes the purpose for which the client is requesting EBRD financing. The client is required to confirm the accuracy of the project description (and hence the Bank's understanding of what is to be financed) through a specific clause in the legal agreement. The client also reconfirms the continued relevance of the use of proceeds in each disbursement request.

In respect of indicators specified for transition-related results the issue emerging from the study is quality rather than quantity. For a single TI criterion the presented benchmarks could include a combination of activities, outputs, indicators, milestones or project covenants, lacking, however, a consistent hierarchy of results. Although benchmarks were most often expected to be reached prior to project completion, there were cases, other than for the demonstration criteria, where benchmarks would not realistically be met until after project completion, which raised questions as to their value for implementation monitoring and who would collect data after completion. While related baseline data were often present somewhere in the OR, in only 37 per cent of ORs were they included for "all" presented benchmarks in the TI monitoring table.

While most TI benchmarks describe indicators to be measured during implementation, it is not unusual to find indicators that could only be measured after project completion. Regarding frameworks or facilities and projects approved under these, there was no consistent pattern in practice but the tendency was to present indicators/benchmarks for the individual project(s) with no reference to a set of results, including transition impact, pertaining to the framework/facility. Exceptions to this included the Regional ETC Non-bank MFI Framework IV first sub-project: MDO FINCA Tajikistan (OPID 43302) where results for both the framework and the individual project are defined; and the DLF Marneuli (OPID 40413) and the Kirishi Rouble Loan (OPID 43011) projects where TI benchmarks from the previous project are updated in the current one. Consistency in presenting results for these types of projects should be addressed.

3.3 Risks and mitigating factors

3.3.1 Statistics from the sample



As required in the Operations Manual (1.5.4, section 5), considerable effort is expended in demonstrating “to Ops Com and the Board that the risks associated with the operation have been addressed and why risks which are not mitigated are acceptable to the Bank ... (and showing) the extent of the EBRD’s control over the occurrence of such negative events (if any)...” On balance, the risk section of the checklist raised relatively few concerns or issues, save for a few observations on the presentation format.

In 71 per cent of the ORs, mitigating factors stated for key risks as per criterion 3.1 included the EBRD’s role in “all” relevant risks and a further 22 per cent in “most” relevant risks. When mitigating factors were not defined, the acceptability of the risk was explicitly explained for “all” relevant risks in 56 per cent of the ORs and for “most” of the relevant risks in the remaining 44 per cent of the ORs. Not a single case was found in which this criteria would not have been met.

3.3.2 Issues to be considered

Interestingly, the checklist criterion 3.2 (explanations on why an unmitigated risk would be acceptable to the Bank) was the only one showing a substantial deviation between the scores applied by the two consultants (see Annex 4 on the rating results per consultant). Consultant 1 assigned “all” ratings in the majority of cases, stating that he generally found a reasonable explanation in the table presented under OR section 5.1. Consultant 2, who rated “mostly” in many of his sample projects, stated that the explanation was sometimes not as explicitly mentioned as he would have expected. Still, he concluded that the acceptance of such risks would be implicit in the EBRD mandate and policies and therefore scored a “mostly”. On the other hand, the evaluation team had some discussions about the benefit of such a risk-related section in the first place which is presented and concluded upon below.

The above could suggest that the OR statements could be even more explicit on whether the Bank controls a risk or not and whether a risk that is not to be mitigated as such should still be acceptable. Efforts could also be made to improve the consistency of the risk analysis table format with a view to distinguishing more clearly between mitigating factors and mitigation measures. For example, OPID 43155 included a standard heading separating the mitigation factors from the risk description in the comments column and in OPID 43069 the scope of risk was also summarised in the first column in the table. These initiatives improved the readability of the risks table.

4. Conclusions

4.1 How useful was the evaluability checklist in the test assessment?

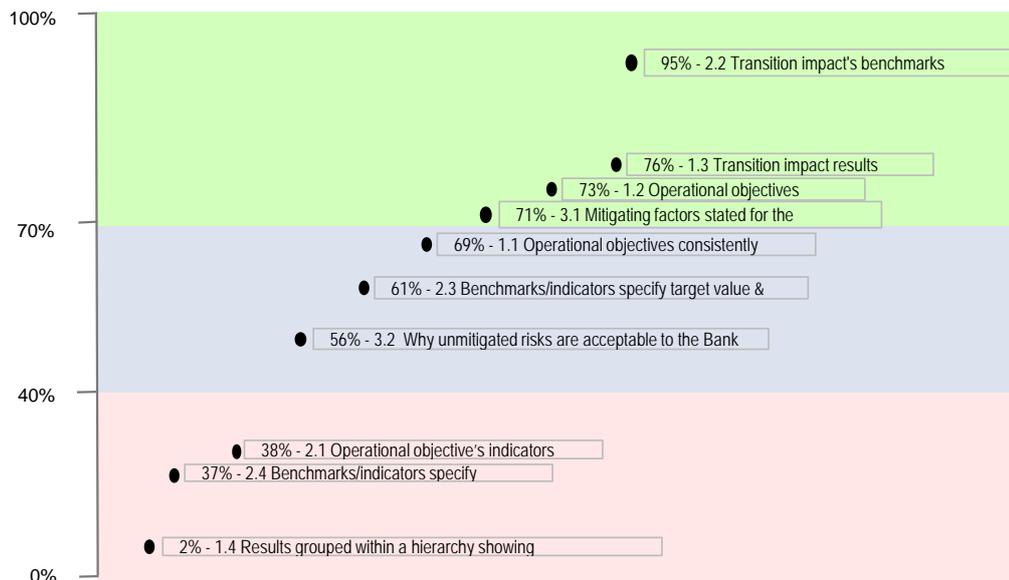
The evaluability checklist proved to be an effective assessment tool overall. It contains the necessary criteria, is straightforward to interpret and was easy to apply to the ORs for the selected sample. This general impression of the evaluation team is supported by the fact that deviations in scorings between the two consultants have been quite low. This is illustrated in detail in Annex 4.

The chosen criteria in the risk section have been difficult to apply uniformly between the two consultants. During the discussions it became clear that this issue does not necessarily pertain to the wording of criteria 3.1 and 3.2 but to the benefit of having a risk section in the evaluability checklist for EBRD projects at all. A fully-fledged system of assigning and monitoring risks associated with transition impact is already in place so it seems there is little value in asking the same for operational objectives. EvD proposes to remove the section 3 on risks from the evaluability checklist.

4.2 How evaluable are EBRD investment operations?

The "all" ratings achieved per criterion in the Bank's sample projects are illustrated in Figure 4.2 below. While it is ideal for each criterion to be rated "all" in 100 per cent of an OR, in practice this is seldom achieved. From comparable experience in other MDBs, ratings of over 90 per cent as "all" for each checklist criterion is considered a high quality benchmark whilst ratings of less than 70 per cent would be a concern. The best result of 95 per cent of projects found in the "all" category, was scored for the presence (not the quality) of transition impact benchmarks. This is followed by the specification of transition impact results (76 per cent) and operational objectives (73 per cent).

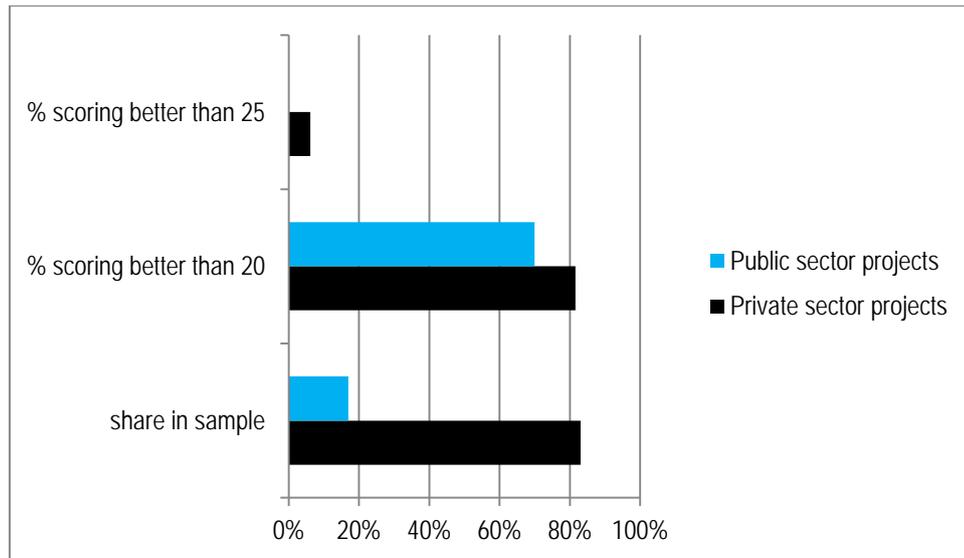
Graph 4.2-1: Ratings "all" achieved per criterion for the EBRD sample



The frequent "mostly" scores found in the sample suggest there is still considerable scope for continued improvements to achieve a higher share of "all" ratings. The challenge for the Bank is to move towards the 90 per cent threshold for all the evaluability criteria in the next couple of years.

One particular observation deserves a mention, given the EBRD’s specific focus on working mainly with the private sector and following a demand-driven approach. In the pre-assessment discussions within the Bank, there was considerable scepticism as to the applicability of the evaluability concept. It was perceived to be better suited to the more development- and public sector-oriented mandate pursued by other MDBs. The sample results were reviewed in this regard by comparing the scores for the public sector projects with those for the private sector projects. As can be seen in the graph below, the private sector projects scored even better than the public sector ones.

Table 4.2-2: Performance of public versus private sector projects in the sample



All = 3, Mostly = 2, Few = 1, None = 0; maximum achievable for the 10 Evaluability-criteria = 30 points

The central observation relevant for both private and public sector projects is that key evaluation information is scattered throughout an OR, including its annexes. This is discussed in greater detail in section 4.3.2 below.

4.3 Recommendations

4.3.1 With a view to the evaluability checklist itself

Whilst the checklist as a whole seems to be fit for purpose, the two criteria related to the assessment of risks have been removed. In addition to fine-tuning the list itself, EvD plans to elaborate some practical guidance on its interpretation and usage for Operation staff, such as:

- Good practice examples could be assembled and presented to Banking teams. It should be noted that the results of the evaluability assessment were deliberately not broken down to team level as the numbers in the sample were too low for a meaningful comparison. That does not mean, however, that such aspects would not be interesting and useful to cover in similar future assessments.
- The evaluability checklist could become an effective quality assurance tool for Operation staff overseeing project preparation. The present list could also serve as a useful instrument for other supporting departments in the Bank such as the Office of the Chief Economist, the

Environmental and Social Department and the Donor Co-financing and Technical Cooperation teams. They could develop their own checklists containing performance metrics for their particular area of responsibility.

Now that the baseline exists, evaluability checks could be run regularly – by EvD or management itself – to review progress made in individual areas. It is recommended that:

- Management decides on the utility of the checklist and what role it might play in the project appraisal process at the EBRD
- EvD develop practical guidance and tools for Operation staff as needed (such as templates and good practice examples for a coherent hierarchy of results for individual projects).

4.3.2 *With a view to the results shown by the sample*

Improve the quality of benchmarks and the systematic collection of baseline data

This depends partly on the hierarchy of results mentioned above, as a systematic methodological approach to defining the anticipated results will have positive effects on the quality of related indicators and benchmarks. Operational and transition-related results could be better distinguished. Baseline data as well as specific targets and time periods could focus on a small (around five) number of key performance metrics. The monitoring of benchmarks beyond project level would be referred to other strategic documents/teams as appropriate. It is recommended to:

- Ensure adequate baseline data is presented for all key results and related indicators/TIMS benchmarks.

Enhance the overview and tracking of performance metrics in the OR

It would be desirable to see the key objectives and indicators comprehensively and concisely presented, such as in the form of an overall results framework in one place in the OR. In general the guidance on the content of ORs in the OM provides sufficient opportunities for results and their relevant performance metrics to be clearly and concisely described. This should be specified further as the interpretation by Operation staff currently varies substantially and should be more effectively enforced in its application. It is, in general, recommended:

- Section 3.5, Measuring/monitoring success, would be the natural fit for presenting an overall results framework. If another section is determined for this role, such as section 3.1, it should also include operational results and not be limited to transition impact.

Strengthen the consistency and coherence of operational results in relation to transition impact

The Bank's focus on the expected TI of its projects seems to overshadow the *ex-ante* definition and implementation monitoring of the other operational objectives. It is recommended to:

- Improve the quality and consistency of operational objectives performance metrics.

Strengthening the links between a project and the relevant country and sector strategies

The presentation of results suffers from the absence of a coherent framework linking results across levels. Naturally, attribution is limited to project relations between inputs, activities, outputs and outcomes. In other words, the OR would benefit from a clearer articulation of the higher level impact the project will contribute to. The work of the current Country Strategy Working Group (CSWG) is crucial in this respect, as is the Joint Task Force on Results Management/TIMS. There are very good prospects for linking project appraisal documents more closely to strategic considerations at country and sector level.

Moreover, this would allow for a more consistent aggregation and reporting of achievements beyond project level in the longer-term.

It is recommended to:

- Develop an EBRD-specific results framework to improve reporting of institutional achievements through a more structured description of operational results and their links to Bank strategies, policies and corporate scorecard. Critical components of such a framework would be:
 - operational results presented with the standard terminology of outputs, outcomes and impacts
 - upward links between the expected project results and their potential contributions to transition impact at country and sector level clearly presented
 - country strategies and sector policies incorporate an aligned results framework
 - enabling institutional learning and matching the requirements from the Bank's wider modernisation agenda.

Sources

Links evaluability and quality-
at-entry assessments from
other IFIs

[IADB](#)

[IADB Development Effectiveness Review](#)

See pages 34-5

[ADB](#)

ADB's [Development Effectiveness Review 2011](#)

See, for example, section (level) 3 of the report

[World Bank](#)

The Bank's Operational Manual

available on request

Briefing Note on Evaluability
(WP2011)

available on request

Approach paper for this study

available on request

Terms of reference and CVs of
consultants

available on request

Results from the evaluability
assessment in detail

available on request

Annex 1: Study methodology

The evaluability checklist is applied to each Operation Report (OR) presented to the Board which typically contains the following table of contents:

Heading/section title	Sub-sections
1 The client and the project	President's Recommendation Summary fact sheet 1.1 Sector/background 1.2 Company history and operation 1.3 Strategy and potential expansion 1.4 Ownership/management/the sponsor 1.5 Financial highlights 1.6 Competitive environment 1.7 Investment programme
2 Key terms and conditions	Table
3 Rationale for the Bank's involvement	3.1 Transition impact and Bank strategy 3.2 Additionality 3.3 Environment 3.4 Projected profitability for the Bank 3.5 Measuring/monitoring success*
4 Financial/economic analysis	4.1 Sensitivity analysis 4.2 Valuation
5 Key risks and sensitivity analysis	5.1 Sensitivity analysis/risks 5.2 Integrity 5.3 Lessons from past experience
Annexes to Operation Report	For example: I Company/ Market Overview II Lessons from past experience III Financial analysis summary IV Offshore jurisdiction annex V Project/procurement implementation

* The template in the OM refers to section 3.5 as "Other relevant issues" and section 3.6 as "Measuring/monitoring success". However, section 3.5 was titled "Measuring/ monitoring success" in almost the entire sample.

Rating system. The rating system was developed to ensure a consistent score could be given when a checklist criterion applied multiple cases in an OR and was applied as follows:

Criterion 1.1 - Operational objectives consistently stated

The project's operational objectives are stated in the following places in the Operation Report:

- the President's Recommendation
- the summary fact sheet as project description/business purpose
- the terms and conditions as Use of Proceeds
- occasionally in the Strategy Section 1.3.

The criteria are rated:

"All" if the operational objectives are stated consistently in all three or four places

"Mostly" if they are stated the same in at least two of the three or two of the four places

"None" if the operational objectives are stated differently each time.

In this case there is no "few" rating.

Criterion 2.3 - Benchmarks/indicators specify target values and time period

This criterion refers to the transition impact monitoring benchmarks table in section 3.1 of the OR and saw the following rating applied:

"All" If all benchmarks state both a target value and time period

"Mostly" If over 50 per cent of the total number of benchmarks/indicators state both a target value and time period

"Few" If less than 50 per cent have both a target value and time period

"None" If benchmarks/indicators have either a target value or a time period but not both.

Annex 2: List of sample projects and related features

OPID	OP Name	Country	Sector team	Instr	P/F Class	BA Amt	Size
39581	Project Bronze	Mongolia	Natural Resources	Debt	PRIVATE	40,873,524	S
40413	DLF-Marneuli	Georgia	Agribusiness	Debt	PRIVATE	4,900,000	S
40782	Golice Wind Project	Poland	Power and Energy	Debt	PRIVATE	19,590,516	S
41031	Hlebprom	Russia	Agribusiness	Equity	PRIVATE	10,000,000	S
41222	Bodrum Water	Turkey	Municipal & Env Inf	Debt	STATE	8,260,128	S
41370	Banja Luka to Dobož Road	B&H	Transport	Debt	STATE	150,000,000	L
41492	Bishkek Public Transport Project	Kyrgyz Republic	Municipal & Env Inf	Debt	STATE	7,960,905	S
41547	MAK Phase II	Mongolia	Natural Resources	Debt	PRIVATE	253,353,827	L
41795	Kazexportastyk - Equity	Kazakhstan	Agribusiness	Equity	PRIVATE	33,629,774	S
41840	Hotel Leipzig	Ukraine	Property and Tourism	Debt	PRIVATE	23,646,252	S
42161	Russian Standard Bank - Senior loan	Russia	Insurance & Financial Services	Debt	PRIVATE	97,273,653	L
42168	Split Wastewater Project	Croatia	Municipal & Env Inf	Debt	PRIVATE	25,000,000	S
42218	Chisinau Urban Road Sector Project	Moldova	Municipal & Env Inf	Debt	STATE	10,300,000	S
42219	Ombla HPP	Croatia	Power and Energy	Debt	STATE	80,000,000	L
42313	Ukraine MSME Lending Framework	Ukraine	Bank Lending	Debt	PRIVATE	78,820,840	L
42387	European Pension Fund	Russia	Insurance & Financial Services	Debt & Equity	PRIVATE	16,207,109	S
42411	Globino	Ukraine	Agribusiness	Debt	PRIVATE	19,705,210	S
42421	EPS Hydropower Plants	Serbia	Power and Energy	Debt	STATE	45,000,000	S
42496	ProCredit Bank Serbia - Senior Loan	Serbia	Small Business Finance	Debt	PRIVATE	30,000,000	S
42551	Aura Centre Surgut	Russia	Property and Tourism	Debt	PRIVATE	45,000,000	S
42552	Ukreximbank SME EE Loan	Ukraine	Bank Lending	Debt	PRIVATE	39,410,420	S
42557	Capljina Water Supply	B&H	Municipal & Env Inf	Debt	STATE	5,000,000	S
42621	Sadovaya Coal Recycling	Ukraine	Natural Resources	Debt	PRIVATE	28,375,502	S
42640	Polkomtel Privatisation	Poland	ICT	Equity	PRIVATE	200,000,000	L
42677	Cersanit Syzran	Russia	Manufacturing and Services	Debt	PRIVATE	23,000,000	S
42754	Croatia Control ATM Modernisation Project	Croatia	Transport	Debt	STATE	47,000,000	S
42796	Astarta IV	Ukraine	Agribusiness	Debt	PRIVATE	46,100,000	S
42831	Hambledon Mining - Equity & Debt	Kazakhstan	Natural Resources	Debt & Equity	PRIVATE	16,464,791	S
42845	Danube	Serbia	Agribusiness	Debt	PRIVATE	35,000,000	S
42870	Georgian Agricultural Finance Facility	Georgia	Bank Lending	Debt	PRIVATE	40,000,000	S
42876	SVL Group	Russia	Transport	Debt	PRIVATE	15,833,530	S
42889	Sarajevo Urban Roads Development Project	B&H	Municipal & Env Inf	Debt	STATE	16,500,000	S
42919	Kronospan Belarus Particleboard/MFC project	Belarus	Manufacturing and Services	Debt	PRIVATE	50,000,000	L
42947	DCS	Romania	ICT	Debt	PRIVATE	15,000,000	S
42991	Severtar	Russia	Manufacturing and Services	Debt	PRIVATE	10,049,657	S
43011	Kirishi Rouble Loan	Russia	Agribusiness	Debt	PRIVATE	11,670,347	S
43017	Trivon - Equity	Russia	ICT	Equity	PRIVATE	23,871,454	S
43067	MoSEFF II - Moldovan Sustainable Energy FF Extension & Re	Moldova	Bank Lending	Debt	PRIVATE	57,000,000	L
43069	Tiryaki	Turkey	Agribusiness	Debt	PRIVATE	39,410,420	S
43094	Roads Reconstruction and Upgrading Project	Azerbaijan	Transport	Debt	STATE	591,156,302	L
43138	Polish Enterprise Fund VII (F.Enterprise Fund 2012)	Regional	Equity Funds	Equity	PRIVATE	80,000,000	L
43155	PSSF - Societe Generale Banka Srbija SME Competitiveness	Serbia	Bank Lending	Debt	PRIVATE	10,000,000	S
43260	Credit Bank of Moscow - Equity Investment	Russia	Bank Equity	Equity	PRIVATE	75,630,265	L
43302	ETC Non-Bank MFI Framework IV	Regional	Bank Lending	Debt	PRIVATE	78,820,840	L
43304	TurAFF - TEB Subordinated Loan	Turkey	Bank Lending	Debt	PRIVATE	101,233,546	L
43306	Procredit Bank Romania Syndicated Loan	Romania	Small Business Finance	Debt	PRIVATE	5,000,000	S
43377	Turkven Private Equity Fund III	Turkey	Equity Funds	Equity	PRIVATE	27,588,050	S
43439	GFT - Getrag FORD Slovakia	Slovak Republic	Manufacturing and Services	Debt	PRIVATE	50,000,000	L
43440	Metsa Tissue - Project Copernicus	Regional	Manufacturing and Services	Debt	PRIVATE	35,000,000	S
43456	Western Balkans SME Platform: ENEF & ENIF	Regional	Equity Funds	Equity	PRIVATE	34,000,000	S
43495	Ohridska Banka - SME CL 2012	Fyr Macedonia	Bank Lending	Debt	PRIVATE	10,000,000	S
43502	ROSEFF SME Energy Efficiency - BRD	Romania	Bank Lending	Debt	PRIVATE	20,000,000	S
43515	Bank Zachodni WBK S.A. Equity Investment	Poland	Bank Equity	Equity	PRIVATE	78,341,812	L
43542	XacBank Syndicated Loan	Mongolia	Small Business Finance	Debt	PRIVATE	7,882,084	S
43703	Baring Voslok Private Equity Fund V	Regional	Equity Funds	Equity	PRIVATE	78,820,840	L
43772	Bulgarian ESCO Fund II	Bulgaria	Manufacturing and Services	Debt	PRIVATE	10,000,000	S
43839	ACBA Credit Agricole MSME Loan	Armenia	Bank Lending	Debt	PRIVATE	11,823,126	S
43875	Khan Bank - parallel senior loan	Mongolia	Bank Lending	Debt	PRIVATE	19,705,210	S
43991	Project Chess	Poland	Manufacturing and Services	Debt & Equity	PRIVATE	40,114,785	S

Breakdown of sample by sector, country, P/F class, instrument and size

Sector	# of ops	%
Agribusiness	8	13%
Bank Equity	2	3%
Bank Lending	11	19%
Equity Funds	4	7%
ICT	3	5%
Insurance & Financial Services	2	3%
Manufacturing and Services	7	12%
Municipal & Env Inf	6	10%
Natural Resources	4	7%
Power and Energy	3	5%
Property and Tourism	2	3%
Small Business Finance	3	5%
Transport	4	7%

Size	# of ops	%
L	16	27%
S	43	73%

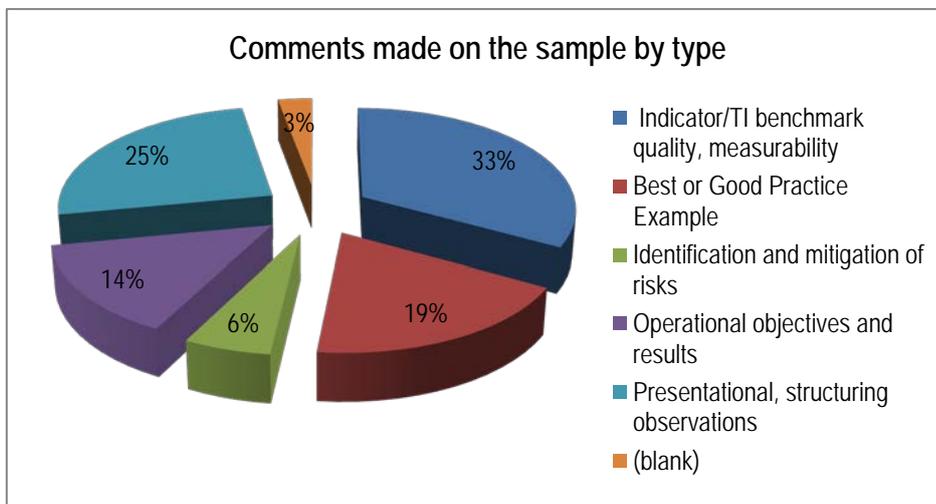
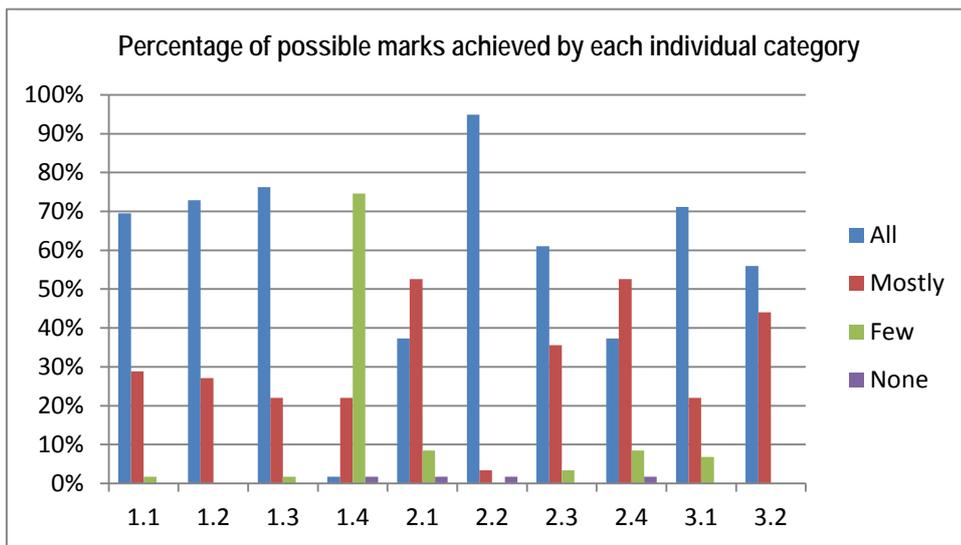
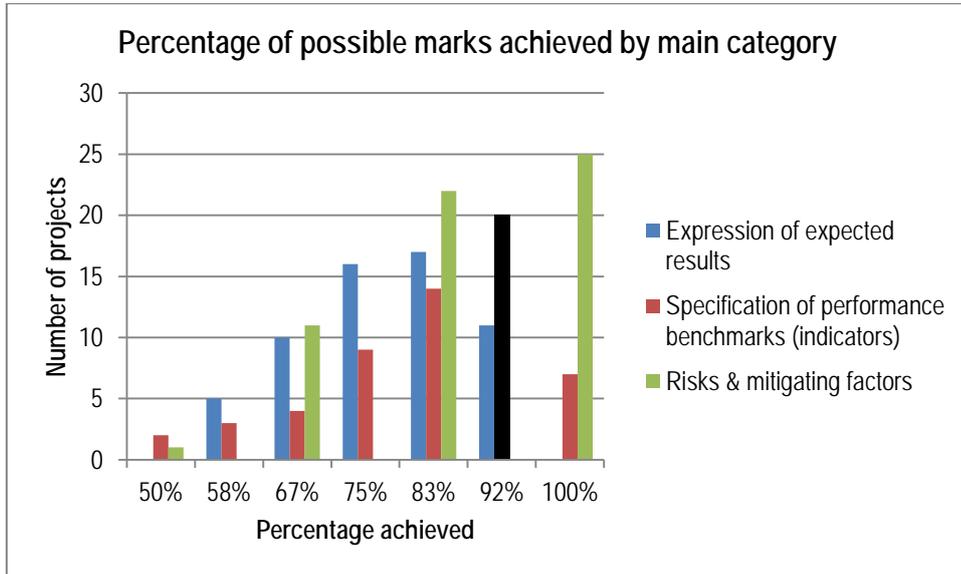
Instrument	# of ops	%
Debt	46	78%
Debt & Equity	3	5%
Equity	10	17%

Total variance (total population) approx. 7%

Country	# of ops	%
Armenia	1	2%
Azerbaijan	1	2%
Belarus	1	2%
B&H	3	5%
Bulgaria	1	2%
Croatia	3	5%
FYR Macedonia	1	2%
Georgia	2	3%
Kazakhstan	2	3%
Kyrgyz Republic	1	2%
Moldov a	2	3%
Mongolia	4	7%
Poland	4	7%
Romania	3	5%
Russia	10	17%
Serbia	4	7%
Slov ak Republic	1	2%
Turkey	4	7%
Ukraine	6	10%
Regional	5	9%

P/F class	# of ops	%
PRIVATE	49	83%
STATE	10	17%

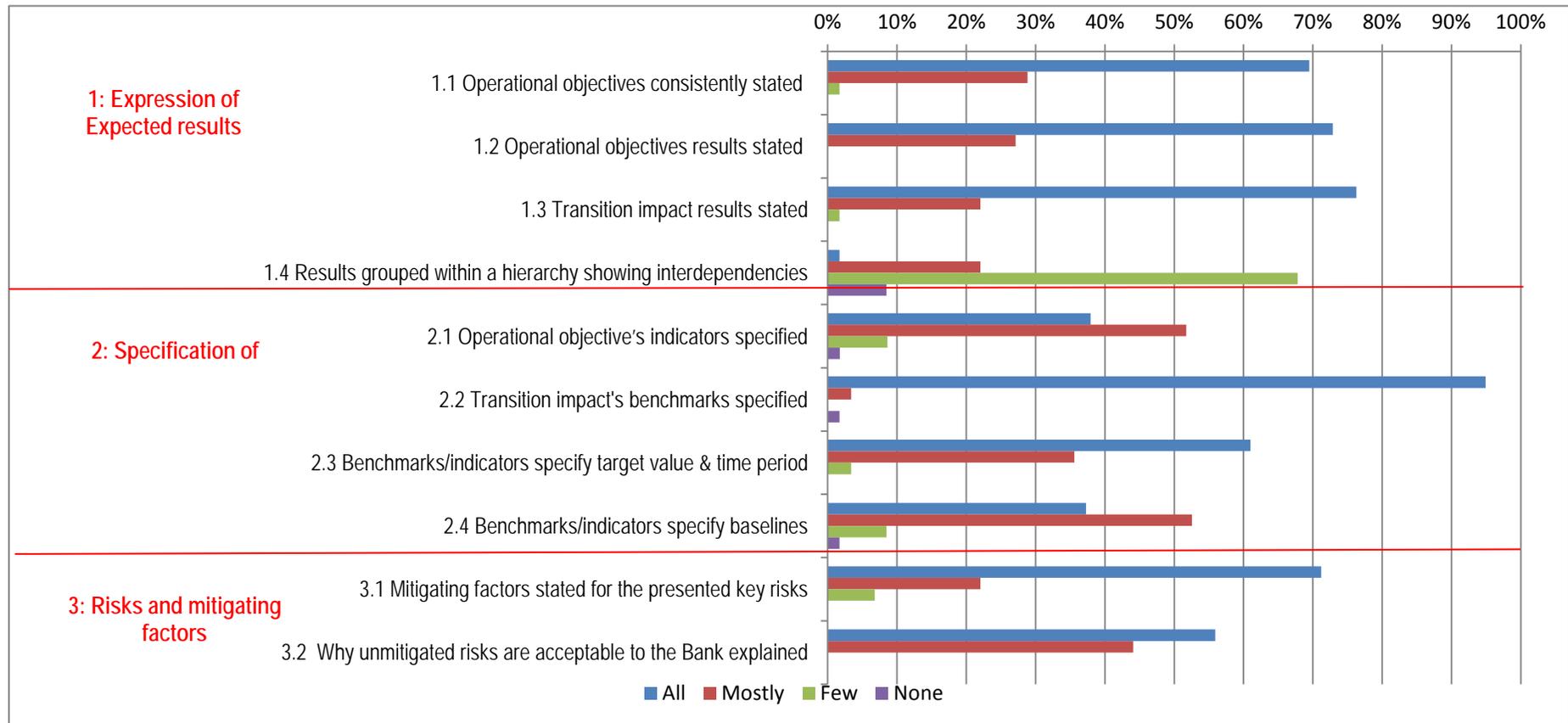
Annex 3: Statistics from the assessment



Special Study:

Performance metrics: how well do EBRD projects specify expected results?

Results per main and individual category



Special Study:

Performance metrics: how well do EBRD projects specify expected results?

Annex 4: Deviations in consultants ratings

