



Mid-term Evaluation of IDB-9 Commitments

The Development Effectiveness Framework and the Development Effectiveness Overview

Background Paper



Inter-American Development Bank
March 2013

Classification: Public Document
Original Version: English



This work is distributed under a Creative Commons license (CC BY-NC-ND 3.0). You are free to copy, distribute and transmit this work to third-parties, under the following conditions:



Attribution - You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).



Non-Commercial - You may not use this work for commercial purposes.



No Derivative Works - You may not alter, transform, or build upon this work.

Waiver - Any of the above conditions can be waived if you get permission from the copyright holder.

© **Inter-American Development Bank, 2013**

Office of Evaluation and Oversight

1350 New York Avenue, N.W.

Washington, D.C. 20577

www.iadb.org/evaluation

ABSTRACT

The Inter-American Development Bank (IDB or Bank) has worked hard to increase its capacity to report on results. The importance of this issue is reflected in many of the requirements that stem from the Ninth General Capital Increase (IDB-9) process. The IDB-9 Agreement mandates that the IDB report on the results of its projects through an enhanced Development Effectiveness Framework (DEF). It requires that the Bank revise the Development Effectiveness Matrix (DEM) as a tool for assessing evaluability in projects; define a minimum threshold for the evaluability score; strengthen the role of the Operations Policy Committee (OPC) to ensure minimum evaluability standards; strengthen the tools for project monitoring and results reporting to provide more useful information; and report on results annually in the Development Effectiveness Overview (DEO). The DEO is also supposed to report on the findings of impact evaluations and on the recommendations of evaluations of the Office of Evaluation and Oversight (OVE).

Many of these requirements have been implemented, and most of those not yet implemented are expected to be finalized soon. Though it is still early in the process, the effectiveness of implementation of the DEF to date appears positive, with some challenges remaining. The DEM for sovereign-guaranteed (SG) projects has been effectively implemented, while the DEM for non-sovereign-guaranteed (NSG) projects is still a work in progress. The focus of OPC meetings on evaluability has increased, though evaluability issues in NSG operations were not identified. The Progress Monitoring Report (PMR) for SG projects has been implemented and is being revised to address some of the challenges it still faces. One significant issue is the lack of integration across these various tools, which appears to limit the usefulness of the PMR system. Finally, the new Project Completion Report (PCR) has not yet been launched but is under preparation. It is critical that it be launched soon, as it is a key part of the Bank's results architecture for SG projects.

The DEO is not yet fulfilling its mandate to report on the Bank's overall development results. The DEOs to date have been effective primarily in highlighting and discussing selective areas of Bank activity. Some of the weaknesses of the DEO result from the incomplete implementation of the DEF, including the absence of reliable PCRs and the failure of the NSG DEM to measure evaluability. In addition, the DEO does not explain the reasons for selecting the cases presented and discussed in the analysis of the Bank's priority sectors, and the impact evaluations promised are not followed up in later editions. Finally the DEO has not to date had an analysis of the economic rates of return of projects approved in the year of reference or a discussion of OVE recommendations, though the latter is expected in 2013.

Moving forward, the Bank should complete the implementation of the DEF and continue to refine the tools that are already in place. To make the most of these tools, it is important to ensure that they are integrated and to promote a culture of planning for results and a willingness to report on problems and failures as well as successes. The DEO should seek to focus on lessons learned from both success and failure and to serve more as a tool for accountability.

PREFACE

The Inter-American Development Bank (IDB) is in a period of rapid change, responding to both the economic dynamism of the Region it serves and the increasing competition in the international financial marketplace. Over the past decade, countries in Latin America and the Caribbean have gained greater access to alternative sources of finance and an increasingly ability to generate and share knowledge among themselves. Like other multilateral development banks, IDB is seeking to adapt to this changing international landscape by ensuring that it is responsive to borrowing countries' needs and putting strong emphasis on effectiveness in its use of scarce resources.

In 2010 the IDB's Board of Governors approved the 9th General Capital Increase of the IDB (IDB-9). The IDB-9 Agreement laid out a series of reforms intended to strengthen the strategic focus, development effectiveness, and efficiency of the IDB to help it remain competitive and relevant in the years ahead. As part of that Report, IDB's Office of Evaluation and Oversight (OVE) was charged with conducting a midterm evaluation—to be presented to the Board of Governors in March 2013—to assess IDB's progress in implementing those reforms. The full evaluation is available at www.iadb.org/evaluation.

This paper is one of 22 background papers prepared by OVE as input to the IDB-9 evaluation. It seeks to determine whether one portion of the IDB-9 requirements has been implemented fully and effectively and to offer suggestions to strengthen implementation going forward. The overarching goal of this paper and the entire evaluation is to provide insights to the Governors, the Board, and IDB Management to help make IDB as strong and effective as possible in promoting economic growth and poverty reduction in Latin America and the Caribbean.

TABLE OF CONTENTS

ABBREVIATIONS AND ACRONYMS

EXECUTIVE SUMMARY

I.	INTRODUCTION	1
A.	IDB-9 Requirements	1
B.	Purpose, scope, and methodology	2
II.	FINDINGS	3
A.	The Development Effectiveness Matrix and evaluability	3
1.	The SG DEM.....	5
2.	The NSG DEM.....	10
B.	The role of the Operations Policy Committee	13
C.	The Progress Monitoring Report	14
D.	The Project Completion Report	22
E.	The Development Effectiveness Overview	28
III.	SUGGESTIONS GOING FORWARD.....	32

LIST OF PERSONS INTERVIEWED

ANNEX A:	<u>THE REQUIREMENTS ON THE DEF AND THE DEO IN THE REPORT ON THE NINTH GENERAL INCREASE IN THE RESOURCES OF THE INTER-AMERICAN DEVELOPMENT BANK (AB-2764)</u>
ANNEX B:	<u>THE BANK’S STRATEGIC PRIORITIES IN THE DEO</u>
ANNEX C:	<u>THE IMPACT EVALUATIONS PROMISED IN THE DEO</u>
ANNEX D:	<u>COMPARISON BETWEEN THE EFFECTIVENESS OVERVIEW OF DIFFERENT DEVELOPMENT BANKS</u>

This background paper was prepared by Anna Crespo (team leader), Kris Hallberg, Nelson Ruiz, Lucia Martin, Daniel Alonso, Daniel Broid, Patricio Zambrano, and Thiago Dal-Toe. It is an input to OVE’s broader IDB-9 evaluation. All background papers were thoroughly reviewed and discussed within OVE and shared with IDB management for comments. The other background papers and full IDB-9 evaluation can be found at www.iadb.org/evaluation.

ABBREVIATIONS AND ACRONYMS

DEF	Development Effectiveness Framework
DEM	Development Effectiveness Matrix
DEO	Development Effectiveness Overview
ECG	Evaluation Cooperation Group
ERM	Eligibility and Risk Review Meeting
ERR	Economic rate of return
GPS	Good Practice Standards
IDB	Inter-American Development Bank
IDB-9	Ninth General Capital Increase of the IDB
KNL	Knowledge and Learning Sector
LAC	Latin America and the Caribbean
NSG	Non-sovereign guarantee
OMJ	Opportunity for the Majority Sector
OPC	Operations Policy Committee
OVE	Office of Evaluation and Oversight
PBL	Policy-based loan
PBP	Programmatic policy-based loan
QRR	Quality and Risk Review
RES	Department of Research and Chief Economist
SCF	Structured and Corporate Finance Department
SG	Sovereign guarantee
SPD	Office of Strategic Planning and Development Effectiveness
VPP	Vice-Presidency for Private Sector
XPMR	Expanded Progress Monitoring Report
XPSR	Expanded Project Supervision Report

EXECUTIVE SUMMARY

The Inter-American Development Bank (IDB or Bank) has worked hard to increase its capacity to report on results. The importance of this issue is reflected in many of the requirements that stem from the Ninth General Capital Increase (IDB-9) process. Of the 13 requirements in the Cancún Declaration, 3 are related to the Development Effectiveness Overview (DEO), the Development Effectiveness Matrix (DEM), and project evaluability. The “Report on the Ninth General Increase in the Resources of the Inter-American Development Bank” (IDB-9 Agreement) describes what is expected from the tools created under the Development Effectiveness Framework (DEF) and reinforces the mandates in the Cancún Declaration. It focuses strongly on the need for the Bank to increase its capacity to report on results by strengthening the tools it has in place: “We recognize that the Bank must not simply become larger, but that it must also become more effective at achieving its mandates. Thus, the Board of Governors agrees to the following reforms: ... improving the results measurement framework to identify for every project whether specific and tangible results have been achieved” (CS-3868-1). This goal is at the heart of the DEF.

The references to the DEF in the IDB-9 Agreement do not appear to take into account the Development Effectiveness Framework for the Bank’s non-sovereign guaranteed (NSG) projects (GN-2471-1) presented to the Board in March 2008, except for recognizing the different DEMs used by sovereign-guaranteed (SG) and NSG projects. The tools described in the Agreement as the ones composing the DEF are used only for SG projects and differ substantially from the ones used by NSG operations.

This paper investigates whether the IDB-9 requirements related to the DEF have been fully and effectively implemented. It addresses two main questions: Has the DEF been fully and effectively implemented? and Is the DEO reporting on the effectiveness of the Bank’s support to Latin America and the Caribbean? The two questions are highly correlated, as the DEO is informed primarily by the information produced by the DEF tools. It is important to note that it is still too early to assess the full effectiveness of the DEF, as most of the projects approved since the implementation of the DEM in April 2009 are still in implementation.

Development Effectiveness Matrix

The Bank has been working with two different types of DEM—one for SG and one for NSG projects. The SG DEM included an evaluability self-assessment from the beginning, while the NSG DEM was initially designed to be a development results assessment tool and an assessment of evaluability was added in 2011. The two DEMs use different formats and criteria.

The Office of Evaluation and Oversight (OVE) carried out an analysis of the SG and NSG DEMs in 2012.¹ The analysis of the SG DEM found significant progress in design and implementation and suggested further improvements to better assess evaluability. Management agreed with most of OVE's recommendations and is implementing them. The first projects with the revised DEM are expected to be sent to the Board in 2013.

The analysis of the NSG DEM concluded that a full revision of this instrument was needed. OVE's assessment of a sample of projects found that the NSG DEM did not properly capture evaluability issues in these projects, and it yielded misleading evaluability scores. The Bank has been working on a new DEM, which is expected to be submitted to the Board for approval by the end of 2012. This new instrument closely follows the SG DEM, with the exception of a few criteria and the weights proposed.

The Operations Policy Committee (OPC) has played a partial role in ensuring minimum evaluability scores. Analysis of the minutes of OPC meetings suggests that since 2008 the frequency of discussions about evaluability issues has increased. Also, the content of the discussions has changed, reflecting the different stages of implementation of the DEM. During the past year the discussions have been less frequent, which may be a result of Management's policy of ensuring that projects reach a minimum evaluability score of 5 before they are sent for approval. However, the evaluability issues identified by OVE in private sector loans and the fact that the tool used to assess their evaluability was inappropriate have not been discussed in OPC meetings.

Progress Monitoring Report

The Progress Monitoring Report (PMR) system has been fully implemented and appears to be an improvement on the prior system. It continues to face some challenges that Management is aware of. The high degree of complexity and uncertainty that characterize the design of many projects, combined with the short period of time for preparation, can make it difficult for staff to properly plan outputs and costs. The Performance Index, which is measured by comparing the actual delivery of outputs and their cost with the planned values, has not been widely accepted by Bank staff. The index relies on the quality of the data put into the system, which in turn depends on the capacity of the executing agency to provide timely and good-quality data—capacity that in practice can be quite low. Also, the tools used in the Bank are not fully integrated, making it more difficult for staff to prepare the PMR.

Project Completion Report

Rather than implement an XPMR system, as envisioned in the DEF to take place in 2011, Management has opted to strengthen the Project Completion Report (PCR) system. This has been delayed, however, and is now expected to be launched in early 2013. In 2012 OVE evaluated the main challenges facing the existing PCR system and recommended

¹ *2011 Evaluability Review of Bank Projects* (RE-397-1), June 2012, available at www.iadb.org/evaluation.

revising the template to integrate the tool with the DEM and the PMR.² OVE also stressed the importance of incentives for honest reporting on the achievement of project objectives. Management agreed with the recommendations and is working to implement them. Management has also revised the budget for PCR production to ensure that enough resources will be allocated for the production of good-quality reports.

Development Effectiveness Overview

The incomplete implementation of the architecture of the DEF has limited the capacity of the Development Effectiveness Overview (DEO) to report on results, as some of the tools in the self-evaluation framework are not providing reliable information (notably PCRs and the NSG DEM). In addition, the DEO does not explain the rationale for selecting the evaluations discussed in it, and thus it is not possible to generate any general conclusion about the Bank's effectiveness. It appears that the DEO has been used largely to highlight the areas the Bank has been working on. The DEO does not follow up on promised impact evaluations and does not include an analysis of the economic rate of return of the portfolio approved in the year of reference. OVE evaluation recommendations were not included in early DEOs but are expected to be included in 2013.

Conclusion and Suggestions Moving Forward

In sum, the Bank has made substantial progress in implementing the IDB-9 requirements related to project-level results reporting. Some critical parts of the DEF—most notably PCRs—are still under construction, and Management indicates that they are likely to be completed in early 2013. In part because of this incomplete results architecture, the DEOs have not met the expectation in IDB-9 of reporting fully on Bank results. OVE makes a number of suggestions going forward:

- ***First, ensure that the various tools are integrated with one another.*** It is not sufficient to have projects that are perfectly evaluable if they are not fully monitored and evaluated. Similarly, it is not possible to use a monitoring tool effectively if results indicators were not properly defined from the beginning. The standards of evaluability required from project teams need to be those that will make them more successful in preparing the PMRs throughout the life of the project and will allow them to have the information to create meaningful PCRs.
- ***Second, support a culture of planning for results and a willingness to report on problems and failures.*** The DEF tools will be informative only to the extent that project teams do proper planning and report problems. The Performance Index relies heavily on the quality of the data put into the system and the planning of output delivery and cost at design. Similarly, the lessons

² Review of the Project Completion Reporting System for Sovereign Guarantee Operations (RE-417), July 2012, available at www.iadb.org/evaluation

learned from PCRs will be most useful when teams are willing to report not only on success but also on failure.

- ***Third, strengthen the evaluation of private sector operations.*** The instruments used to monitor and evaluate NSG operations are different from those used for SG projects. The lack of attention given to the NSG DEF in the IDB-9 indicates that these projects have generally been off the radar screen with regard to results reporting. However, knowledge about what works and what do not work in NSG operations will be critical as the Bank's engagement in this area grows.
- ***Finally, strengthen the DEO as an accountability tool.*** It is also important that the DEO report accurately on the full range of Bank experiences, both successes and failures. Refining the tools under the DEF will provide better information for the DEO, and including the results of impact evaluations and OVE recommendations will also help in this regard.

I. INTRODUCTION

- 1.1. The Inter-American Development Bank (IDB or Bank) has worked hard to increase its effectiveness. Part of this effort is reflected in the approval of the Development Effectiveness Framework (DEF) in 2008 and the annual publication of the Development Effectiveness Overview (DEO) report. The DEF aimed at improving the Bank's effectiveness by "generating a body of knowledge about 'what works' in meeting the region's development challenges" (DEO 2008/09). It consists of tools through which projects are assessed, monitored, and evaluated from their design until their completion. Each part of the DEF generates inputs that are expected to be presented and discussed in the DEO, the only corporate report on the general effectiveness of the IDB.
- 1.2. The DEF is composed of three tools:
 - The Development Effectiveness Matrix (DEM) is used to assess a project's ability to report on results at completion—that is, its evaluability.
 - The Progress Monitoring Report (PMR) is used to monitor outcomes, outputs, and output costs throughout the execution of the project.
 - The Project Completion Report (PCR)³ is produced for each project at completion, showing evidence of, and discussing, the achievement of the development objectives of the project vis-à-vis their intended objectives, and the sustainability of those achievements.

A. IDB-9 Requirements

- 1.3. The IDB-9 Agreement emphasized the need to increase the Bank's capacity to report on results: "We recognize that the Bank must not simply become larger, but that it must also become more effective at achieving its mandates. Thus, the Board of Governors agrees to the following reforms: ... improving the results measurement framework to identify for every project whether specific and tangible results have been achieved" (CS-3868-1). The agreement mentioned four aspects of the framework: standards of evaluability, the role of the Operations Policy Committee (OPC) regarding evaluability, the design of tools, and the report on results.
- 1.4. The Cancún Declaration (AB-2728) had three specific recommendations about the DEO and the DEM. For the DEO, the IDB was expected to "disclose project-level reporting with stronger metrics" and "strengthen...Development

³ In GN-2489, the DEF suggested that the completion report would be called not a PCR but an Expanded Progress Monitoring Report, but this has not been put into practice.

Effectiveness.” As regards these reforms, the Declaration states that the DEO should be published annually and should include such information as ex-ante project evaluability, compliance with institutional priorities, economic rate of return (ERR) calculations for projects approved in the year of reference, and results of ex-post impact evaluations completed in the year of reference. It should also include the analysis of recommendations by the Office of Evaluation and Oversight (OVE) and discuss how they can help to improve loan quality. The third reform—“Ensure quality of loan portfolio through the use of the new development effectiveness matrix”—requires strengthening the role of the OPC to ensure that projects meet a minimum evaluability score.

- 1.5. Similar reforms and additional mentions of the DEF and DEO were in the “Report on the Ninth General Increase in the Resources of the Inter-American Development Bank” (AB-2764). This document’s requirements regarding the DEM, the role of the OPC, and the DEO (see Annex A, Box 1) are similar to those of the Cancún Declaration: annual disclosure of the DEO, strengthening the role of the OPC to ensure minimum standards of evaluability, and defining the steps to increase the DEM’s capacity to assess evaluability. An annex to the report shows what is expected from the DEF (see Annex A, Box 2), clearly defining the DEM, the PMR, and the Expanded PMR (XPMR).
- 1.6. The IDB-9 agreement report does not mention the tools under the Development Effectiveness Framework for Structured and Corporate Finance Department (SCF) projects (GN-2471-1), which was presented to the Board in March 2008, except for recognizing that the DEM should be used by SG and NSG projects. While the evaluability assessment is done by tools with the same name for SG and NSG projects, the monitoring and self-evaluation tools for NSG are different from those for SG operations. NSG operations use two tools: the PSR (Project Supervision Report), which is an annual monitoring report tool, and the XPSR (Expanded Project Supervision Report), which is prepared when projects reach early operational maturity. The procedures for and format of these tools are distinct from those used for SG operations and were not analyzed in this report.

B. Purpose, scope, and methodology

- 1.7. This report aims to answer two main evaluative questions: Has the DEF been fully and effectively implemented? And is the DEO properly reporting on the effectiveness of the Bank’s support to Latin America and the Caribbean (LAC)? (It is important to note that it is too early to assess the effectiveness of the DEF: since most of the projects approved since the implementation of the DEM in April 2009 are still disbursing,⁴ there is not yet a representative group of projects to

⁴ To assess the effectiveness of the DEF, it is necessary to be able to observe whether projects have increased their capacity to report on results. This means that one would need to observe projects approved and completed under the DEF. So far only 87 projects, of which 76% are policy-based or programmatic policy-based loans, fulfill that criterion; this is not a representative sample.

evaluate.) The two questions are strongly correlated, as the DEO is mostly fed by the information produced by the DEF tools. Therefore, to answer these questions, this report will describe and analyze the progress in implementing the DEM and the PMR, and the challenges of the PCR system.

- 1.8. This paper is based on an independent review of a sample of DEMs, PMRs, and PCRs,⁵ a desk review of all relevant documents, and interviews with Bank staff. The past editions of the DEO were also analyzed.

II. FINDINGS

C. The Development Effectiveness Matrix and evaluability

- 2.1. In 2008, following the launch of the DEF, Management designed the DEM to be incorporated into all IDB projects before they are submitted to the Board for approval. Two different types of DEMs were developed: one for non-sovereign guarantee (NSG) projects, and one for sovereign guarantee (SG) projects.⁶ The SG DEM included an evaluability self-assessment tool; the NSG DEM was initially designed to be a development results assessment tool, but in 2011, an assessment of evaluability was added to it. The two types of DEM have completely different formats and questions.
- 2.2. Most of the focus on the DEM is related to its assessment of projects' evaluability. The SG DEM, developed by the Office of Strategic Planning and Development Effectiveness (SPD) as a "review checklist," is divided into three sections: Strategic Alignment, Development Outcome – Evaluability, and IDB's Role – Additionality. The evaluability section of the DEM is detailed in RE-397-1 and is the only part of the tool considered in this report. The evaluability score is computed through a weighted average of a set of binary criteria to which a YES (1) or NO (0) should be assigned. The criteria refer to the project diagnosis, project logic, results matrix, economic analysis, monitoring and evaluation plan, and risk assessment. The final score ranges from 0 to 10.
- 2.3. The NSG DEM serves two purposes. First, it justifies an IDB intervention by rating the project's potential development outcomes, its potential contribution to IDB's/country's strategic priorities, and IDB's additionality. Second, it attempts to assess project evaluability through a review of its Development Outcome

⁵ The PCRs reviewed in this report were completed in 2008 and 2009. None of them was directly affected by the implementation of the DEM or the PMR. However, since the system has not been revised, the analysis included in this report remains valid.

⁶ The NSG DEM, which is based on ECG-GPS, was defined in the SCF Development Effectiveness Framework (GN-2473-1) and has been adopted by SCF and OMJ. The SG DEM was developed by SPD, and its original guidelines can be found in GN-2489.

Section. An evaluability score from 0 to 10 is given to each indicator,⁷ and the overall evaluability score is computed on the basis of a weighted average of the indicator scores.

- 2.4. The IDB-9 Agreement had very specific requirements regarding the DEM. In response to these requirements, the SG and NSG DEMs were revised in February 2011, and the changes—including a new template for the DEM and a revised process to prepare the DEM—were implemented in April of the same year (see Table 1).⁸ Despite the delay of about half a year, all the requirements were addressed.

Table 1. Requirements on the DEM and Evaluability in the IDB-9 Agreement^a

Requirement	Implementation
All SG and NSG projects must be rated for evaluability.	Since the implementation of the SG DEM, all projects have been rated for evaluability before approval. The NSG DEM added evaluability scores in February 2011.
The evaluability score includes only the dimensions of evaluability of the DEM.	In either SG or NSG projects, the evaluability score is computed through the answers given to the criteria in the DEM with respect to evaluability and, therefore, includes only the dimensions of the DEM.
The Office of Strategic Planning and Development Effectiveness (SPD) will support teams in meeting evaluability standards from project profile to project proposal, and will validate the final evaluability score for OPC consideration.	SPD has actively supported SG project teams since the project profile, sending comments to the Eligibility and Risk Review Meeting (ERM) and the Quality and Risk Review Meeting (QRR), and validating final scores before they are submitted for OPC consideration. For NSG projects, the Development Effectiveness Officers in the Vice-Presidency for the Private Sector (VPP) have the same role.
The Department of Research and Chief Economist (RES) will review the existing methodologies for scoring evaluability to determine any required improvements.	In December 2010 RES submitted to the Board a review of the existing methodology, with a high focus on the weights used in the DEM.
OVE will report annually to the Board of Executive Directors on project evaluability.	In 2012 OVE presented to the Board the assessment of the 2011 projects based on a sample of projects, which included all the latest revisions of the DEM.
A minimum evaluability threshold of 5 will be required for all operations to be submitted to the Board of Executive Directors.	Before being submitted to the Board for approval, all projects need to score at least 5 in each section of the DEM related to evaluability, which implies that they all need to score 5 or above overall.

⁷ The score is derived from an assessment of (i) the quality of indicators/results, (ii) the rationale for the project score, (iii) the analysis/evidence included, and (iv) the monitoring and evaluation arrangements.

⁸ In this new process, SPD (for SG projects) or the VPP Development Effectiveness Officers (for NSG projects), are expected to provide comments to the teams at the ERM and the QRR. Before projects are sent to the OPC, SPD or the VPP Development Officers validate the final version of the DEM and prepare an evaluability note to be annexed to the Project Proposal.

Requirement	Implementation
In cases of extraordinary humanitarian or financial crisis (acute payment pressures), and upon the presentation of a properly justified request by the President, the Board of Executive Directors may waive the required threshold before approving an operation, as long as the request is accompanied by a timetable giving the shortest possible time period, not to exceed 90 days, for achieving the said threshold.	According to the 2011 DEO, in 2010—the first year for which evaluability was rated—4% of the projects approved were rated below 5 in evaluability. No project approved in 2011 scored below 5 in evaluability. According to the 2011 DEO, “Only 14 [SG] projects of a total of 122 approved were below the threshold at Quality and Risk Review (QRR) ... and had to improve the information and analytics in the loan proposal before being submitted to the Board of Executive Directors for approval.” All projects in the NSG sector, according to the DEO, scored in the highly evaluable range.
DEM is used as a common instrument for rating the evaluability of projects at entry (before approval).	The DEM has been the common instrument used by Management and OVE to assess evaluability. However, there are two different DEMs—the NSG and the SG. SCF and Opportunity for the Majority Sector (OMJ) are working to harmonize the NSG DEM criteria with those of the SG DEM. VPP is expected to present a new version of the SG DEM to the Board in the coming months.

Note: This table was prepared on the basis of the following paragraphs of the IDB-9 Report:

- Para. 4.7: “Governors endorse a further strengthening of the Operations Policy Committee (OPC) and the programming process, by the President of the Bank and Senior Management, to ensure that projects meet minimum evaluability thresholds. In this respect, Management will amend operational procedures, by end of Q3 of 2010, according to the following criteria: (i) all SG and NSG projects must be rated for evaluability; (ii) the evaluability score includes only the dimensions of evaluability of the DEM; (iii) SPD will support teams in meeting evaluability standards from project profile to project proposal, and will validate the final evaluability score for OPC consideration: RES will review the existing methodologies for scoring evaluability to determine any required improvements; OVE will report annually to the Board of Executive Directors on project evaluability (ex-ante), as well as validate achieved results in completed projects (ex-post); (iv) a minimum evaluability threshold of 5 will be required for all operations to be submitted to the Board of Executive Directors; (v) in cases of extraordinary humanitarian or financial crisis (acute payment pressures), and upon the presentation of a properly justified request by the President, the Board of Executive Directors may waive the required threshold prior to approving an operation, so long as the request is accompanied by a timetable giving the shortest possible time period, not to exceed 90 days, for achieving the said threshold.”
- Para. 6.4: “Development Effectiveness Matrix (DEM) is used as a common instrument for rating the evaluability of projects at entry (before approval). The DEM has three categories: strategic relevance, development outcome, and additionality of the Bank’s contribution. Country strategies, public and private sector operations have their specially designed DEMs.”

1. The SG DEM

- 2.5. The effectiveness of the implementation of these requirements can only be judged once the projects under the new DEM are complete, which will not be for a couple of years. However, an analysis of the DEM can shed some light on this issue. OVE has recently submitted to the Board the results of its first evaluability review based on the validation of sample of DEMs of projects approved in 2011,

and in this report (RE-397-1) it also analyzed the quality of the DEM as a tool to assess evaluability. OVE validated about one-third of the projects approved under the new DEM and got final scores similar to those of Management. In most cases of disagreement, OVE had a somewhat lower rating on criteria related to problem diagnosis and project logic. Overall, this is a positive result that indicates significant progress in the SG DEM criteria and tutorial. However, OVE's analysis also suggests that the SG DEM can be further improved to better assess evaluability. Each area for improvement identified in RE-397-1 is discussed below.

- ***“First, the DEM has some criteria with greater subjective content which require more guidance.”*** During the validation exercise, OVE identified problems in validating some of the criteria. Each DEM criterion requires a YES or NO answer, but in many cases OVE evaluators could not reach consensus on a straight answer.⁹ “Indeed, in 21.6% of the cases OVE evaluators judged that the criteria could not be easily answered with a clear YES or NO for the projects they were analyzing. Most of these cases (68.75%) involved one of the first six questions of the DEM [which refer to the Program Diagnosis]” (RE-3971).
- ***“Second, the criteria regarding the indicators are inadequately considered independently of each other.”*** Nearly all questions in the DEM are answered independently,¹⁰ but this approach is inappropriate for the criteria involving indicators. The DEM asks whether at least one indicator has been identified, whether indicators are SMART (specific, measurable, attributable/achievable, relevant, and time-bound), and whether they have baselines, targets, and means of verification. These questions are defined for impact, outcome and later for output indicators separately. A project can have indicators that are not SMART, which by definition are not pertinent, and still get credit for having indicator baselines and targets. Similarly, a project can get credit for having SMART indicators even if there is no baseline—which means that indicators will not be usable for evaluation at the end.
- ***“Third, the criteria on internal and external validity are in reality inter-related, and as currently formulated they do not allow for innovation projects to get credit.”*** Two DEM criteria are #8—“Evidence of the

⁹ OVE evaluators were given the option to point out situations in which neither a straight YES nor a straight NO could be assigned. The two intermediate categories were given as an option to these evaluators, Yes- and No+, which were supposed to be used when the evaluators judged that either most or some of the information was in the project, but not enough for the criteria to be rated as a YES.

¹⁰ When answering, the DEM teams should fill in the cost-benefit analysis or the cost-effectiveness analysis section; the score of the Economic Analysis section is computed as the maximum score of the two. The same is done for the method chosen by teams for the evaluation. More than one method might be chosen, and the score is given as the maximum among the options chosen: the magnitude of the score goes up from “before and after without attribution” to “random assignment.”

effectiveness of the intervention(s) is based on existing evaluations of interventions in other or similar contexts (internal validity),” and #9—“Information about the applicability of the intervention in the country where it is implemented is provided (external validity).”, These criteria need to be considered together, as external validity is related to the extent to which internally valid results will be held to be true for other interventions in similar cases.¹¹ Additionally, the tutorial of these criteria needs to be clarified in terms of how innovative projects will be considered. A pioneering project, which defines well what model is to be tested and how it is to be tested, may not be less evaluable than a project that is replicating what other interventions have done in the past.

- ***“Fourth, the use of the term impact can lead to confusion, and OVE observed that many projects did not properly identify impacts.”*** The DEM requires identification of impacts and SMART indicators to monitor them. DEM questions 12 to 17 are related to the identification of medium- or long-term impacts and their indicators. Yet what several projects define as the impact of the project is, in fact, an outcome. This may be partly because the term *impact* has been used with “multiple and sometimes conflicting meanings” (Evaluation Cooperation Group [ECG] Public Sector Good Practice Standards [GPS], 3rd edition).¹² In addition, impacts and the corresponding indicators have very little weight in the DEM: together questions 12 to 17 add to less than 0.05 of the overall score. This generates an incentive for teams to give little importance to this section of the DEM. As a consequence, staff might attempt to identify impacts, but without much rigor.
- ***“Fifth, the Risk Management section of the DEM is only a compliance checklist rather than an assessment of the quality of the risk analysis, and the high weight given to this section in the absence of a better assessment mechanism tends to inflate projects’ evaluability scores unjustifiably.”*** The Risk Management section of the DEM “rates the completeness of the information presented in the project’s risk matrix, namely if all risks are

¹¹ During the validation exercise, OVE found evidence that this is not the case: 3 out of 16 projects did not consider them together, as they did not score positively on internal validity but did so on external validity.

¹² The DEM tutorial has the following definition: “Impact of the Project: Positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended. Impacts generally refer to changes in living conditions or institutions and are not under the direct control of those responsible for project management.” The term *impact evaluation* is usually based on attribution and measures changes in outcomes. For instance, the International Initiative for Impact Evaluation (3ie) defines it in the following way: “Impact evaluations measure the net change in outcomes amongst a particular group, or groups, of people that can be attributed to a specific program using the best methodology available, feasible and appropriate to the evaluation question(s) being investigated and to the specific context.” Outcome is defined in the latest public sector ECG-GPS as the final level in the results chain, reflecting the objectives of the project, which is similar to the definition in the DEM: “The outcome (s) should describe what is expected to be different as a result of the delivery of project outputs.”

identified, mitigation measures have been defined and indicators to monitor mitigation measures have baselines and targets to track their implementation” (GN-2489-4).¹³ This section does not include questions about the quality of the risk assessment quality (although many in the Bank believe that it does).¹⁴ In addition, the high weight given to this section tends to inflate the projects’ evaluability score. Since the Risk Matrix is a mandatory part of the project preparation process,¹⁵ most of the projects tend to score 10, regardless of the quality of the analysis in the matrix.¹⁶

- ***“Sixth, it is difficult to assess a PBL or a PBP using the DEM, since some of the criteria cannot be directly applied and the unit of analysis is unclear.”*** Although policy-based loans (PBLs) and programmatic policy-based loans (PBPs) are required to present some sort of economic analysis, neither a cost-benefit analysis nor a cost-effectiveness analysis could be done, as the costs of the proposed reforms could not possibly be measured. In some cases the alternative used by the project teams was to perform the analysis for one of the components of the project for which the measurement of the costs was possible. Although these are relevant analyses, they cannot be considered sufficient for the project as a whole, as many benefits go beyond the components for which a cost-benefit or cost-effectiveness analysis could be done. Similarly, the criteria on results frameworks were not applied uniformly across the PBPs analyzed. The conditionalities were accepted as output indicators in some cases and rejected in others; but OVE could not find documents related to the standards for accepting or rejecting them as

¹³ This section is composed of four criteria: (i) identified risks have been rated for magnitude; (ii) identified risks have been rated for likelihood; (iii) major risks have proper mitigation measures; and (iv) mitigation measures have indicators for tracking their implementation. Except for the third criterion, the report contains no assessment of the quality of the risk analysis in the project.

¹⁴ For instance, the message sent by the Office of the Executive Vice President to Bank staff members on April 28, 2011, regarding Risk Procedures and Fiduciary Arrangements stated, “The DEM will also continue to review the quality of the risk assessment.” This confusion might be a result of the fact that when revising this section of the DEM the intention was to “provide the Board of Executive Directors with an overall assessment of the severity of the risks identified for the project.”

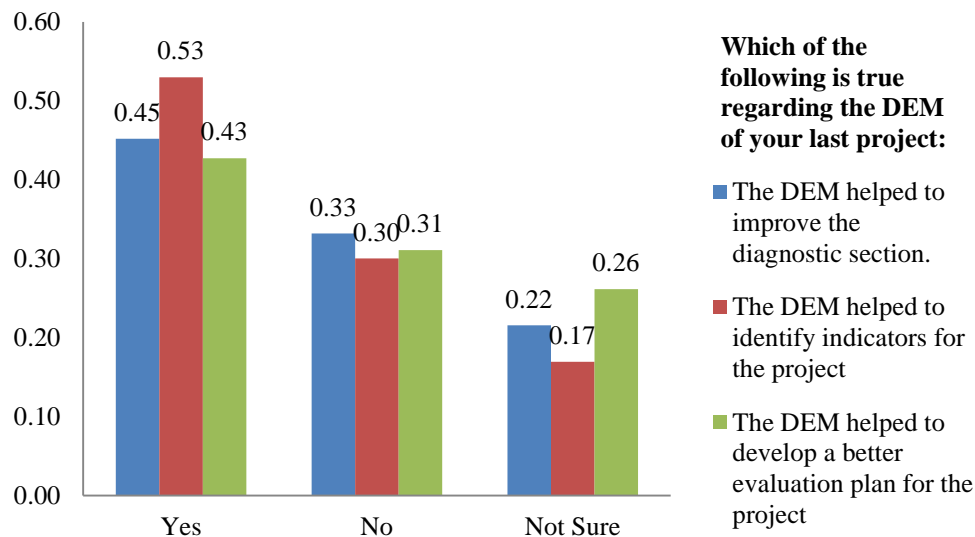
¹⁵ The Procedures for Processing Sovereign Guaranteed Operations, updated in July 2010, lists all the necessary parts of the Proposal for Operation Development to be sent for QRR. Among them is the Updated Risk Assessment Template, “which is not part of the document that will be approved, but it is an essential tool to facilitate the quality of risk review process.”

¹⁶ Among the 122 projects approved in 2011 and analyzed in the DEO, almost 70% scored 10, while two projects scored 5 and the remainder scored 7.5, mostly because they lacked of indicators to track the mitigation of the identified risks. Given that each section has equal weight, scoring 10 in the Risk Management section implies already reaching 2.5 of the total score, which is 50% of the threshold. Although GN-2489-4 argues that “the evaluability score is robust to the specific weights,” giving 10 in the Risk Management section to most projects without an assessment of quality inflates the score of those projects with poor risk analysis. While the 2011 DEO defined that about 90% of the projects were classified as highly evaluable, if the same grouping were done for projects when considering the evaluability score as only the average of the first three sections (i.e., excluding the risk management section), only about 80% would fall in the highly evaluable category.

indicators. Other particularities of PBLs and PBPs are related to how to assess the evaluability of a second phase, or of projects that are expected to be complemented by an investment loan. In both cases, there are many issues in applying the DEM directly, as it is unclear whether the DEM should consider the program as a whole or the individual loans separately and independently.

- 2.6. Despite the problems identified in the instrument, Bank staff have an overall positive perception of the usefulness of the SG DEM. Figure 1 displays the results of an online survey in which OVE asked staff their opinion about the DEM.¹⁷ Of 283 respondents, 45% believed that the DEM helped them to improve the diagnosis of their projects; the remaining 55% was divided among those who were not sure and the ones who did not believe the DEM was useful in this aspect. The majority of the respondents (53%) reported that the DEM helped to identify indicators, while 43% recognized that the DEM helped to improve the monitoring and evaluation plan.

Figure 1. Bank Staff Perceptions of the DEM



- 2.7. One final issue to be noted is that even SG projects with high DEM scores may ultimately not be evaluable if the right incentives are not in place to follow up on the monitoring and evaluation needs after approval. In the report submitted to the Board, OVE presented the review of the Loan Results Reports of projects analyzed in the evaluability review 2009. The main finding was that among the very few projects for which a Loan Results Report was prepared according to schedule, staff did not follow up on the evaluability issues identified by OVE. Similarly, the loan contracts of the projects for which the DEM was validated

¹⁷ The survey was sent to all VPS and VPC staff, of whom 409 (69%) answered the question “Have you ever prepared a DEM?” A follow-up question on their perceptions about the DEM was given to the respondents who answered YES to this first question.

demonstrated little commitment to the evaluation plans in the Loan Proposal, indicating that even high evaluability might not be acted on after approval.

- 2.8. Management agreed with most of OVE's recommendations and has already started to implement them—specifically, revising the DEM tutorial to make the Project Diagnosis criteria less subjective and the criteria on internal and external validity clearer; keeping the identification of impacts in the Results Matrix but removing the criteria on impact from the evaluability dimensions; keeping the Risk Management information in the DEM but removing the score from the evaluability dimensions; and revising and refining the criteria on PBL/PBPs. OVE has not yet seen these changes in the DEM. They will be visible in projects sent for approval in 2013.

2. The NSG DEM

- 2.9. When analyzing the NSG DEM, considering the standards agreed upon in the Review of the SG and NSG DEM in 2011, OVE concluded that it is not an appropriate tool for measuring evaluability.¹⁸ In contrast to the SG DEM, the NSG DEM focuses on the “evaluability” of each final outcome indicator—that is, it requires an assessment of each indicator separately and independently. It does not provide a broad view of project evaluability, as it does not include an overall diagnosis of the problem to be addressed, an analysis of the context in which the project takes place, or a discussion of the project's vertical logic. It does not provide a complete analysis of monitoring and evaluation plans, and it limits teams in the choice of indicators. RE-397-1 presented a list of evaluability issues identified in the projects analyzed¹⁹ that were not reflected in their evaluability scores. Each of these issues is discussed below.

- ***“A first weakness in NSG project evaluability concerns problem diagnosis.”***
The Loan Document for NSG projects is supposed to present the justification for the Bank's intervention based on an analysis of the market context of the project, the market failures to be addressed, and the means to address them.²⁰ NSG operations are expected to produce externalities with positive impacts on

¹⁸ The analysis presented in the report took into consideration the logic of private sector operations and the constraints they face, which affects their evaluability. First, many project teams indicated that information that would be relevant for evaluability is often excluded from project documents because of confidentiality issues. Second, private clients often lack incentives to pay for monitoring and evaluation. If a project provides a learning opportunity, resources from the Bank may be needed to finance an evaluation. Finally, it is difficult to identify final beneficiaries of financial intermediary operations because of the fungibility of money and the size of the loans relative to client portfolios. This leads teams to focus on the financial transaction or on monitoring a client's overall performance rather than the expected outcomes specified in the Loan Document.

¹⁹ The projects analyzed do not constitute a representative sample of the NSG projects approved in the year; they were just used to exemplify the evaluability problems that were not properly captured in the evaluability tool used for such operations.

²⁰ Guidelines for SCF Staff (Draft) - Using the Development Effectiveness Matrix, version 5.0, February 2011.

identified development problems; therefore, the intervention can be fully justified only if the Loan Document provides a clear explanation of the problem, presenting evidence and describing its causes and dimension.²¹ The NSG loan proposals analyzed by OVE did not provide clear diagnoses of the market failures to be addressed and their origins, or explain how the proposed solution is an appropriate means to address them.

- ***“A second issue is the identification of project objectives.”*** The objectives of most of the reviewed projects were not stated in a specific or consistent way in the Loan Document. This compromises evaluability, since it is unclear what development outcome is to be evaluated at the end.
- ***“A third issue identified in OVE’s analysis is that, as a result of poor diagnoses and lack of specificity in the definition of objectives, the vertical logic of the projects analyzed was not fully developed and justified.”*** The projects reviewed by OVE in RE-397-1 did not establish a clear link between the activities that would be performed and the expected development results of the interventions. DEM guidelines state that Loan Documents should present a table linking outputs to outcomes; however, none of the projects contained such a table. The documents also failed to provide evidence that the proposed interventions could produce the externalities—that is, development outcomes—that the Loan Documents claimed.
- ***“A fourth issue in project evaluability concerns indicators, which might result in part from the format of the DEM, which limits the number of indicators teams can use.”*** Although the DEM worksheet focuses on indicators, OVE found serious problems related to the definition and identification of indicators. The analyzed projects contained a good overall assessment of business performance, but lacked proper indicators to measure development impact.²² Quantitative indicators are required only in the Project Business Performance and Contributions to Economic Development sections, while quantitative indicators and qualitative benchmarks are accepted in the Environmental/Social Risk & Mitigation and Contribution to Private Sector Development sections.²³ There is a clear need to further strengthen the

²¹ A project that will have benefits only for the borrower can be perfectly evaluable, even if irrelevant to the Bank’s development mandate.

²² Despite the fact that monitoring of financial intermediaries’ return on equity, return on assets, and nonperforming loans is essential—and important for the Bank to measure credit risk—the indicators that constitute the Project Business Performance section of the DEM do not help to measure the projects’ contributions to economic development.

²³ The section of the DEM on Private Sector Development requires indicators/results for a list of areas: competition, market expansion, private ownership and transition impact, technology and know-how transfer, demonstration effect from innovation, standards for corporate governance, legal and regulatory framework, and physical or financial market infrastructure. Yet in the projects analyzed by OVE, most of the subsections had no indicators or adequate qualitative benchmarks to assess whether the expected results had been achieved.

indicators to assess private sector development in the DEM and to provide more guidance to teams in identifying indicators and measuring performance against them. In addition, in the first two sections of the DEM—Business Performance and Development Outcome—the projects used no more than four indicators.²⁴ The DEM appears to create unnecessary constraints that compromise rather than facilitate evaluability.

- ***“A fifth issue affecting NSG project evaluability is the DEM scoring system, which leads projects to be labeled as highly evaluable despite all the evaluability issues noted above.”*** In practice, the application of the NSG DEM generates negative incentives for project teams, as the highest evaluability score is given when a criterion does not apply to the project, inflating the final evaluability score of the project. The categories analyzed in the Private Sector Development Section of the DEM were defined following the list in the ECG-GPS.²⁵ However, the GPS notes that this is an indicative list of areas that contribute to private sector development; not all projects are expected to consider all of them. In fact, OVE observed that teams have argued that some of these dimensions do not apply to their project, even though the project receives an evaluability score of 10 in that area. In fact, a score of 10 is always assigned when an area is not applicable, even though this rule is not made explicit in the NSG DEM Guidelines. This practice has inflated evaluability scores: almost all projects score above 9.

- 2.10. A final issue observed in projects but not addressed in the DEM is the failure to assess the adequacy of monitoring and evaluation. The guidelines describe the standard monitoring and evaluation procedures for all projects.²⁶ If the project has evaluation activities that go beyond these standard procedures (e.g., in-depth ex-post economic analyses or impact evaluation), the Loan Document should specify them, along with a description of how the Bank will support the client in monitoring the development results of the project. Yet the DEM does not assess the appropriateness of the evaluation plan or the availability of resources. When the project team does not foresee evaluation activities other than the standard practices (annual supervision reports and Expanded Project Supervision Reports,

²⁴ Additional indicators would affect how the score is computed, and, according to interviews with the teams responsible for the projects considered in OVE’s analysis, teams have been advised not to add indicators.

²⁵ “In assigning ratings for [private sector development], the following factors *may* be considered: Competition...; market expansion...; private ownership and entrepreneurship...; frameworks for markets...; transfer and dispersion of skills...; demonstration effects...; standards for corporate governance and business conduct...; development of financial institutions and financial/capital markets...; and development of physical infrastructure...”(ECG-GPS for private sector, 3rd edition, emphasis added).

²⁶ Monitoring activities: annual project supervision exercises based on the DEM, with preparation of Project Supervision Reports. Evaluation activities: self-evaluation exercises through the preparation of the Expanded Project Supervision Reports (XPSRs) when projects reach early operating maturity. XPSRs are independently validated by OVE.

or XPSRs), the NSG DEM guidelines state that they need not lay out a detailed monitoring and evaluation plan, but should include a description of how the indicators included in the project will be monitored throughout the life of the project. This provision is insufficient. First, the XPSR does not require the provision of evidence of results attributable to the intervention. This means that a complete XPSR is necessary, but not sufficient, to show a causal relation between the project and the development impacts. Second, the evidence from OVE's XPSR validation exercises shows that when projects reach early operating maturity, the information necessary for a complete assessment of project performance is often unavailable. Even though none of the projects for which DEMs were prepared has been validated, the analysis of the case studies suggests that the DEM will not be able to resolve the problem highlighted by OVE's XPSR validation team.

- 2.11. The conclusion is that although improvements to the SG DEM have consolidated its ability to assess evaluability, the NSG DEM needs to be fully revised. OVE had only a few recommendations to strengthen the SG DEM, while it recommended fully revising the NSG tool. In particular, OVE recommended that the NSG tool be harmonized with the SG, since the concept and the standards of evaluability should be the same regardless of the nature of the project—public or private sector.
- 2.12. SCF and OMJ teams are working on a new version, not yet implemented or evaluated by OVE. This new version closely follows the SG DEM, with some changes, including in some of the criteria and the weight distribution. SPD and OVE have been consulted throughout the process on the design of this new NSG DEM. The proposal for this new version is expected to be submitted to the Board of Executive Directors for approval at the end of the current year.²⁷

D. The role of the Operations Policy Committee

- 2.13. One of the IDB-9 requirements about evaluability is to “further [strengthen the role of] the Operations Policy Committee (OPC) and the programming process, by the President of the Bank and Senior Management, to ensure that projects meet minimum evaluability thresholds.” To assess whether and to what extent this requirement was implemented, the team analyzed the minutes of OPC meetings; in particular, the minutes of meetings that took place before the IDB-9 Agreement were compared with the minutes of meetings that took place after it. The OPC

²⁷ The proposal for the new NSG DEM should have been discussed in the Board by the end of September. However, the Board has requested that, in addition to the evaluability section of the DEM, Management harmonize the tool (assessing the likely success of the project according to ECG criteria, currently used by SCF and OMJ) with the one used by IIC (DIAS), which has delayed the finalization of the new DEM.

met 41 times between the beginning of 2008 and the date of the IDB-9 Agreement, and 54 times since the Agreement.²⁸

- 2.14. There appears to have been some change in the focus on evaluability. While before the IDB-9 Agreement, evaluability or the need for proper evaluations was mentioned in less than 18% of the meetings, after March 2010 this share increased to more than 35%. In fact, the analysis of the minutes suggests that in more than 28% of the meetings there has been in-depth discussion of evaluability. This is a clear reflection of the higher importance given to evaluability and the need to report on results in the Bank.
- 2.15. In addition, the content of the discussion has changed since the Agreement. In 2008, before the implementation of the DEM, all the discussion was around the DEF and the use of a tool to assess the evaluability of different Bank products. After 2009, with the DEM already in use and with OVE's third evaluability assessment under way, the discussions became more frequent. But it is only after 2010 that a concern with projects' evaluability can be observed in the meetings. For example, the January 2011 discussion of PE-X1006, Contingent Loan for Natural Disaster Emergencies, actually led to the recommendation for "the team to review the evaluability aspects of this operation with the SPD team to ensure that th[e] integrated approach [proposed in the loan] was fully captured in the development effectiveness matrix." Management's procedure of ensuring that SG projects with evaluability scores below 5 would need to be improved before being submitted for approval, which became effective in 2011, probably explains the reduction in the number of discussions on evaluability during the past year.
- 2.16. Although OPC meetings have given more attention to evaluability, OVE's evaluation identified some private sector projects with serious evaluability issues, none of which was discussed in the meetings. The cause might be their misleading high evaluability scores. In any case, this shows that more attention could be given to project evaluability.

E. The Progress Monitoring Report

- 2.17. Implementation of the PMR system for SG projects, intended to improve the Bank's capacity to monitor its portfolio in execution, began in the second half of 2009 and has already had six cycles. Staff are expected to produce two PMRs per year for each project in execution: one in March, reporting on the achievements of the entire previous year, and one in September, reflecting the achievements of the first half of the year. According to AB-2764, Annex A, the PMR "will incorporate the results framework, the risk mitigation matrices, the Environmental Review Report, the Loan Results Review and the DEM in order to enable project teams to monitor the outputs and outcomes of the project, and their delivery in terms of both cost and time. Another new aspect to improve the Bank's monitoring process

²⁸

This analysis was carried out in October 2012, and the last meeting of the OPC was on July 18, 2012.

is the new methodology to classify projects on ‘Alert’ and ‘Problem’ included in the PMR.” (Table 2 summarizes the content of the PMR according to the IDB-9 Agreement).

- 2.18. The PMR system was rolled out largely within the promised deadlines. When the DEF was written, it was envisioned that the new self-evaluation system would be implemented over the subsequent 36 months. By June 30, 2009, all projects approved under the New Project Cycle were expected to monitor their execution using the new PMR; the other SG projects in the portfolio would be phased into the PMR within 24 months. According to the 2008/2009 DEO, the PMR system was implemented in September 2009, already including all SG projects. Although the PMR system is still under revision, its essence has been the same since implementation began, including the use of a methodology to classify project performance.

Table 2. PMR content: Requirements and implementation

Features listed in the IDB-9 Agreement	Implementation
Results framework	Teams should flag indicators that are part of the results framework.
Risk Mitigation Matrix	The system automatically annexes the Risk Matrix to all PMRs.
Environmental Review Report	The project’s environmental and social impact category is included in the Basic Data section of the PMR .
Loan Results Review	The Loan Results Review is produced 18 months after the first disbursement and has the PMR as one of its annexes.
Information on the DEM	There are four criteria in the DEM related to the PMR.
New methodology to classify projects on “Alert” and “Problem”	Using information about project delivery and costs, a Performance Index is computed for each project, comparing actual with planned values. Depending on the value of the Performance Index, a project might be classified as “Satisfactory,” “Alert,” or “Problem.” The Performance Index is included in the PMR.

Note: This table was prepared on the basis of the following paragraphs of the IDB-9 Report:

- Para 4.7: “**Progress Monitoring Report (PMR)** will incorporate the results framework, the risk mitigation matrices, the Environmental Review Report, the Loan Results Review and the DEM in order to enable project teams to monitor the outputs and outcomes of the project, and their delivery in terms of both cost and time. Another new aspect to improve the Bank’s monitoring process is the new methodology to classify projects on “Alert” and “Problem” included in the PMR.”

- 2.19. The PMR has a very simple premise: once a project is designed with a Results Matrix, including indicators with metrics, baselines, milestones, and targets, it should be easy to monitor its progress vis-à-vis the initial plan. Teams are expected to follow up on the indicators defined in the Results Matrix at approval, and on their costs. “Although information on outcomes and outputs is already collected in a project’s results matrix, project monitoring was not properly linked to the results matrix. Monitoring outcomes and outputs requires a gradual change,

which will provide the Bank with information to quantify results and to enable capturing of lessons.”²⁹ Once actual values are available, they are to be contrasted with the planned ones, and this becomes an assessment of the implementation progress of the project. The Performance Index, computed on the basis of the difference between planned and actual delivery of outputs and their costs, is reported for each project.³⁰

- 2.20. In addition to the follow-up of the Results Matrix, the PMR has sections on Issues, Risk, and Lessons Learned. In the Issues section, “the qualitative section of the PMR,”³¹ project teams should discuss the problems they face during project execution. The section is divided into four subsections: outcome, output, changes to the matrix, and general issues; the PMR guidelines provide guidance on how to use each of the subsections. For this report only the Issues section was analyzed. The Risk section is fed from the OPMAS, which goes beyond the scope of this evaluation, and the Lessons Learned section has been evaluated by the Knowledge and Learning Unit (KNL).³²
- 2.21. A desk review of a sample of PMRs and interviews with Bank staff have yielded information about the implementation of the PMR. Since the system was launched, 771 projects have had at least one PMR. Of these projects, most had been designed before the implementation of the PMR system and thus many were not ready to properly use the system for reporting. From the projects that had less than 10% disbursed as of December 2009 and all projects approved since then, OVE randomly selected a sample of 81 (20% of the group) and reviewed the last three cycles of the PMR—March 2011, September 2011, and March 2012.³³ OVE also interviewed a selected group of Bank staff, including SPD staff and consultants, PMR authors, and the PMR focal point in some units. OVE also sent out an online survey. The results are discussed below.
- 2.22. The first observation is that the overall quality of the reporting in the PMRs has been increasing. In particular, the reporting on the outputs, including the

²⁹ DEO 2008/2009.

³⁰ The Performance Index is computed as the product of multiplying the Schedule Performance Index and the Cost Performance Index. The first one gives the percentage of physical output delivery vis-à-vis the original plan, and the second is equivalent but takes into account the information about actual costs and original plan for output costs. More information on the Performance Index can be found in the PMR Concept Document.

³¹ Progress Monitoring Report Guide.

³² KNL has done one evaluation of the lessons learned section of the PMR in 2011 and is currently conducting another. Among the main findings of this evaluation is the fact that teams are not filling in this section of the PMR, so that lessons learned throughout project execution are not being captured in the system.

³³ The PMR system was changed after the first three cycles. To avoid confusion, OVE opted to review only the last cycles, when most of the changes were already consolidated. Some of the projects in the sample did not have the March 2011 cycle, or even the September 2011 cycle. Of the 81 projects in the sample, 62 had the first cycle, 62 the second cycle, and 81 the third cycle.

definition of the indicators, has improved over the cycles. OVE developed a quality index,³⁴ which ranges from 0 to 100, and the average increased over time from 84.3 to 85.3. Although this is not a big difference, the index does not capture certain aspects of the reports' quality,³⁵ which could be observed during the assessment of the sample of PMRs. In addition, when comparing the PMRs in the latest cycle of projects approved before and after 2010, the difference in the quality index is higher (83.8 for projects approved before 2010 versus 86.5 for projects approved during 2010 and after). This might be related to staff's being better prepared to input the information into the system³⁶ and feeling more comfortable with the new system after a period of transition.

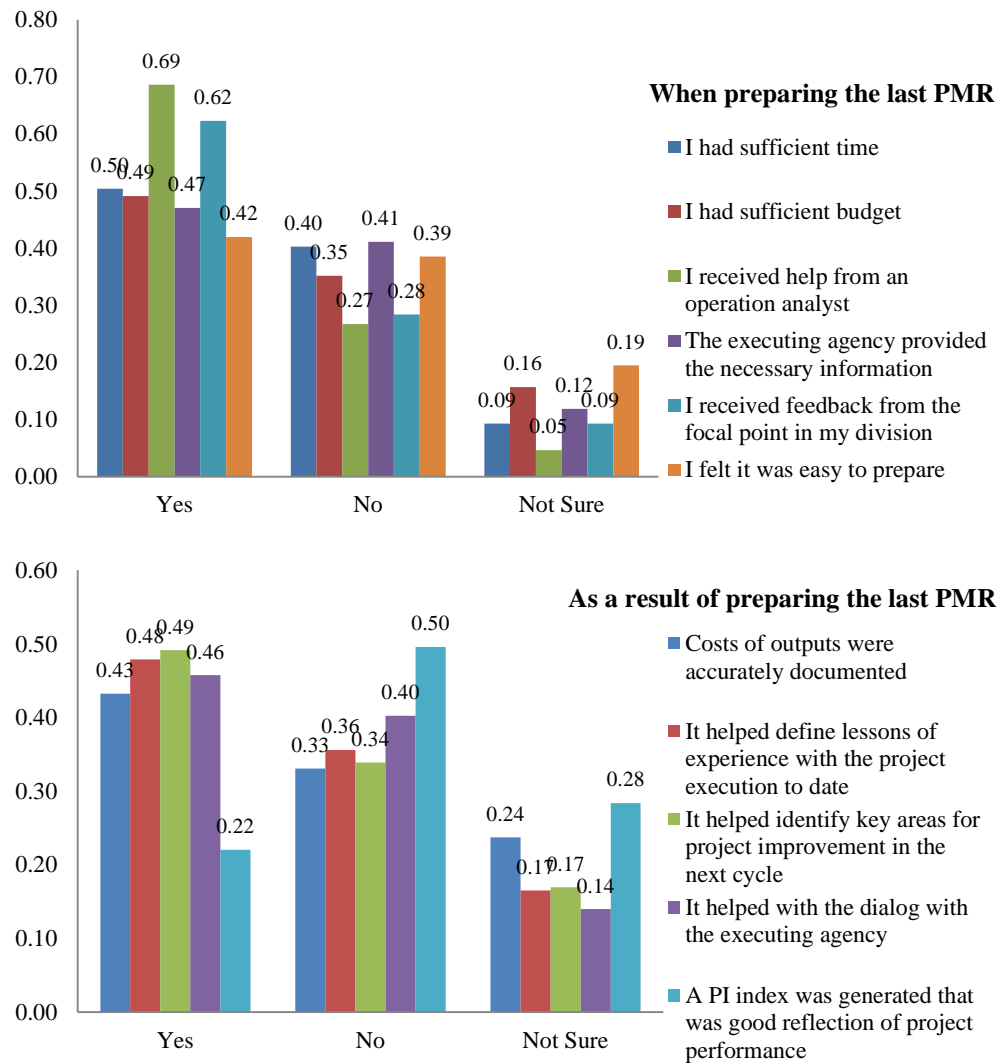
- 2.23. In fact, staff revealed a generally positive perception about the PMR system. The online survey asked staff to compare the old PPMR system with the new PMR system. Among the 236 respondents, 45% were not sure; but of those who expressed an opinion, the majority (62%) preferred the PMR as a tool for tracking project performance. Staff have also been questioned about their general perceptions of the PMR and, as Figure 2 shows, overall more respondents have a positive view of the system, except for the use of the Performance Index. The majority identified issues with the Performance Index, as they do not see it as a proper reflection of their project's performance.

³⁴ The Quality Index is the simple average of the score of 10 different questions that capture the quality of the PMR report in two main aspects. The first set of questions document whether the PMRs include mandatory information, such as the outcomes defined in the project Results Matrix or the actual values of the outcome indicators. The second set of questions aims to assess the quality of the different elements that compose the PMR, including the outcomes, outcome indicators, and outputs. The information collected in the second set of questions helps judge whether the units of measure used are appropriate for the outcomes and outputs stated, whether the indicators contribute to measure the outcome stated, or whether the outcome indicators and outputs are measurable. The exact methodology used is available upon request.

³⁵ The methodology used to compute the quality index has limitations that explain why the average index for the 81 PMRs evaluated is relatively high and why there is not a significant variation across cycles. The first limitation is that the standardized set of questions is focused only on issues that are applicable for the universe of PMRs evaluated and may not represent the full set of issues for which a given PMR reports. For example, it does not control by the project sector, the date of approval, or the percentage of disbursement. These elements may have a significant impact on the definition of the outcomes and outputs, but also on the actual and planned values reported. The second limitation is that the standardized set of questions cannot capture the effort of the teams in reporting issues related to the execution of the project. This certainly diminishes the power of the index to capture changes across cycles, as the project information defined in the project Results Matrix during project design may not change over time, but additional information regarding problems in the execution of the project would increase the quality of the PMR.

³⁶ The fact that numerous training sections have been offered since the implementation of the PMR system is also likely to affect how teams prepare projects and how ready they are to input the information in the system. In addition, in 2009 the DEM was introduced, and projects approved after that were judged on the quality of their monitoring and evaluation systems.

Figure 2. Bank Staff Perception of the PMR



- 2.24. The use of a Performance Index as an assessment of project performance during execution is not new; many other organizations and governments have used such an index. The problems identified in interviews with Bank staff match those commonly discussed in the literature. The Performance Index is based on the earned value method, “a tool that supports the management of project scope, time, and cost.”³⁷ Most users of the method agree that it helps to improve cost, schedule, and technical performance.³⁸ However, effective use of the Performance

³⁷ Anbari, F. T. (2003). “Earned Value Project Management Method and Extensions.” *Project Management Journal*, December 2003.

³⁸ Fleming, Q. W., and Koppelman, J.M. (2000). *Earned Value Project Management* (2nd ed.). Newton Square, PA: Project Management Institute.

Index as a management tool requires that certain conditions be observed. The main problems³⁹ Bank staff noted are as follows.

- ***Users do not have a proper understanding of what the Performance Index is.*** SPD has offered several training sections on the index and its importance, but although many staff have grasped the general concept, many also find that how it is computed remains unclear. This might help explain Bank staff's resistance to accepting the index. In addition, many staff said in interviews that as the index only contrasts the actual delivery of outputs with the original plan, it can only be used as a measure of how badly outputs have been planned.
- ***According to many interviewees, the Performance Index fails to explain the project.*** Project teams do not believe that the Performance Index is a good reflection of their project's performance because they believe they have more information about the project's execution than what is in the index. As one of the online survey respondents said, "Indicators are TOO numerical, inflexible. They help describe a small part of project results, and leave out many others." The general perception is that the Performance Index is one index and it cannot completely capture the essence of what is happening with the project. It is true, however, that the index can only be useful if it is associated with an explanation of the causes of the problem—which project teams do not usually include in the PMRs. In most of the PMRs analyzed by OVE, the issues leading to a low Performance Index are not explained, or are unclear.
- ***When designing projects, staff face difficulty in properly planning the delivery of outputs and their cost, which is essential for the Performance Index.*** Interviewees have pointed out that one reason Bank staff disregard the Performance Index is that a tool that takes into account the original plan will not be a good reflection of project's performance. Many times staff are dealing with high uncertainty, complex projects, and the pressure to prepare projects in a very short period of time. It is important to point out that the New Project Cycle shortened the period of project preparation, allowing staff to complete any missing analysis within 18 months after first disbursement, through the Loan Results Report. It has also been mentioned in the interviews that staff might not be prepared to plan properly, as there is not much training on how to do so. As the literature about the earned value method points out, "The key here is having a quality control process in place, which should include review of project plan deliverables such as the schedule and estimate by experienced project personnel. Everybody makes mistakes, so it's

³⁹

Only problems mentioned or discussed in the majority of the interviews conducted by OVE are listed in this report.

important to have a process in place to catch as many mistakes as possible before the project plan is finalized.”⁴⁰

- ***Staff feel pressure to report only on “good news,” which might lead to inadequate incentives when preparing the PMR.*** Not only in the context of the PMR, but in general in the organization, there is a belief that news about problems or failures is not well received. As an example, early this year the Vice-Presidency for Sectors checked on all projects in execution for which the Performance Index fell in the “alert” and “problem” ranges, and required project teams to explain why their projects had been classified in this way. However, what has been ignored is that even some projects with high Performance Index scores might also have problems, as the index also reflects the data that are inputted into the system. In fact, the correlation between the Performance Index and the quality index is extremely low (13% in the latest cycle).⁴¹ This creates an incentive to report only positive results. A more comprehensive look at the portfolio needs to be used to better inform Senior Management.
- 2.25. Problems other than the Performance Index have also been identified. Many users view the PMR as a tool for “bean counting”—a tool for reporting on, rather than managing, the project: “El PMR no es una herramienta de planificación ni de gestión, es únicamente un registro.” Staff who do not perceive the tool as being useful for them tend to be less motivated while preparing the PMR.
- 2.26. Another issue relates to the lack of integration among the many tools used in the Bank. For instance, there is no apparent connection between the supervision plan and what is being reported in the PMR. In theory both instruments should address the same goal, but they are not connected. Interviewees frequently mentioned that the format of the PMR is very different from that of the other DEF tools, and in many cases staff do not see the relation between it and the Annual Operations Plan or other tools: “The format is not consistent with the Loan Document/Results Matrix,” “The information provided [by executing agencies] does not come in the same format as the PMR requires it.” Another respondent commented, “El PMR debería ser un tema a ser abordado con las contrapartes desde la misión de identificación y el diseño de la Matriz de Resultados”.
- 2.27. The PMR can only be as good as the data that are used in the report, which depends on the capacity and interest of the executing agencies. According to the interviews, the PMR may not be important to the executing agency, and the executor does not always have the incentive to properly report the data necessary

⁴⁰ Lukas, J.A. (2008). “Earned Value Analysis – Why it doesn’t Work.” AACE International 2008 Transactions.

⁴¹ The low correlation indicates that poorly prepared PMRs might also have a high Performance Index. However, the Performance Index is less informative in a project for which the quality of the data used is low.

for the PMR. Another respondent pointed out, “Porque el PMR se nutre de la Matriz [de Resultados] y la Matriz satisface las necesidades del Banco, pero no las del ejecutor, el ejecutor no necesariamente puede proporcionar la información de la Matriz de Resultados y entonces quedan huecos de información”. This has led some to conclude that PMRs should be done by the executing agencies. But whoever prepares the PMR, the issue is that it depends on the capacity of the executor and on the agreements made in the loan contract about the role of the executing agency in monitoring indicators. If the Bank is to have good data, it needs to ensure that the executing agency has the capacity to collect the necessary data and that, before execution begins, there is a clear agreement on what needs to be monitored.

- 2.28. In addition to the problems with data collection, OVE found improper use of the Issues section in many of the PMRs it analyzed. The Issues section is very often not used. For example, although there is specific guidance to report on changes to the Results Matrix in the Issues section, in 63% percent of the cases in which the outputs reported in the PMR do not come from the Results Matrix, there is no explanation for this in the Issues section.⁴² The analysis of outcome indicators leads to an even higher proportion: in 84% percent of the projects analyzed, in the latest PMR for which the outcome did not come from the Results Matrix, there was no explanation in the Issues section.⁴³ Only about 54% had some information in the outcome subsection of the Issues section, and OVE found that in about 10% of the cases the information provided was not clear.⁴⁴ Similar results were found with respect to the output and changes to the matrix subsections. Interviews with staff suggest that this is a result of the fact that the most “visible” part of the PMR seems to be the Performance Index. Therefore, teams preparing the PMR focus on the information to generate the index, neglecting the importance of reporting on the issues identified.
- 2.29. If the Bank is to get the most out of the PMR system, it needs to adjust some aspects of the system, and the Bank itself has to undergo some changes. Staff should have appropriate incentives to use the Issues section and, most importantly, to be very consistent in reporting all changes to the Results Matrix. This would be an essential piece of information for the preparation of PCRs. There is also a need to balance the Performance Index with other indicators of project progress, so that the analysis of project execution would not rely on a single indicator and there would be less incentive to attempt to manipulate it. In addition, the Bank needs to develop a culture of proper planning, accountability for planning, and willingness to report on problems and failures.

⁴² Among all PMRs analyzed, approximately 43% of the outputs reported do not come from the Results Matrix.

⁴³ Among all PMRs analyzed, approximately 33.5% of the outcomes reported do not come from the Results Matrix.

⁴⁴ Clarity was judged by whether or not the evaluator would need more information to understand the issue described.

F. The Project Completion Report

- 2.30. The last piece of the architecture of the DEF for SG operations, a system to report on project performance at completion, is not yet finalized. The DEF had in its design the concept of an Expanded Progress Monitoring Report (XPMR), which would use the information of the PMR system and would include a discussion on the achievement of final outcomes. As AB-2764 described in Annex A, the “XPMR compares actual results achieved with the expected results presented at entry in the DEM, which includes the recalculation of the DEM score.” The XPMR was intended to report on (i) time and cost of outputs achieved from the last PMR; (ii) results achieved and evaluation methods used; and (iii) the Bank’s performance during the life of the project. The XPMR system was supposed to be launched the end of 2011, according to GN-2489. Instead, Management has decided to keep the PCR system after revising it to increase the quality of the reports produced and to make sure teams are using the information generated by the PMR system. The new system is expected to take effect in 2013.
- 2.31. The PCR should be the main instrument in IDB’s self-evaluation system for reporting on the effectiveness of completed operations.⁴⁵ The PCR documents progress in implementing the project, the achievements of the project against its stated objectives, the sustainability of results, and the performance of the Bank and the borrower/executing agency. The PCR should serve two purposes: an accountability tool for the Bank’s shareholders and clients, and a means of learning from experience to inform the design of future operations. Currently, PCRs are governed by guidelines approved in mid-2006.⁴⁶ This section draws on findings in Review of the Project Completion Reporting System for Sovereign Guarantee Operations, RE-417, to discuss the main challenges faced by the current system and the problems the DEO faces when reporting on the effectiveness of completed projects.
- 2.32. As OVE reported in its last assessment of the PCR system, a desk review of PCRs approved in 2008 and 2009 found that the overall quality of such reports is very low. According to the PCR template, PCRs are required to compare planned results with achieved results through the analysis of the outcome indicators. However, the evaluation found that the evidence provided is mostly inadequate to document projects’ achievement of their development objectives. Only a few PCRs contained acceptable analysis of results. OVE recognizes that the severe

⁴⁵ For SG operations, PCRs are required for projects supported by IDB loans of \$150,000 or more, with execution periods of at least 12 months. Certain nonreimbursable technical cooperation activities also require a PCR.

⁴⁶ IDB (2006), *Project Completion Report*. Version 1.1, July 25; IDB (2006) *Project Completion Report: Policy-Based Loan*, June; and IDB (2006) *Quick Reference Guide for Preparing and Processing the Project Completion Report*, August. PCRs also are governed by OP-305, *Ex-Post Evaluation Operations*, October 2003, and CO-309, *Project Completion Report Manual*, May 2006.

data limitations identified are likely to be related to deficiencies in project design.⁴⁷

- 2.33. OVE's evaluation also pointed out a number of shortcomings in the current PCR system, which might help to explain the low quality of the reports and their lack of data. The report submitted in June 2011 by the Independent Review Panel studying IDB's evaluation system⁴⁸ pointed out negative perceptions of the PCR system: "While a large majority of respondents view these reports as relevant to their work..., the evidence in support of their usefulness and influence was weaker. This is consistent with the feedback received during the focus groups that indicated that PCRs received little attention or importance." By the same token, in interviews conducted by OVE in 2011, Bank staff indicated that the existing PCR guidelines provide insufficient direction on evaluation criteria and format, and staff and consultants have received too little training on how to write PCRs. As a result, PCR authors have often sought help from colleagues or invented their own solutions, resulting in reports that are very heterogeneous and in many cases of poor quality.
- 2.34. Another issue frequently pointed out by staff during interviews is the insufficient funding to produce PCRs.⁴⁹ They are usually financed with the supervision resources remaining at the time of project closure. As respondents mentioned in the staff survey, "The funding for the PCR comes out of whatever is left in the supervision budget. It is almost assumed that no money is needed for the PCR." "People don't have time for PCRs. The time allocated to preparing projects has been reduced and the number of projects approved has increased, with the same resources." On average, according to the results of the online staff survey, PCR authors spent less than 14 days preparing their most recent PCR.
- 2.35. However, despite the general perception that resources are insufficient, it is not possible to determine how much PCRs actually cost. Thus it is also impossible to measure how much more needs to be allocated for PCR production. The costs related to PCRs are not entered as a line item in the division's budget, as they are

⁴⁷ OVE reviewed the quality of the results frameworks at approval of 5 of the 32 projects in the sample. All projects lacked well-defined outcome indicators. The introduction of the DEM and the focus given to evaluation after project approval, as discussed in RE-397-1, have the potential to improve the data environment in which PCRs will be produced in the future.

⁴⁸ In August 2010 the Board commissioned an external review of IDB's evaluation function to assess its quality and recommend ways to strengthen the role of evaluation in achieving development effectiveness. See IDB (2011), *Strengthening Evaluation to Improve Development Results: Report of the Independent Review Panel on Evaluation at the Inter-American Development Bank*, Washington, DC.

⁴⁹ Interviews with staff suggested that the average cost is between \$10,000 and \$20,000. Data from the staff survey are less indicative: the reported cost of a PCR consultant ranged from \$2,500 to \$80,000 (the latter was for a firm), along with between \$500 and \$5,000 for consultant travel and a similar amount for staff travel. Respondents spent an average of two to three weeks preparing the PCR. The amount spent for additional data collection ranged from \$0 to \$5,000, with the typical amount being between \$0 and \$1,000.

financed by the remaining supervision budget. Since some of these resources are used for other post-project activities (e.g., workshops), they are not a good proxy for the cost of the PCR. Another difficulty in estimating the actual cost of PCRs comes from the fact that PCRs for investment projects can begin to be produced even before the project closes.

- 2.36. Management's 2012 budget proposal⁵⁰ calls for a "gradual incorporation of quality enhancement requirements in operational cost parameters." It requests additional resources to ensure adequate monitoring of the DEF and its component tools.⁵¹ However, if costs related to PCRs are not a separate line in the budget, how much is actually spent and how much increase is needed to achieve better quality reports will remain unclear.
- 2.37. In addition, the system of quality review does not appear to ensure high quality. Of the 79 PCRs approved in 2010, 29 included the Quality and Risk Review (QRR) comments as an annex; in the other cases OVE had no access to the QRR information. The analysis of available QRR information revealed that in about 25% of the cases the QRR was attached but did not generate comments. Moreover, only slightly more than half of the survey respondents said that the comments received at the QRR meeting were useful. Only about a third of respondents said that QRR comments addressed the findings, conclusions, and recommendations in the PCR or important topics that had been overlooked. Problems with data or the empirical basis of conclusions drawn in the PCR were covered only 23% of the time. As a respondent to the staff survey said, "We organize a QRR meeting to review the PCR, but very few people show up. The ones that do have very little to say about the content of the PCR—the discussion is mostly about grammar." Another respondent observed, "Nobody reviews PCRs, so there is little incentive for quality. There isn't much quality control at the level of the Division Chief either, and not much push at the Board for results." In fact, 36% of respondents said the PCR received no division review or review by their manager.
- 2.38. The PCR format compounds the problem. As many staff pointed out in interviews, the PCR template contains separate sections for the discussion of project outputs and project outcomes. This type of organization breaks the presentation of the results chain—the sequence of events linking project activities, outputs, and outcomes—making it difficult to determine whether the intended chain of events did in fact take place and thus whether observed outcomes can be attributed to the project. In addition, the PCR template asks only for data on the project performance indicators defined at entry or modified during execution. This gives the appearance that other relevant evidence may be ignored. Limiting evidence to project-defined performance indicators is particularly problematic if

⁵⁰ IDB (2011), "2012 Program and Budget Proposal" (GA-248-2), October 14.

⁵¹ The amount of resources devoted to monitoring and evaluation in the 2012 budget proposal represents a 26% increase over 2009, but includes expenses for the DEMs and PMRs.

the project's monitoring and evaluation system was poorly designed or implemented. Some of these issues were noted during interviews, and others were reflected in explicit suggestions for changes in the template in the online staff survey.⁵²

- 2.39. A comparison of the PCR template with the ECG-GPS reveals a number of areas in which the PCR guidelines are inconsistent with the GPS⁵³—mainly in the definitions of criteria by which a project is assessed. The GPS calls for four “core” evaluation criteria to be assessed and rated: Relevance, Effectiveness (achievement of objectives), Efficiency, and Sustainability. The PCR methodology calls for three evaluation criteria to be assessed and rated: Achievement of Development Objectives, Implementation Progress, and Sustainability. Of the three, only Sustainability is the same as the GPS core criterion. Development Objectives corresponds roughly to the Effectiveness criterion in the GPS, but, as noted above, the discussion of outcomes is separated from the discussion of outputs, splitting the discussion of the results chain into two parts. Development Objectives also includes some elements of project efficiency. The third PCR criterion, Implementation Progress, does not correspond to a GPS criterion, although it could usefully be included in the effectiveness section of the GPS as part of the discussion of the causal chain.
- 2.40. Data from the staff survey indicate that the Exit Workshop provides some useful information for the PCR, but that more could be gained from these reviews. Of the 78 PCRs for projects that finished disbursing in 2010, 84% included some Exit Workshop material in an annex, although the content of the annex varied: some were the minutes of the meeting, some were copies of the Bank's presentation, and some provided the list of participants. This seems to corroborate the results of the survey: 83% of respondents held an Exit Workshop, and 79% used information from the workshop in PCR preparation. However, in the staff survey a few respondents pointed out problems they faced with the Exit Workshop. For example, “Depending on the project, Exit Workshops tend to happen years after the project has ended. That also means that executing units are not there to provide feedback. The Government has no motivation to participate actively in the workshop. It would be more appropriate to have the Exit Workshop at the same time (or during) the final evaluation.”

⁵² For example, one of the respondents of the staff survey mentioned, “Taking care of the noted deficiencies, and consolidating the information with the indicators and goals included in the project risk, results, and DEM matrixes thus avoiding the need to prepare 4 separate reports that, worse yet, could have diverging or contradictory information and conclusions.”

⁵³ The GPS calls for an assessment of the project's relevance, including both the Relevance of Objectives and the Relevance of Design. Under the current PCR methodology, Relevance is not included as a separate evaluation criterion, although to some extent, it may be discussed in the Development Objectives section. The GPS calls for Efficiency to be assessed and rated as a separate criterion on the basis of both cost-benefit analysis (if feasible and practical) and cost-effectiveness analysis. In the current PCR format, evidence on efficiency—ex-post calculations of economic and financial rates of return and/or cost effectiveness—is buried in the Development Objectives section.

- 2.41. For most projects a final evaluation is planned at approval, but the PCRs contain little information about such evaluations. Of the PCRs approved in 2010, 38% included the final evaluation as an annex; for the remaining 62%, it is unclear whether or not a final evaluation was done. Responses to the staff survey indicated that a final evaluation was produced before the PCR 58% of the time, and 78% of those final evaluations were used as an input to the PCR. This indicates the final evaluation was not always produced before the PCR.
- 2.42. A key underlying problem is that many Bank staff regards the self-evaluation system, including PCRs, as having limited usefulness and as being too inflexible. As a result, according to the Independent Review Panel report, the system is not being used as it should be to establish a culture of learning. This perception of PCRs was confirmed in staff interviews. For instance, only 22% of survey respondents had used a PCR in discussions with country counterparts or stakeholders regarding public policy or project design. However, many staff commented in the survey that the PCR could potentially be useful to the Bank: “The PCR could be a useful tool to generate institutional knowledge regarding the feasibility, challenges, and expected outcomes of different types of projects. Bank employees acquire qualitative information regarding these aspects, and the PCR could be a good vehicle to document this information.” In fact, 46% of survey respondents suggested that staff’s motivation to produce good-quality PCRs and deliver them on time would be greater if PCRs were used more inside the Bank.
- 2.43. The PCR is also not used by counterparts. Interviews with counterparts revealed that they did not remember the contents of the PCR and saw it as another contractual requirement made by the IDB to obtain funds. Of 79 PCRs approved in 2010, 66 included the counterpart PCR, but the contents were brief and did not give evidence or reasons for assigning scores to the Bank. Of the 66 counterpart PCRs, only one (BR0372) was an extended and well-documented borrower PCR.
- 2.44. The absence of a regular and sustained process of OVE validation has further compromised staff incentives to improve the quality of PCRs. OVE did not play a central role in setting the right incentives to produce good-quality PCRs by validating them annually. Since the last revision of the PCR guidelines in 2006, OVE has sent to Management only one cycle of PCR validations, which were reflected in the 2010 DEO; but after that, additional validations were not done or made public.
- 2.45. The limited use of PCRs makes the production of high-quality PCRs difficult. Interview evidence indicates that staff do not feel a strong incentive to produce high-quality and timely PCRs because self-evaluation is not a priority of Management and no further use of the final PCR is approved.⁵⁴ More than 60%

⁵⁴ Of 256 PCRs approved between 2006 and 2010, 33.2% were not found on the Internet, limiting further external use of the PCR.

- of those surveyed felt that staff are not rewarded or recognized for producing high-quality PCRs by their manager, in departmental meetings, or in their performance appraisal: “PCRs are not visible, and the work required to produce a high-quality PCR is not recognized.” During 2006-2010, 60% of PCRs were delivered more than 180 days after the last disbursement date. For PCRs approved in 2008-11, the average delivery date was 250 days after the date of the last disbursement; some were delayed more than a year.
- 2.46. Conflict of interest is another potential issue. According to interviews with PCR authors, when preparing PCRs, some of them hesitate to criticize their colleagues who worked on the design or on part of the execution of the project.⁵⁵ This can result in unjustifiably high ratings, particularly on Development Objectives and Bank Performance. The share of PCRs with Development Objectives ratings that were *probable* or *highly probable* (the upper two of four rating categories) was 80% in 2008, 85% in 2009, and 87% in 2010. PCR authors praised the performance of the Bank even more: in 2010, fully 93% of PCRs rated Bank Performance *highly satisfactory* or *satisfactory*.⁵⁶
- 2.47. Many of the problems with the PCR system are not unique to the IDB. Other international financial institutions and development organizations have had similar experiences with project self-evaluation: lack of incentives to deliver quality reports, inattention by management and staff, and missed opportunities to apply lessons learned from past operations. For example, a study by the International Development Research Centre (IDRC) of its project completion report process noted, “When staff realized that their project completion reports were unlikely ever to be read, they understandably became reluctant to put time and effort into preparing them. As a result ...when these people left the program or the organization, IDRC suffered a ‘knowledge drain.’”⁵⁷
- 2.48. Management and the Board of Directors are aware of the deficiencies of PCRs, and, according to Management’s response to the OVE evaluation that was presented to the Board in August 2012, Management is working on a new version of the PCR that is to be launched early in 2013. Management is also working to strengthen the process of preparing the PCRs and the incentives for staff. Finally, Management has proposed a separate budget line and an increase in resources for PCR preparation: “resources would be specifically allocated for PCR preparation, including \$25,000 for non-personnel costs per PCR.”⁵⁸ Management’s willingness

⁵⁵ The PCR author usually is the person who was responsible for the design and/or implementation of the project, or a consultant who reports to that person. This puts the person responsible for the PCR in the difficult position of having to criticize his or her own work.

⁵⁶ 2008-09 data from DEO 2010; 2010 data from OVE.

⁵⁷ International Development Research Center (2005). *Deepening a Culture of Reflection: IDRC’s Rolling Project Completion Report Process*. November, p. 2.

⁵⁸ Management’s response to the “Review of the Project Completion Reporting System for Sovereign Guarantee Operations,” RE-417-1.

to address the problems of the PCRs may represent an opening to significantly improve the Bank's monitoring and evaluation system and its incentive structure.

G. The Development Effectiveness Overview

- 2.49. The DEO is Management's annual report on IDB's effectiveness. The first DEO was published in 2010, reviewing the data of the 2008/2009 report. While the first DEO had a simpler format and used an annex for most of the discussion on the effectiveness of the operations supported by the Bank in the Region, the latest two DEOs are very much alike in their structure and the nature of their content.
- 2.50. The IDB-9 Agreement had very specific requirements regarding the DEO. In particular, the Cancún Declaration refers to the need to "disclose project-level reporting with stronger metrics," which implied "public disclosure in the Development Effectiveness Overview (DEO) of ex-ante project-level evaluability analysis, compliance with institutional priorities, and ERR calculations for projects approved that year, and ex-post impact evaluations for any projects evaluated in that year, including for NSG projects." This requirement was to take effect immediately, affecting so far the 2010 and 2011 DEOs. Table 3 shows the status of implementation of the IDB-9 requirements related to the DEO, and Annex B of this report shows detailed information on the DEO's content.

Table 3. Implementation of IDB-9 requirements related to the DEO

IDB-9 Requirement	Implementation
Published annually.	The DEO is published each year before the Annual Meetings.
Provide a consolidated report on the achievement of development results of the Bank's interventions in line with the institutional strategy of IDB-9.	The DEO is partially reporting on the achievement of IDB's interventions, as not all interventions have been properly evaluated.
Include analysis of compliance with institutional priorities and include a review of lessons learned in IDB strategic priority areas. It should also report an assessment of development effectiveness in a specific area, with emphasis on policy responses and the role of the Bank. Finally, it should also provide evidence on the effectiveness of the intervention models applied in Bank interventions.	The compliance with institutional priorities is done partially. Each institutional priority is analyzed in light of the projects for which Management has information.
Include analysis of project-level evaluability at approval.	Since its first edition, the DEO includes a discussion on the evaluability of the portfolio approved in the year of reference and an analysis of progress in improving evaluability.
Include analysis of ERR (or if ERR cannot be calculated, use cost-benefit and other calculations for projects approved that year).	The DEO only reports on the share of projects that included such analysis, but does not provide an actual analysis of the computations.
Include analysis on ex-post impact evaluations	This is partially done, as most of the

IDB-9 Requirement	Implementation
for any projects evaluated in that year, including for NSG projects.	discussions on impact evaluations are related to proposed impact evaluations to be completed in the future. The share of projects that include a rigorous impact evaluation has increased over time, but the results are not yet available. (Annex C lists the impact evaluations promised in each DEO).
Consider how to incorporate into the design and implementation of operations those OVE recommendations endorsed by the Board of Executive Directors to improve loan quality.	The only mention of OVE's recommendations is in the 2010 DEO regarding projects related to citizen security. Management is working on a system to track OVE's recommendations, which will facilitate the process of reporting on their implementation status.
Report on progress of all the targets of the Results Framework.	IDB-9 defined lending targets, regional development goals, and the Bank's outputs to contribute to such goals. The DEOs report progress on all of these dimensions. The quality of what is reported is assessed in the background paper on the Corporate Results Framework.

Note:. This table was constructed on the basis of the following parts of the IDB-9 Report:

- Para. 4.6: "The progress in development effectiveness will be reported annually through the Development Effectiveness Overview (DEO), which will be disclosed publicly. Among the analyses that are to be included in every DEO are: project-level evaluability at approval, compliance with institutional priorities, ERR (or if ERR cannot be calculated, cost-benefit and other calculations for projects approved that year will be applied), and ex-post impact evaluations for any projects evaluated in that year, including for NSG projects. Additionally, the DEO will consider how to incorporate into the design and implementation of operations those OVE recommendations endorsed by the Board of Executive Directors to improve loan quality."
- Para. 6.11: "The Development Effectiveness Overview (DEO) to be published annually has as its purpose to provide a consolidated report on the achievement of development results of the Bank's interventions in line with the institutional strategy of IDB-9. The DEO aims to support the work to build a body of research into effective approaches to policies and programs that provide sizable development outcomes. It will include: a review of lessons learned in IDB strategic priority areas, and an assessment of development effectiveness in a specific area, with emphasis on policy responses and the role of the Bank, as well as to provide evidence on the effectiveness of the intervention models applied in Bank interventions."

2.51. A number of weakness were identified in the analysis of the DEO. The most important one is the lack of accountability for the information provided. This problem is partially related to the issues identified with the tools that make up the Bank's self-evaluation framework, particularly the information related to the PCR and the NSG DEM. In addition, the reports do not provide any criteria to justify selecting the projects for which they present evaluation results, and evaluations have been promised and are not followed up in later DEOs. These issues are further discussed below.

2.52. As has been pointed out, the lack of a proper system to report on the achievements of completed projects has undermined Management's ability to report on results

in a reliable way. The DEO has annually reported on the scores of the PCRs. The large majority of the PCRs rate their achievement of development objectives as *satisfactory* or *highly satisfactory*. As OVE shows in its latest review of PCRs, the scores in the PCRs approved in 2008 and 2009 have no correlation with the quality of evidence provided, which makes them unreliable. As the PCR system has not changed since then, and the vast majority of projects completed in 2010 and 2011 did not directly benefit from the DEM and the improvements in the PMR system, it is unlikely that the results of the analysis of the PCR scores would lead to a different conclusion on the last two cohorts of completed projects. Therefore, the section of the DEO on the PCR data can be misleading. This issue has been acknowledged in the DEOs.

- 2.53. Similarly, the improper way of measuring evaluability among private sector projects, and the categories used to label projects, have been misleading. Overall, the portfolio of the Bank as reported in the last DEO is mostly highly evaluable, despite the many evaluability problems identified by OVE in NSG projects. OVE's assessment of the NSG DEM has shown that it is not a proper tool for assessing evaluability, and that the high evaluability scores of these projects do not properly reflect the capacity of these projects to report on the results at completion. Similarly, the SG projects are classified as highly evaluable when their scores are above 7. Some of the projects analyzed in OVE's report have scores above 7 but have shown serious evaluability problems, which suggests that they should not be labeled as highly evaluable. The classification of the SG projects according to their scores needs to be revised to avoid misconception.
- 2.54. Additionally, the choice of evidence to be presented in the discussion of the Bank's strategic sectors is unclear, leaving the impression that it is just a collection of stories rather than a representative sample of results (see Annex B for examples). For instance, several subsectors within the mandated IDB-9 sections are chosen for extensive focus. To illustrate this point, in the thematic area of energy there is a focus on hydropower plants (which represent 3 of the 10 projects approved in 2010). There is no explanation of this choice, and of what the IDB is doing in the other sectors in energy. The 2011 DEO does not report even the number of projects and the financing allocated in each of the priority sectors—information that was presented in the two previous editions of the DEO.
- 2.55. Another issue identified is the lack of follow-up on the promised evaluations of previous DEOs. To compensate for the lack of reliable information in PCRs, the DEO has presented the analysis of projects for which proper impact evaluations have been designed.⁵⁹ The discussion of impact evaluations of Bank-financed operations is mostly focused on the design of future impact evaluations, to be completed in the next couple of years, particularly because the Bank's culture of

⁵⁹

These evaluations do not always refer to projects financed by the Bank; see Management's response to the "Review of the Project Completion Reporting System for Sovereign Guarantee Operations," RE-417-1.

evaluating projects through rigorous methods is young, dating only from the implementation of the DEM. However, the DEO does not include expected deadlines for the impact evaluations to be completed. The impact evaluations announced in 2008-09 were not followed up in 2010 or 2011, and the ones announced in 2010 were not followed up in 2011 (Annex C lists all of them). Currently, there is no visible tracking system to follow up on such evaluations. In the next years Management expects to have a website with a more comprehensive analysis of the progress of the impact evaluations proposed, so that the information provided in the DEO could be tracked.

- 2.56. OVE compared the DEO with the overview reports produced in the World Bank, the Asian Development Bank, and the African Development Bank (Annex D).⁶⁰ This exercise has revealed that the DEO is more complete than most of them in several aspects. In particular, the IDB is the only Bank to report on projects evaluability⁶¹ and on the share of projects with economic analysis. All organizations rely heavily on the results of completion reports (PCRs or equivalent), and only the IDB and the overview report produced by the World Bank (the World Bank for Results, or WB4R) illustrate their analysis with impact evaluations. In the WB4R, as in the DEO, the selection criteria for the impact evaluations mentioned is not clear; instead, they are used as examples, and a more extensive analysis of the World Bank impact evaluations can be found on a website—similarly to what IDB Management is proposing for the coming years.⁶² In none of the overview reports is it possible to identify the recommendations from their independent evaluation offices,⁶³ except for of the overview report produced by the World Bank’s Independent Evaluation Group.
- 2.57. In conclusion, the DEO is not yet fulfilling the role envisioned in IDB-9 of reporting on Bank results. It lacks an analysis of the ERR of projects in the portfolio, except for information on the share of projects that included such a computation. And there are issues with the reporting on evaluability and project results, due in part to weaknesses in some DEF tools. However in some aspects it seems to be reporting better and more extensively than similar reports produced in

⁶⁰ OVE compared the 2011 DEO with the 2011 World Bank for Results (WB4R) produced by the World Bank’s Management, the 2011 Results and Performance produced by the World Bank Independent Evaluation Group, the 2011 Development Effectiveness Review from the Asian Development Bank, and the 2012 Annual Development Effectiveness Review from the African Development Bank.

⁶¹ The other reports provide some analysis of the quality at entry of their portfolio; however, the evaluability analysis prepared by IDB’s self-evaluation system goes beyond to address the question of how well projects can be evaluated at completion.

⁶² The World Bank’s impact evaluation website is managed by the Independent Evaluation Group.

⁶³ The Asian and African Development Banks include in their reports a list of actions defined by Management to respond to the recommendations of several parts of the Bank. However, they do not address all the recommendations from their independent evaluation offices. The Asian Development Bank has a tracking system to follow up on the recommendations of their independent evaluation office, but it is not directly translated into their Development Effectiveness Review.

other organizations; and it could be greatly improved by addressing the weakness identified in this report.

III. SUGGESTIONS GOING FORWARD

- 3.1. Knowledge about what works and what does not work is essential to inform the decisions of the Bank. This understanding has generated many requirements in the IDB-9 and a particularly strong focus on the need to design projects that can be properly evaluated at completion. As this report has discussed, for SG projects Management has taken important steps in this direction. Every project is now rated on evaluability and needs to be followed up in the PMR system. However, it is clear that for the purpose of generating knowledge about the effectiveness of the Bank's operations, it is not enough to have projects that are evaluable or monitored. Projects need to be analyzed at completion, the achievement of results should be contrasted with what was originally expected, and their sustainability should be considered. Management is also committed to finish the implementation of the DEF with a proper system to report on the results of completed projects, which is essential to move forward.
- 3.2. The following suggestions for Management emerge from the analysis in this paper:
 - ***First, ensure that the various tools are integrated with one another.*** It is not sufficient to have perfectly evaluable projects if they are not fully monitored and evaluated. Similarly, it is not possible to use a monitoring tool effectively if results indicators were not properly defined from the beginning. The standards of evaluability required from project teams need to be those that will make them more successful in preparing the PMRs throughout the life of the project and will allow them to have the information to create meaningful PCRs.
 - ***Second, support a culture of planning for results and a willingness to report on problems and failures.*** The DEF tools will be informative only to the extent that project teams do proper planning and report problems. The Performance Index relies heavily on the quality of the data put into the system and the planning of output delivery and cost at design. Similarly, the lessons learned from PCRs will be most useful when teams are willing to report not only on success but also on failure.
 - ***Third, strengthen the evaluation of private sector operations.*** The instruments used to monitor and evaluate NSG operations are different from those used for SG projects. The lack of attention given to the NSG DEF in the IDB-9 indicates that these projects have generally been off the radar screen with regard to results reporting. However, knowledge about what works and

what does not work in NSG operations will be critical as the Bank's engagement in this area grows.

- ***Finally, strengthen the DEO as an accountability tool.*** It is also important that the DEO report accurately on the full range of Bank experiences, both successes and failures. Refining the tools under the DEF will provide better information for the DEO, and including the results of impact evaluations and OVE recommendations will also help in this regard.

LIST OF PERSONS INTERVIEWED

In preparing this report, to gain information on the DEM, the PMR, or the PCR, OVE interviewed a total of 50 Bank staff and consultants, two executing agencies, and selected Executive Directors. In addition, OVE sent out two online surveys. The first one, focused on the PCR and sent to all authors of PCRs completed in the past 5 years, had a response rate of about 35%. The second survey, focused on the DEM and the PMR and sent to all VPS and VPC staff, had a response rate of more than 60%.

Among the people OVE met are the following (in alphabetical order):

<i>Name</i>	<i>Title</i>	<i>Department</i>
Carmen Pages	Unit Chief, Labor Markets	SCL/LMK
Carola Alvarez	Div. Chief, Strategy Development	SPD/SDV
Fernando Regalia	Div. Chief, Social Protection and Health	SCL/SPH
Flora Painter	Div. Chief, Competitiveness and Innovation	IFD/CTI
Graciela Schamis	Sector Manager	KNL/KNL
Hector Malarin	Div. Chief, Environment and Rural Development	INE/RDH
Jaime Enrique Vargas	Operations Senior Advisor	EVP/EVP
Karen Mokate	Div. Chief, Knowledge Management	KNL/KNM
Kurt Focke	Div. Chief, Retirement Plans	HRD/SRP
Morgan Doyle	Sector Senior Advisor	VPS/VPS
Vicente Fretes	Div. Chief, Fiscal and Municipal Management	IFD/FMM

Of the Bank staff and consultants interviewed by OVE, five are in SPD; the others were selected for either being the author of a recently prepared PCR or PMR, or for being a focal point in their units.

**ANNEX A. THE REQUIREMENTS ON THE DEF AND THE DEO IN THE REPORT ON THE
NINTH GENERAL INCREASE IN THE RESOURCES OF THE INTER-AMERICAN
DEVELOPMENT BANK (AB-2764)**

Box 1. Requirements in the IDB-9 related to the DEO and evaluability

“Among the analyses that are to be included in every DEO are: project-level evaluability at approval, compliance with institutional priorities, ERR (or if ERR cannot be calculated, cost-benefit and other calculations for projects approved that year will be applied), and ex-post impact evaluations for any projects evaluated in that year, including for NSG projects. Additionally, the DEO will consider how to incorporate into the design and implementation of operations those OVE recommendations endorsed by the Board of Executive Directors to improve loan quality.”

“Governors endorse a further strengthening of the Operations Policy Committee (OPC) and the programming process, by the President of the Bank and Senior Management, to ensure that projects meet minimum evaluability thresholds. In this respect, Management will amend operational procedures, by end of Q3 of 2010, according to the following criteria: (i) all SG and NSG projects must be rated for evaluability; (ii) the evaluability score includes only the dimensions of evaluability of the DEM; (iii) SPD will support teams in meeting evaluability standards from project profile to project proposal, and will validate the final evaluability score for OPC consideration: RES will review the existing methodologies for scoring evaluability to determine any required improvements; OVE will report annually to the Board of Executive Directors on project evaluability (ex-ante), as well as validate achieved results in completed projects (ex-post); (iv) a minimum evaluability threshold of 5 will be required for all operations to be submitted to the Board of Executive Directors; (v) in cases of extraordinary humanitarian or financial crisis (acute payment pressures), and upon the presentation of a properly justified request by the President, the Board of Executive Directors may waive the required threshold prior to approving an operation, long as the request is accompanied by a timetable giving the shortest possible time period, not to exceed 90 days, for achieving the said threshold.”

Box 2. Description of the DEF in the Report on the Ninth General Increase in the Resources of the Inter-American Development Bank

*“The Bank has **monitoring and management instruments** to ensure results. The Bank has been making improvements in the way projects are monitored and evaluated. The DEF has three instruments for monitoring at entry (DEM), during implementation (PMR) and at exit (XPMR).*

***Development Effectiveness Matrix (DEM)** is used as a common instrument for rating the evaluability of projects at entry (before approval). The DEM has three categories: strategic relevance, development outcome, and additionality of the Bank’s contribution. Country strategies, public and private sector operations have their specially designed DEMs.*

***Progress Monitoring Report (PMR)** will incorporate the results framework, the risk mitigation matrices, the Environmental Review Report, the Loan Results Review and the DEM in order to enable project teams to monitor the outputs and outcomes of the project, and their delivery in terms of both cost and time. Another new aspect to improve the Bank’s monitoring process is the new methodology to classify projects on “Alert” and “Problem” included in the PMR. It is based on a performance indicator and percentage of disbursement which will provide information focused on the management of results delivery during project execution. The PMR is updated twice a year in September and in March. The **PMR** will collect most of the information for reporting on progress of the RF outputs. Furthermore, it will facilitate reporting on the links between project outputs and outcomes, outcomes and priorities in country strategies.*

*The **Expanded Performance Monitoring Report (XPMR)** will soon replace the Project Completion Report (PCR). The XPMR compares actual results achieved with the expected results presented at entry in the DEM, which includes the recalculation of the DEM score. It will report on: (i) time and cost of outputs achieved from the last PMR; (ii) on results achieved and evaluation methods used; and (iii) Bank’s performance during the life of the project.”*

ANNEX B. THE BANK'S STRATEGIC PRIORITIES IN THE DEO

DEO 2010

Review of results by IDB strategic priority areas	What is the report focused on? What types of results are reported?	All impact evaluations of Bank's projects in the DEO.
1. Social policy for equity and productivity.	<p>This chapter reports on 5 main thematic areas:</p> <ol style="list-style-type: none"> 1. Investing in early childhood - Education in IDB-9 1) Improving school quality - Education in IDB-9 2) Improving the functioning of labor markets and extending the coverage of social security - LK in IDB-9 3) Fostering social inclusion -Violence against women and women in politics – Partially in IDB-9 4) Addressing youth at risk - Not part of the IDB-9 <p>The focus in some of the areas reported in this chapter does not correspond exactly with the strategic priority areas of the IDB-9. The DEO 2010 mentions that the missing areas were covered in the previous DEO or will be covered in future reports.</p> <p>The DEO 2010 reports on the total number of projects and total commitments made in 2010 for each of the sections. However, these sections do not correspond to the sections included in the DEO 2010. It stands out that there was no project executed or for which a disbursement was allocated in the gender and diversity area.</p> <p><u>Investing in early childhood.</u></p> <p>The DEO 2010 reports on interventions in early childhood health, instead if early childhood education as it is suggested by the IDB-9.</p> <p>In reports on two interventions in Guatemala and Honduras regarding reproduction, maternal and neonatal health (RMNH). The section focuses on the description of the projects, even though it does not report the results. It mentions that impact evaluation will be conducted.</p> <p>Additionally, this section reports on Bank's projects that generate "knowledge" regarding early childhood, although it does not include the results of the projects. It also presents a more detailed description of projects in Colombia, Nicaragua, Jamaica and Guyana. Once again, some results are mentioned, although without presenting concrete figures.</p> <p><u>Improving School Quality.</u></p> <p>Despite the increase in school enrollment, LAC test scores in Mathematics and Natural Sciences are still below other regions. It reports on the results of experimental teaching methods in Math and Science in Argentina: Science, Technology and Creativity (CTC), Scientific Literacy Program (PAC), and Math for all.</p> <p>It also reports on <i>Enseña Chile</i>, a program that intends to improve school quality by improving teacher quality. The program quality was underway by the time the DEO 2010 was published, although it mentions that the results will report on student achievement using value-added measures, student intellectual and interpersonal abilities, and student behavior. Overall, the report describes a project that has no available results.</p> <p><i>Moving from School to Work. Cognitive and non-cognitive abilities affecting transition from school to work.</i></p> <p>This section reports on existing knowledge regarding the skills for a successful transition from school to work. It presents a survey called Employers Survey of Demand of Skills about the more relevant skills for hiring workers less than 25 years. It describes the results of the survey by</p>	<p><u>School quality</u></p> <p>Experimental teaching programs in Argentina</p> <p>-Statistically significant difference between treatment schools and control schools which used the new Math teaching method.</p> <p>-The difference was of 44 points (half standard deviation). Within the learning module arithmetic was the subject that reported more advancement compared to non-treatment groups.</p> <p>-In the case of the science programs: CTC was statistically significant different in Buenos Aires, and PAC in Tucumán. This result is not significant across the whole sample. As a result PAC (more cost effective) and math for all will be scaled nationwide.</p> <p><u>Functioning of labor Markets</u></p> <p>-PROCAJOVEN on-the job training in Panama (pg 85). Two modalities youth aged 18 to 29, and first-time job seekers between ages 16-23. PAIL nationwide legal-age, active job seekers already registered in the public employment service.</p> <p>-An impact evaluation done in 2006 which includes only PROCAJOVEN, show that there was an increase in 5% of employment for the entire sample.</p> <p>-IDB is preparing a second evaluation including both programs, preliminary evidence using PSM show a</p>

Review of results by IDB strategic priority areas	What is the report focused on? What types of results are reported?	All impact evaluations of Bank's projects in the DEO.
	<p>identifying the cognitive and non-cognitive availability demanded by the firms in hiring a first time worker. The connection between the results of the survey and development effectiveness is not clear, as a survey would be only the diagnostic of the problem.</p> <p><u>Improving the functioning of labor markets.</u> This section describes the program <i>Bécate (Becas de Capacitación para el Trabajo)</i>, implemented in Mexico, and an Employment program in Bolivia. It provides the description of program evaluations associated with each of the programs, although no results were available by the time the DEO was published. Finally, it reports on some of the preliminary results of the job training program PROCAJOVEN in Panama.</p> <p><u>Fostering social inclusion.</u> This section describes previous IDB projects regarding Violence against women (VAW), although it does not present the associated results. In order to generate knowledge on VAW the Bank has designed a Knowledge and Capacity Building product (KCP). It also mentions the program Sol y Luna implemented in Medellin Colombia, describing it as a “good practice example.” However, any empirical evidence is provided to sustain this claim. An impact evaluation was designed, but no results were available by the time the DEO was published.</p> <p><u>Addressing youth at risk.</u> Following the structure of the previous sections, the program Music and Life program in Colombia and PROLEAD (Program for the Support of Women's Leadership and Representation) are described, without presenting systematic evidence on the beneficiaries of the programs. Again, future impact evaluations to assess the programs are promised, along with some background evidence not necessarily related with the IDB operations.</p>	<p>significant increase of employment in 22% for both modalities, and a large 37% for the PAIL modality. This result is observed because the program requires hiring 80% of beneficiaries of the program and measurement was done 18 months after the start of the program.</p>
<p>2. Infrastructure for competitiveness and social welfare.</p>	<p>This chapter reports on three main thematic areas:</p> <ol style="list-style-type: none"> 1. Water and Sanitation – Basic Services in IDB-9 2. Transport Infrastructure – Productive Infrastructure 3. Energy and Urban Development – Not part of the IDB-9 <p>The third area does not correspond with what is mandated in the IDB-9. As in the previous chapter the DEO provides information about the total number of projects approved in this priority area, as well as the total commitments made in 2010.</p> <p>It states that no impact evaluations were done at the moment the DEO was published, as they were planned to be implemented in the future. It presents a description of the programs with evaluations in the pipeline.</p> <p>The DEO reports on the infrastructure projects that had cost-benefit analysis and those that calculated an economic rate of return. Yet, it does not report on the average ERR.</p> <p>It describes the impact evaluation of <i>Favela II Bairro</i> program in Rio, which finished being executed in 2007. Additionally, the Bank implemented the program <i>Habitat</i> in Mexico, which aims to reduce urban poverty. The Mexican government mandated annual evaluations of the program.</p> <p>Regarding the water and sanitation thematic area, it presents the Spanish Cooperation Fund for Water and Sanitation program in Paraguay. There is only a mere description of the program, as no impact evaluation was available.</p> <p><u>Energy.</u></p>	<p><u>Favela II Bairro</u> -According to the impact evaluation conducted there was an impact on connection to the city water system 81% compared to the 55% control group. -There was a small 3% effect on formal ownership. 13% of the treated had an increase of informal documentation on ownership (such as bills of sale). -There was also an increase in 74% of value per square meter of treated homes. -There was no impact found on participation of favelas activities, or time spent in transportation. -Small but statistically significant impact on school</p>

Review of results by IDB strategic priority areas	What is the report focused on? What types of results are reported?	All impact evaluations of Bank's projects in the DEO.
	<p>In the thematic area of Energy there is a special emphasis on hydropower plants, where the Bank is financing the rehabilitation of three plants. It presents an <i>ex-ante</i> cost benefit analysis of the programs, although it does not provide information regarding <i>ex-post</i> cost benefits or outcomes of the programs.</p> <p>This section presents a detailed description of an energy private sector in Brazil. This description contains several outcome indicators of projects approved since 1994. The selection criteria of these projects it not clear as the report does not mention whether they provide an overall view of the energy private sector financing in Brazil.</p> <p>Lastly, a description of future sustainable energy programs in the Caribbean, specifically in the Caribbean and Bahamas. No information regarding the approval date, the committed amounts, or the potential disbursement dates is provided, and all the results reported are only <i>expected results</i>.</p> <p>There is also a brief description of sustainable energy programs in the Caribbean. First it describes a program in Barbados that will be executed. Here it is not clear when the project was approved, the amount that the project will disburse, and the results reported <i>are expected results</i>. There is also a description of a program in Bahamas, once again it is not clear when the project was approved, the amount committed, and the results reported are expected results of the program.</p> <p><u>Transport.</u></p> <p>The main focus of the thematic area of transportation is in rural roads in El Salvador. This part presents a causal chain to report on the results of these projects. Even though there is an explanation of the indicators, they do not contain specific values of baselines, milestones or targets.</p> <p>This section concludes with an unproved statement that the DEM provides better indicators at approval, which will allow to measure benefits during project execution and provide better evidence for doing an ERR at the end of project execution. The project Puerto Cortés in Honduras is used as an example of an ERR of 20%, however the ERR is calculated ex-ante and no information on <i>ex-post</i> calculations is provided.</p>	<p>attendance among aged 5-20 year old.</p> <p>-An increase in daycare attendance. Most importantly income of households increased in 15%.</p> <p>-There was no effect on employment or type of work.</p> <p>-There was also a marked improvement in household perception of quality of life in their community as a result of the program.</p> <p><u>Habitat program.</u></p> <p>-At the <i>polygon</i> level there was an increase of access to sewage services of 3% compared to the treatment group.</p> <p>-They explain several methodological problems the evaluation: the control group could have been treated, and the unit of analysis polygon was not sufficient given that the intervention was focused on several blocks within each polygon.</p> <p>-An evaluation of the second phase is being prepared with experimental design to appropriately identify the effects of the program. The results of the baseline are presented, and no statistically different result between treatment and comparison group were found.</p>
3. Institution for growth and social welfare.	<p>This chapter reports on five main thematic areas:</p> <ol style="list-style-type: none"> 1. Credit Markets – Financial Services in 2. Fiscal Efficiency and Sustainability – Fiscal Efficiency and Sustainability in IDB-9 3. Citizen Security – Citizens Security in IDB-9 4. Institutional Strengthening - Fiscal Efficiency and Sustainability in IDB-9 5. Country Systems - Not part of the IDB-9 <p>As in the other chapters of the DEO, this chapter provides information about the number of projects approved and the disbursed amounts. It draws the</p>	<p><u>Public Credit Policies in Ecuador</u></p> <p>-This study estimates the effect of public credit on firms' investments. It Is done by using a firm level panel data from 1996-2007.</p> <p>-PSM was used to find a counterfactual in the sample, and using DID</p>

Review of results by IDB strategic priority areas	What is the report focused on? What types of results are reported?	All impact evaluations of Bank's projects in the DEO.
	<p>attentions that a significant share of the commitments made in 2010 belongs to the Fiscal Efficiency and Sustainability section with \$2,610 USD millions, while only \$5 million were allocated to the Citizen Security division. The Bank has centered its efforts in institutions that promote economic growth, and address both market and State failures. A total of 14 diagnostics were conducted in each country to address specific problems on the countries' context.</p> <p>The first part of this section focuses on credit constraints, specifically to SMEs. The Bank operations are concentrated in the provision of global credit lines, which provide resources to financial intermediaries (IFI). The idea behind this is that the IFI assume the risk of repayment by firms, having the incentive of selecting only those projects and firms that are the most viable and productive. This is under the assumption that IFIs have a better understanding of the local market.</p> <p>In order to explain its operations with more detail, the case <i>Bancoldex</i> in Colombia is presented. This entity has a special line for MSMEs since 2004, called <i>Progresar</i>, through which the IDB provides financing. There is a description of the impact evaluation to assess this program, although no results are presented.</p> <p>In the same line, this section provides another example of Credit Policies in Brazil, with the corresponding results of an evaluation that was conducted to estimate the effect of the project. A final example on this is the program conducted by OMJ, which allocated a \$10 million loan to <i>Empresas Publicas de Medellin (EPM)</i> in order to increase financing to low income population. It presents a description of the program, which is based on the fact that EPM has information on costumer payments, reducing the binding constraints such as lack of credit histories by its financing. Once again, a potential impact evaluation is promised.</p> <p><u>Fiscal management</u></p> <p>The Bank uses the <i>Strengthened Approach</i> developed by the Public Expenditure and Financial Accountability initiative (PEFA). This approach gives metrics to assess the progress of the reforms. In 2010 the IDB was involved in PEFA assessments for eight countries in the region. Through these assessments the Bank was able to structure loans to Nicaragua and Peru. Throughout a cost effective analysis it estimates an IRR of 26% for Nicaragua and 23% for Peru. It presents an <i>ex-ante</i> analysis, without providing information of <i>ex-post</i> results.</p> <p>A study financed by the Bank in Jamaica (2004) pointed out that extensive tax exemptions generated forgone revenue that could have be used for expenditure. Based on this diagnostic, a second operation of The Fiscal Consolidation Program was designed, which aims to address the weaknesses in the Jamaican Fiscal sector. Once again there are no results associated with this example. As in most of the previous examples, the description of the evaluation is presented, without available results.</p> <p><u>Citizen Security.</u></p> <p>The DEO recognizes an overall lack of quality to assess the impact of projects conducted by the Bank regarding Citizen Security. For this reason the Bank pushed forward three projects in Belize, el Salvador and Costa Rica in this area that included the design of rigorous impact evaluations. The potential evaluations are described, without including results.</p>	<p>methodology it was estimated that the program generated an effect of 24% on employment, and 40% on exports and almost no significant impact on average standardized wages.</p> <p>-The effect on exports is through existent exporting firms, which suggest that the program had no effect on the probability of becoming an exported, but on the volume of exports of the exporters at the beginning of the program.</p> <p><u>Remittances in the Region. Salvadorian Migrants.</u></p> <p>-In a first experiment, a reduction on remittances fees was assigned randomly to remittances sent by Salvadorian Migrants in Washington DC.</p> <p>-The evidence shows that the ones that received a reduction in transaction fees increased frequency of transactions done but not the amount.</p> <p>-Evidence also suggests that there is no inter-temporal substitution: more remittances were not sent early in order to benefit from the reduction of the fee.</p> <p>-In a second experiment, Salvadorian migrants and the receiving households were chosen randomly to create or facilitate access to various saving mechanisms.</p> <p>-Surveys conducted to both migrants and receiving households suggested that migrants are more likely to choose savings than remittance-receiving households.</p> <p>-For the migrant, there is more attention given to the</p>

Review of results by IDB strategic priority areas	What is the report focused on? What types of results are reported?	All impact evaluations of Bank's projects in the DEO.
		<p>fraction that is saved from remittances. The higher the control for the migrant, the more savings they preferred. -Despite an effect on savings, this control over savings did not change the total amount of remittances sent to the home country.</p>
<p>4. Competitive regional and global international integration.</p>	<p>This chapter reports on five main thematic areas:</p> <ol style="list-style-type: none"> 1. Business Environment Improvement - Not part of the IDB-9 2. Promoting Productivity through Clusters and Local Development – Trade and Integration 3. Fostering Science – Not part of the IDB-9 4. Innovation and Technology Adoption - Not part of the IDB-9 5. Export Investment, Promotion and Enhancement – Trade and Integration in IDB-9 <p>It provides a detail description with examples in each of the areas, with the exception of Business Environment Improvement. It is not clear the reason behind this exclusion. Besides to the fact that only two of the areas is actually part of the IDB-9, conceptually there is not an explicit connection between competitive regional and global integration, and the selected categories.</p> <p><u>Export and Investment Promotion and Trade Enhancement.</u></p> <p>This section reports on Bank's knowledge products done by the Bank, with special emphasis on <i>Odyssey in International Markets</i> and an executive course to brainstorm alternatives to increase effectiveness of trade promotion in the region. The products are described in terms of plans, without providing concrete results. Thus, even though there have been some improvements in generating knowledge about trade enhancement, this has not produce specific results yet.</p> <p>There is a research agenda to understand the economic consequences of 20 years of opening LAC economies to external competition as the results of the abandon of the import substitution model. The results of the research were expected to be ready by 2011 and published on the next year's DEO.</p> <p>Also there is mention on a research agenda of the bank to further understand the economic consequences of 20 years of opening LAC economies to external completion as a result of changing the import substitution model.</p> <p>A report on India and LAC is briefly described. At the end of the description there are policy recommendations for both maximizing opportunities in bilateral trade with India. It emphasizes the need to address the deficiencies in education, access to credit, science and technology and infrastructure. The DEO is systematic in reporting the results of this investigation, but it is unclear if these recommendations are addressed by the Bank to evidence further use of the research product (as suggested by the IDB-9 requirements). Also it is unclear why a report on India and LAC is included as part of the development effectiveness overview.</p> <p>It describes the <i>Salud Mesoamérica</i> 2015 initiative, which consists of a joint effort of several development agencies. The strategy is presented by delineating the objectives and the funding system, although it is not clear the reason why this initiative is presented as part of the DEO considering that it will be developed in the future.</p>	<p><u>COLCIENCIAS</u> (Crespi, Maffioli and Melendez, 2010).</p> <p>-The program had a positive impact on firms' investment on R&D, and an impact on performance. These effects remain, and increase over time.</p> <p>-Impact on new products 12% and labor productivity 15%. The data is from 1995-2007.</p> <p><u>Programa de Apoyo a la Reestructuración anual (PRE)</u> (Castillo et al (2010)).</p> <p>-The PRE program was implemented between 1999 and 2007 and aimed to foster firm competitiveness by developing the market of professional services for SMEs and promoting access to those services.</p> <p>-The results show that the program was effective in increasing employment, real wages, and the probability of exporting.</p> <p><u>MIF</u></p> <p>-TRAZ.AR was a program implanted in Santa Fe, Argentina and consisted on developing software to store and administer information associated with the displacement of livestock.</p> <p>-Livestock identification and registration. The idea was to increase efficiency and administration processes of livestock</p>

Review of results by IDB strategic priority areas	What is the report focused on? What types of results are reported?	All impact evaluations of Bank's projects in the DEO.
	<p>The DEO also describes an initiative to give private financing to trade. Although it explains the nature of the fund to finance this initiative, no information regarding the size of the fund or the amount of the loans is provided. Moreover, it does not include information about the results from the private window financing.</p> <p>This section identifies a knowledge gap regarding the relationship between innovation and trade. In recognition of this problem, the Bank supported an initiative to address the knowledge gap, in which several researches with counterfactual were going to be used. The first country to analyze was Chile to fine-tune the evaluation methodology. The description about this initiative is short. Thus it is not clear either what the knowledge product will produce or the methodology it will use to produce it.</p> <p><u>Fostering Science, Technology adoption and Innovation.</u></p> <p>This section explains some of the evaluations about the use of technology development funds. In the case of the COLCIENCIAS evaluation, it uses data of 1995-2007.</p> <p>Another research gap area in the region pushed forward by the Bank is the one that explores the effect of innovation on employment in Latin America and the Caribbean. Innovations in LAC are characterized by imitation of technology from developed countries. Also the technology used depends on the size of the firm utilizing it. Smaller firms informally search for existing technologies while large firms' innovation comes from formalized in R&D laboratories. As part of its research agenda, the Bank is conducting evaluations on how innovation policies foster employment. Three potential impact evaluations in Chile, Argentina and Colombia are announced.</p> <p>This section presents the description of potential impact evaluations in Colombia, financed by a funding program of \$2 million USD. With the help of IDB and COLCIENCIAS, the impact evaluations cover three strategic areas: building human capital, promoting research in the sciences; elementary and middle school science education; and innovation by Colombian firms.</p> <p>Lastly, it describes a program with the Argentine Government to strengthen the national innovation system. Once again, a potential impact evaluation is announced.</p> <p><u>Challenges on evaluating cluster policies.</u></p> <p>This section is not required by the IDB-9, and is it not clear why it is included in this section of the DEO: "Competitive regional and global international integration". The selected projects correspond to local economic development, which conceptually can be placed in a different section.</p> <p>It identifies a knowledge gap regarding the effect of cluster policies. Therefore the Bank aims at closing this knowledge gap by producing a methodological toolkit for the evaluation of cluster policies, and a pilot evaluation of the cluster policy in Brazil. It provides a description on how the methodological toolkit will operate and, a description of the pilot evaluation in Minas Gerais and Sao Paulo. The idea here is to look for a counterfactual in municipalities where there was no treatment by the program.</p>	<p>activities, and increase the probability of selling in higher value markets.</p> <p>-The program participants had higher capacity to adapt to external shocks such as weather related issues, and an effect of 13 to 14 % in the stock of livestock and an increase in the employment of qualified workers.</p> <p><u>Business Development Program for Software Industry in Uruguay.</u></p> <p>-The program provided inputs for quality enhancement, managerial strengthening, and standard setting for the software industry.</p> <p>-The program provided 1) training for 160 SMEs in the software sector. 2) an estimated 600 SME staff trained on internationalization, management, and quality control.</p> <p>-PASS participants show higher rates in sales and exports as well as employment compared to the control group. (No magnitude of the coefficients reported).</p>
5. Protect the environment, respond to climate change and enhance food security.	<p>This chapter reports on five main areas of the IDB-9 thematic classification:</p> <ol style="list-style-type: none"> 1. Agriculture and Food Security – Food Security in IDB-9 2. Disaster Risk Management – Not part of the IDB-9 3. Climate Change – Environment and Climate Change in IDB-9 4. Sustainable Tourism – Not part of the IDB -9 5. Natural Resources Management – Not part of the IDB-9 <p>The chapter reports on the total number of projects and total financing</p>	<p><u>PATCA</u></p> <p>-The Program to Support Technology Innovation in Agriculture was initially implemented in 2002.</p> <p>-OVE did an impact evaluation of the first phase</p>

Review of results by IDB strategic priority areas	What is the report focused on? What types of results are reported?	All impact evaluations of Bank's projects in the DEO.
	<p>approved in this area, although the information does not correspond with the thematic areas above-mentioned. Furthermore, the IDB-9 asks for sustainable transport, alternative fuels, renewable energy and energy efficiency, which is not included as part of the thematic classification and no information about the number of programs or financing is provided.</p> <p><u>Food security: Enhancing Productivity in Agriculture.</u></p> <p><i>PATCA</i> program provided financial support to farmer's through vouchers and private providers in order to promote different types of technology adoption. It presents the results of the impact evaluation of the first phase of <i>PATCA</i>. For <i>PATCA II</i>, information about the design of the impact evaluation and the data collections methods is included, although no results were available by the time the DEO was published.</p> <p>Additionally, this section includes the description of the program <i>Programa de Apoyos Directo al Campo</i> implemented in Mexico. The program was a response to the decline in prices of agricultural products after the start of the North-American Free Trade Agreement (NAFTA). This program provides subsidies, and gives financing for productive investments. There is no results available, although an impact evaluation of the program is promised.</p> <p><u>Tourism for development and environmental sustainability.</u></p> <p>This section presents a brief description of the work that the Bank has done previously in the tourism section, and how it has evolved to the current strategy. The Bank has developed a strategy to fill knowledge gaps related to the role of tourism in development and poverty reduction. A brief literature survey about the possible effects of tourism and growth and poverty reduction is described, although it is not related to the Bank operations. Likewise, it presents the challenges of doing impact evaluations in the tourism sector, where the lack of reliable data is one of the most important. For this reason, the Bank started the implementation of a TC to support the harmonization of tourism statistics and creation of Tourism Satellite Accounts.</p> <p>It also describes possible methods of impact evaluation in the tourism sector, as the computable general equilibrium model method used by the Bank in the case of PRODETUR project in RIO. The Bank used this methodology <i>ex-ante</i> to determine the anticipated impacts of the project, and depending on the data collected it will be used to assess the impact of the project <i>ex-post</i>.</p> <p><u>Promoting environmental sustainability and planning for climate change impact.</u></p> <p>This section presents a program that seeks to reduce vulnerability to climate change of rural communities in Nicaragua. As it occurs on most of the sections, a potential impact evaluation is announced without presenting specific evidence or results.</p>	<p>of the program. The program had an impact on rice productivity and breeding, but no significant impact on other crop-productivity measures or on milk productivity. (There are no coefficients reported, nor controls used).</p>

Are there lessons learned presented in the 2010 DEO? How are they presented? Are they presented in line with IDB strategic areas?

R/. There are lessons learned as the DEO 2010 uses a theoretical and empirical background to justify the interventions of the Bank. However, these lessons do not necessarily come from the Bank projects, but from an overall policy experience in the sector. The DEO is not systematic in presenting lessons learned in a separate section, or in each thematic area. This is in part due to the structure of the DEO, which combines several unrelated examples and projects without lessons with some other that actually provide lessons learned.

DEO 2011

Review of results by IDB strategic priority areas	What is the report focused on? What types of results are reported?	List all impact evaluations of Bank results. (And main results presented.)
<p>1. Social policy for equity and productivity.</p>	<p>The chapter presents six examples of projects that the Bank approved in 2011:</p> <ol style="list-style-type: none"> 1. Sanitation project in Mexico's primary school – Health in IDB - 9 2. Haiti's plan for education reform – Education in IDB - 9 3. Salud Mesomerica 2015 – Safety nets for the poor and Health in IDB - 9 4. Social safety net in Trinidad and Tobago – Safety nets for the poor in IDB - 9 5. Women's right and empowerment in El Salvador – Gender and Diversity in IDB – 9 6. Youth Training in Dominican Republic – Labor Markets in IDB - 9 <p>The chapter does not include information regarding the number of projects and the total financing allocated to this strategic priority area in 2011. Likewise, it is no included the rationality behind the selection of these projects, although it is mentioned on the introduction that the projects were chosen due to a unique path in their design (unusual development challenge, out of the box solution or innovative impact evaluation).</p> <p><u>Water, Sanitation and Hygiene Program in Basic Education (PASHEB) – Mexico</u></p> <p>PASHEB was developed to reduce the risk of illness caused by unsafe water and bathroom facilities through the improvement of school infrastructure and promotion of good hygiene. There are three main objectives in evaluating the program: effectiveness, evidence-based results and sustainability. One of the components of the evaluation will be a comparative study between two groups of parents for which financial support to maintain the bathrooms will be given at different stages over 3 years.</p> <p>This section does not provide information about the role of the IDB in this program. Even though some aspects of the evaluation are explained, the dates when the evaluation will be implemented, completed or presented are not mentioned. No additional results on other evaluations of similar projects are presented.</p> <p><u>New and Improved Haitian Education System - Haiti</u></p> <p>In 2011, the IDB approved a second grant to the government's reform plan (US \$50), which includes an education reform, whit advice from the designers of the education reconstruction plan after Katrina. The expected outputs of the program for the academic year 2012-13 are 220,000 tuition waivers for non-public schools and supplies kits for 30,000 children and 2,000 teachers.</p> <p>The interactive radio instruction (IRI) program to provide educational materials in remote areas will be also implemented. An experimental evaluation will be conducted to evaluate the radio program. In the first phase 200 schools will receive the program for grades 2nd, 3rd and 4th, while the other group will serve as control. It is included evidence for the program “<i>little mathematicians</i>” in Paraguay supported by the Bank. No information regarding the dates for the implementation or publication of the results is provided.</p>	<p><u>Little mathematicians – Paraguay</u></p> <p><i>Evaluation year/coefficients presented:</i> Not specified/No</p> <p>-The program develops solid foundations in basic math among children aged for to six.</p> <p>-Positive results in terms of achievement in math.</p> <p><u>Juventud y Empleo- Dominican Republic</u></p> <p><i>Evaluation year/coefficients presented:</i> 2011/Partially</p> <p>-Positive impact in terms of job quality, in particular for men.</p> <p>-Participants have higher expectations about the future than those that did not participate in the program. This was particularly important for women and youngest.</p> <p>-Teenage pregnancy decreased about 5 percentage points from a teenage pregnancy rate of eleven points among young women that did not participate in the program.</p> <p>Improvements of about ten percent of a standard deviation attributable to the seventy-five hour life</p>

Review of results by IDB strategic priority areas	What is the report focused on? What types of results are reported?	List all impact evaluations of <u>Bank</u> results. (And main results presented.)
	<p><u>Salud Mesoamerica 2015 (SM2015)</u></p> <p>The section presents some literature about conditional cash transfers in health. It describes the design of the program, which is not financed but executed by the IDB. The program includes rounds of survey data and an impact evaluation agenda. It presents the case of El Salvador, where outcomes in municipalities financed by the initiative will be compared with municipalities which receive unconditional funds. In Nicaragua, the initiative will conduct an intervention consisting of Community Based Agreements in health indicators to test if commitments are stronger than cash.</p> <p>Although SM2015 was announced since the DEO 2008-09, any of the three reports have presented results on the interventions.</p> <p><u>Social Safety Nets – Trinidad and Tobago</u></p> <p>Since 2010, the government is implementing a program of social safety net reforms, based on cash grants and diverse methodologies for enrolling beneficiaries. The impact evaluation of the program will focus on “second generation” questions for conditional cash transfers, i.e. how targeting results in adjustments to the composition of household receiving benefits, the impact of withdrawing ineligible households, and how to do this effectively. However, the details on how the evaluation will be conducted are not explained.</p> <p>The section does not explain the role of the IDB either in the implementation or the evaluation of the reforms. Although a potential evaluation is explained, it does not mention the dates for the implementation or publication of the results.</p> <p><u>Ciudad Mujer – El Salvador</u></p> <p>The program consists of one-stop centers to provide critical services for health services, economic autonomy, violence against women and community education. An experimental impact evaluation will be conducted. The proposed methodology is to randomize the promotion of the services offered and complement this by exploiting the gradual implementation of the project. As in the previous section, no information regarding the IDB’s role or when the evaluation will be conducted is provided.</p> <p><u>Youth Training – Dominican Republic</u></p> <p>The program <i>Juventud y Empleo</i> was designed in 1999 and had an experimental evaluation design from its start, using the results to incorporate the feedback to the program. The most recent evaluations, conducted by the IDB, show a positive impact in terms of job quality, positive employment effects for women, and participants had higher expectations than those that did not participate in the program. The evaluation also showed that most of the benefits from the program are attributable to the life skills component and not the technical. Based on this, the program will double the hours dedicated to life skills, and will limit technical training only to those individuals that are likely to benefit from them.</p>	<p>skills training module.</p>

Review of results by IDB strategic priority areas	What is the report focused on? What types of results are reported?	List all impact evaluations of <u>Bank</u> results. (And main results presented.)
<p>2. Infrastructure for competitiveness and social welfare.</p>	<p>The chapter presents five examples of projects that the Bank approved in 2011:</p> <ol style="list-style-type: none"> 1. Housing in Nicaragua – Not mentioned explicitly in IDB - 9 2. Water and Sanitation in Bolivia – Basic Services in IDB - 9 3. Sustainable Energy in Barbados – Productive Infrastructure in IDB - 9 4. Air Transport system in Bahamas – Not mentioned explicitly in IDB - 9 5. Beltway <i>Rodoanel</i> in Brazil – Productive Infrastructure in IDB - 9 <p>It is mentioned that these examples were chosen because they were based on sector diagnosis supported by the IDB, and for which evaluations arrangements were defined to learn about their effectiveness. However, the chapter does not include information regarding the number of projects and the total financing allocated to this strategic priority area in 2011.</p> <p><u>Provision of Housing in Nicaragua</u></p> <p>In 2011 the government of Nicaragua requested a loan from the IDB to help address the housing problem of low-income families (US\$40 million). A diagnostic study of the sector was commissioned to have a clear understanding of the housing needs. The program will fund housing to those that are least served and where the deficit is highest (the modalities are explained). It also ensures that other programs are attending the higher income levels, not included in the program. It is also mentioned that the program includes a monitoring and evaluation framework, although the details of the evaluation methodology are not explained.</p> <p><u>Water and Sanitation in Bolivia</u></p> <p>The Bolivian Government, in partnership with the IDB and the Spain Government, is implementing an impact evaluation that will measure how the water and sanitation interventions have affected the health of young children and the economic lives of families. The program will conduct a public lottery in the presence of community leaders to select which communities will be offered financing in each year of the program. Communities receiving the interventions at the beginning of the program will be compared to those elected for funding in the latter stages to evaluate causal effects. The evaluation will measure health outcomes amongst children related to diarrhea, parasites and respiratory diseases. The chapter does not provide results on any other project, and it does not mention the dates when the evaluation will be implemented.</p> <p><u>Sustainable energy for Barbados</u></p> <p>Barbados spent 10% of its GDP on oil imports and its own electricity bill in 2010. Before launched a program to save energy, the Government asked for technical assistance funded by the IDB. With the information of the study, the Government and the IDB designed a programmatic operation that would promote sustainable use of energy. The Program's Monitoring and Evaluation Plan include indicators to capture project-level output information, like electricity generated and electricity saved. The report does not present results or indicators, as it can be assumed that the implementation of the project was at an early stage by the time the report was published.</p>	

Review of results by IDB strategic priority areas	What is the report focused on? What types of results are reported?	List all impact evaluations of <u>Bank</u> results. (And main results presented.)
	<p><u>Air Transport in Bahamas</u></p> <p>In 2011 the IDB supported the Air Transport Reform Program, which aims to separate the functions performed by the Civil Aviation Department (CAD). In designing the reform, the government supported by the IDB, will take previous experiences from Canada, Australia and Zealand, which had a similar air transport systems as Bahamas before a reform took place. Indicators such as the Intervention Air Transportation connectivity index, the average operating expenses per traffic unit, and the number of incidents per year will be used to evaluate the reform. Although the reforms in the above-mention countries are consider successful, no impact evaluations are presented in order to support this affirmation.</p> <p><u>Rodoanel in Brazil</u></p> <p>The Bank has supported Brazil with financial and technical support in the construction of the northern section of Sao Paulo's beltway (the IDB also supported the construction of the west part). The technical analysis, supported by the Bank, designed the best alternative to complete the beltway, considering social and environmental issues and incorporating lessons from the construction of the rest of it. This section presents results of the construction from the operation of the 93 km of <i>Rodoanel</i>: a reduction of truck traffic on the city's main roads by 43%, and the level of airborne particulate matter within this area by 47%. During the rush morning time, the average travel time has been reduced from 88 minutes from Bandeirantes and Imigrantes expressways to 40 minutes. However, the source of these results is not mentioned, and no future evaluation of the northern section is presented.</p>	
<p>3. Institution for growth and social welfare.</p>	<p>The chapter presents four examples of projects that the Bank approved in 2011:</p> <ol style="list-style-type: none"> 1. Access to credit for microenterprises in Ecuador – Financial Services in IDB - 9 2. Combating corruption in Brazil – Citizen's Security in IDB - 9 3. Reform of pension institutions in Honduras – Fiscal efficiency and sustainability in IDB - 9 4. Strengthening police effectiveness in Ecuador – Citizen's Security in IDB - 9 <p>It is mentioned that these examples were chosen because they will strengthen key institutions in the region. The chapter does not include information regarding the number of projects and the total financing allocated to this strategic priority area in 2011.</p> <p><u>Global Credit for Microfinance Support Program (GMSP) in Ecuador</u></p> <p>GMSP will provide credit to microfinance institutions that lend to microenterprises. The goal of the GMSP is to raise incomes, increase hiring, and improve business performance among clients of participating institutions. Available evidence on previous evaluations of similar interventions is used to support the design of the evaluation of the GMSP. The GMSP evaluation will randomly pick borrowers from a pre-screened list of potential credit clients to receive the offer of a loan. By introducing random variation in the offer of credit, the evaluation design will make it possible to measure the average impact of having a loan on microenterprises that would take a loan when it is</p>	

Review of results by IDB strategic priority areas	What is the report focused on? What types of results are reported?	List all impact evaluations of <u>Bank</u> results. (And main results presented.)
	<p>offered to them. The section does not give an expected time to conduct the evaluation or have the results available.</p> <p><u>Strengthening of the Federal Government's Audit Agency in Brazil</u></p> <p>Since 2003, the Brazil government began implementing random audits on municipal government use of federal funds, particularly in health and education. A lottery on which municipalities get an audit is publicly broadcasted three times a year, and a few days later a group of auditors descends on the chosen municipalities. The section presents results on studies conducted in 2009 and 2011 by external evaluators, not the IDB, with positive results in education and corruption. The IDB support seeks to contribute to the program to continue improving public sector integrity. It will provide training, and access to a new transparency portal that will be interconnected with similar portals at the state and municipal levels. The section does not present information about a future evaluation conducted by the IDB of this program.</p> <p><u>Reform of Pension Institutions in Honduras</u></p> <p>In 2011, the Bank supported the Government of Honduras with a program that will help to alleviate the fiscal impacts of the actuarial and liquidity problems of the National Pension Institute for Teachers (INPREMA) and the National Institute of Retirement and Pensions for Public Officials and Government Employees (INJUPEMP). It also seeks to protect the financial position and improving operations management in the Honduran Social Security Institute (IHSS). This section explains the main proposed actions. It is expected that when the first programmatic loan ends, it will have contributed to the long term stability of the pension systems and to reduce the contingent fiscal liabilities. Also, it is anticipated that approval of the draft legislation will result in a reduction in the actuarial deficit of 27.7% of GDP in 2010 to a range between 9-13%. No information regarding the amount of the loan, the end date of the first programmatic loan, or a potential evaluation of the intervention is provided.</p> <p><u>Strengthening police effectiveness in Ecuador</u></p> <p>The Government of Ecuador and the IDB partnered to design a program that will gather data and create information systems, arguing that a key problem in the region is the lack of adequate diagnosis. The program entails training in the criminological analysis of the data, which will be incorporated in the design and implementation of evidence-based solutions. The project incorporates an evaluation to determine whether the new protocols have the intended effects and how well are implemented. The approach of the evaluation will be based on three methodologies: evidence-based-policing, problem oriented policing, and hot spot policing operatives. This section explains all the details regarding these methodologies, although it does not provide more information on the amount committed or the publication of results.</p>	
4. Competitive regional and global international integration.	<p>This chapter is focused on the assistance the Bank will provide to Uruguay to enhance its access to global export markets. It does not explain the reasoning behind the selection of this example. Although it brings an overview of the Uruguayan economy, the potential it has to become a strong player in the global export services market, and the different problems it has to overcome.</p>	

Review of results by IDB strategic priority areas	What is the report focused on? What types of results are reported?	List all impact evaluations of <u>Bank</u> results. (And main results presented.)
	<p>As it occurs in the other chapters of the report, no information regarding the number of projects or the total financing allocated to this strategic priority area is mentioned.</p> <p>The expected outcomes include an improved international positioning in the global services industry, greater number of skilled workers, and improved times and costs associated with foreign trade processes. By the end of the program, exports in the sector are expected to have increased 55%, investment 66% and employment 28.4%, although the end date of the program is not provided.</p> <p>The evaluation of the program will incorporate three approaches: a difference-indifference matching approach to evaluate the effect of the program on service exports and employment to supported firms; a regression discontinuity design to measure impacts of the creation of the registry of skilled service sector workers; a difference-indifference approach to evaluate the impact of actions designed to improve the efficiency of the logistics sector. No additional examples or results from other evaluations are provided as it states that no rigorous impact evaluations of public programs for export of services exists so far. Thus, the evaluation of this program will be the first of its kind.</p>	
<p>5. Protect the environment, respond to climate change and enhance food security.</p>	<p>The chapter presents six examples of projects the Bank approved in 2011:</p> <ol style="list-style-type: none"> 1. Promote Food Health in Dominican Republic – Food energy in IDB - 9 2. Land Formalization in Paraguay – Food security in IDB - 9 3. Nautical Tourism in Uruguay – Not explicitly mentioned in IDB - 9 4. Climate Change in Panama and Trinidad y Tobago – Environment and Climate Change in IDB – 9 5. Disaster Risk Management in Peru - Environment and Climate Change in IDB - 9 6. Solar Energy in Chile - Environment and Climate Change in IDB - 9 <p>In contrast with the other chapters, Chapter V elaborates further on the selection criteria of the projects presented. No information regarding the number of projects or the total financing allocated to this strategic priority area in 2011 is provided.</p> <p><u>Promoting Food Health and Safety in Dominican Republic</u></p> <p>The program aims to strength all components of food and health system. On the producer side, it will promote exports and increase domestic food safety through certification of good practices. On the consumer side, it will foster consumer's understanding of the value of domestic food certification. The expected impacts are: raise agricultural productivity, improving sanitary standards (reducing export rejections) and reduce consumer's health through lower pesticides residues. The evaluation has the problem that the adoption of new production techniques is correlated with unobserved factors that may also affects outcomes. The evaluation will deal with this problem through the application of cutting edge experimental and quasi-experimental techniques. The details of the evaluation are well explained throughout the section. Nevertheless, no information about the expected availability of results is provided. It does not mention any results of similar interventions arguing that this intervention is unique and new in its policy area.</p>	

Review of results by IDB strategic priority areas	What is the report focused on? What types of results are reported?	List all impact evaluations of <u>Bank</u> results. (And main results presented.)
	<p><u>Land formalization in Paraguay (PROCAR II)</u></p> <p>PROCAR II will follow up in modernizing the cadaster and the registry started by PROCAR in 2004. It will also incorporate two elements: 1) it selected nine districts in which to pilot field verification and expansion of land titles and 2) it will help to establish the National Cadastre and Property Registry Directorate (DINACARE). The expected results are: increases in agricultural productivity, and through this, an improvement in the well-being of rural households. To conduct the impact evaluation, four districts will be randomly selected to be included at the start of the program (treatment group) and the remaining five districts will be incorporated two years later (control districts). In contrast with the other sections of the report, this section specifies that the evaluation will not be completed at least until 2013.</p> <p><u>Nautical Tourism in Uruguay</u></p> <p>The program supports the implementation of the nautical tourism master plan. The Tourism Satellite Account (developed with support of a Bank loan in 2006) was used to prepare an ex-ante analysis of the tourism sector. It is mentioned that the program includes an ex-post impact evaluation strategy, but it does not provide any detail on the methodology of such evaluation.</p> <p><u>Climate change policy-based loans: Panama and Trinidad y Tobago</u></p> <p>The policy-based loan in Panama will focus on integrating climate change considerations into the country's disaster risk management frameworks with emphasis on water management. In Trinidad y Tobago the goal is to mainstream climate change into national development planning and to promote carbon emissions growth strategy. In both cases, the Bank used an ex ante cost-benefit analysis. In Panama it showed that the investments are likely to have favorable net returns. In Trinidad y Tobago the analysis focused on a carbon sequestration project on afforestation and reforestation on coastal wetlands, with negative returns when based solely on the value of the carbon being sequestered for a range of prices. It does not mention any future impact evaluation for these projects.</p> <p><u>Reduce the Vulnerability of the State to Disasters Program in Peru</u></p> <p>Two programmatic policy based loans took place in 2010 and 2011. The major accomplishment was the establishment of the National Disaster Risk Management System (SINAGERD). The System of Indicators of Disaster Risk and Risk Management (which generates the RMI index) is used to evaluate the project's impact and results, although any specific project or result is mentioned. It is said that, in terms of evaluation, they will compare the indicators within the RMI before and after the project is in place, but no project or evaluation is including in this section.</p> <p><u>Solar Energy in Chile</u></p> <p>Through a NSG loan, the IDB supported <i>Subsole's</i> capital expenditure program for 2011-14. The project will develop one of the first photovoltaic (PV) plants and the first to be used in the agribusiness sector. As part of the project, the RedLAC network was created to share knowledge about PV technology. Funds from the Sustainable Energy and Climate Change Initiative were used to support a study for the PV plant, which determined that it has a capacity of 307 kWp during the first phase. However no study or evaluation to test this expected results is included in this section.</p>	

Are there lessons learned presented in the 2011 DEO? How are they presented? Are they presented in line with IDB strategic areas?

R/. DEO 2011 incorporates some lessons from both projects implemented by the IDB and similar projects not implemented in the region. While in some cases the Bank incorporates these lessons including technical assistance from the designers of other similar programs (Katrina), in other cases they add components that the previous programs did not have (PROCAR II), or modify the program based on studies conducted by the Bank (*Juventud y Empleo*).

Particularly, the report includes the following lessons:

1. To help the government to rebuild the Haitian educational system the Bank is using technical assistance of the reconstruction plan for the education sector of New Orleans after Katrina.
2. Based on a previous evaluation, the program *Juventud y Empleo* will double the hours dedicated to life skills, and will limit technical training only to those individuals that are likely to benefit from them.
3. The Bank uses previous experiences in Canada, Australia and New Zealand to design the reform of the air transport system in Bahamas. These countries had similar structure for airport operations, air traffic control, and technical regulations to what currently exists in Bahamas.
4. In designing the best alternative to build the northern section of Sao Paulo's beltway *Rodoanel*, the Bank followed lessons learned from the construction of the Southern and Eastern sections, in which the IDB had participation.
5. Two components derived from PROCAR were incorporated in the program PROCAR II. It will select districts in which to pilot field verification and expansion of land titles and it will help to establish the National Cadastre and Property Registry Directorate.

ANNEX C. IMPACT EVALUATIONS PROMISED IN THE DEO

Table 1.C List of Promised Impact Evaluations

2008-09 DEO	2010 DEO	2011 DEO
SOCIAL POLICY FOR EQUITY AND PRODUCTIVITY		
<p>1) Mexico, Colombia, El Salvador, Ecuador: Social protection through conditional cash transfers. There is an operative research agenda that, through innovations and evaluations, will answer some of the most pressing questions concerning urban CCTs. Pp. 66-70</p>	<p>1) Guatemala (SM2015): Improved Access and Quality of Health and Nutrition Services. The evaluation will focus on two main interventions: the strengthening of the community program of integral attention to women and childhood (AINM-C) and the expansion of the institutional mobile groups (GMI) with an educational component to isolated rural communities. Pp. 68-69</p> <p>2) Honduras (SM2015): Program to Strengthen Decentralized Management and Supply of Health Services. The impact evaluation will assess the effectiveness of a decentralized model for health-service provision. P. 70</p> <p>3) Colombia, Nicaragua, Jamaica and Guyana: The Bank's work on Early Childhood Development (ECD). The agenda of the Bank includes impact evaluations to build knowledge on ECD interventions in Colombia, Nicaragua, Jamaica and Guyana. Pp. 72-73</p> <p>4) Bolivia: Evaluating the Employment Programs in a Random Encouragement Design. The proposed evaluation strategy will use random encouragement given to potential users to use the Public Employment Service (SEP). This evaluation will help to strengthen the institutional capacity of the Ministry of Labor. P. 84</p>	<p>1) Mexico: Water, Sanitation, and Hygiene in Basic Education (PASHEB). One particularly aspect of the PASHEB evaluation is the comparative study of two groups of parents. This comparative analysis will allow to see how much the educational and institutional components promote community participation and what role financial resources play in incentivizing participation. Pp. 52.</p> <p>2) Haiti: lessons from Katrina for a new and improved Haitian education system. An experimental evaluation will be conducted alongside the Interactive Radio Instruction Program to provide evidence of the levels of success of this alternative delivery method for lessons. Pp. 52-54.</p> <p>3) El Salvador (SM2015): Using results-based financing to impact the health outcomes of the poorest. In order to evaluate the impact of Results-based financing at the national level, outcomes in municipalities financed under the initiative will be compared to similar municipalities which receive unconditional funds. Pp. 55-56.</p> <p>4) Nicaragua (SM2015): Are Commitments Stronger than Cash? The evaluation will assess whether CBAs based on a "moral" commitment can produce improvements in health, or if liquidity constraints and lack of motivation need to be overcome with monetary incentives. P. 56.</p>

2008-09 DEO	2010 DEO	2011 DEO
	<p>5) LAC: Building the Capacity of State and Local Actors to Respond to Violence Against Women. The KCP will design and conduct rigorous impact evaluations employing both experimental and non-experimental design, as well as quantitative and qualitative data-collection methods, where application of each design and method is deemed ethical, feasible and appropriate to program context. The programs to be evaluated will meet several criteria. Pp. 87-88</p> <p>6) Colombia: <i>Sol y Luna</i> (SyL), Curbing Teenage Pregnancy in Latin America. An initial analysis will consider the impact of SyL on attitudes towards sexual behavior and early childbearing, and the follow-up study will evaluate the impact of the project on sexual activity, adolescent pregnancy and fertility. Pp. 88-90</p> <p>7) Colombia: Music and Life: Music Schools for Youth. To conduct the impact evaluation, a randomized experimental design will be implemented to estimate the causal impacts of the music program on the outcomes of interest. Pp. 90-92</p> <p>8) LAC: Promoting Women's Political Empowerment (PROLEAD). Evaluation of the impact of two key intervention models funded by the initiative: (1) media campaigns to promote women's participation in politics; and (2) training of journalists, or people working in the media sector, to promote less gender-biased coverage of the elections. Pp. 92-93</p>	<p>5) Trinidad y Tobago: Improving efficiency and effectiveness of Social Safety Nets. The impact evaluation will focus on "second generation" questions for conditional cash transfers, for example how targeting results in adjustments to the composition of households receiving benefits, the impact of withdrawing ineligible households, and how to do this effectively. P. 56.</p> <p>6) El Salvador: <i>Ciudad Mujer</i>, Empowering Women to Promote Development. The proposed methodology is to randomize the promotion of the "one-stop centers" health, education, and prevention of violence services offered and possibly complement this by exploiting the gradual implementation of the project. P. 58.</p> <p>7) Dominican Republic: How Evaluation Findings on Socio-emotional Skills have Improved the Youth Training Program. Based on the second round of evaluations, the new <i>Juventud y Empleo</i> program will double the hours dedicated to life skills, and will limit technical training only to those individuals that are likely to benefit from them. Pp. 58-61.</p>

2008-09 DEO	2010 DEO	2011 DEO
INFRASTRUCTURE FOR COMPETITIVENESS AND SOCIAL WELFARE		
<p>1) LAC: Provision of Sustainable, reliable and affordable Energy. Within the cluster, project DR-L1026 (Dominican Republic) has formulated a concrete evaluation indicator (cash recovery index) to capture project impact across energy producers and users. Pp. 85-86</p>	<p>1) Paraguay. Water and Sanitation, Measuring their Impact on the Health of Rural Populations. The main goal of this program is to improve living conditions in rural communities by providing access to safe water supply and sanitation facilities. The evaluation design will use the natural timing and logistic limitations of the project to construct a counterfactual to the program. Pp. 105-106.</p> <p>2) El Salvador: Learning from Rural Roads for Development. The Government of El Salvador with support will finance eight transport projects. To measure program impacts, household and community surveys will be applied in the area of influence of the project, and of other eligible projects that will not be financed due to lack of resources. Pp. 113-114</p>	<p>1) Bolivia: Water and Sanitation for Small Towns and Rural Communities. The evaluation will measure a comprehensive set of outcomes: health outcomes amongst children, particularly those related to diarrhea, parasites and respiratory diseases and reduction in the cost of accessing water. Pp. 66-68.</p>
INSTITUTIONS FOR GROWTH AND SOCIAL WELFARE		
<p>1) Brazil, Chile, Paraguay, Peru: Access to credit and Small and Medium Enterprise Productivity. The Bank is developing methodological guidelines to support the adoption of state-of-the-art impact evaluation techniques in the areas of access to credit and SME productivity. In addition, it is initiating pilot impact evaluations of the effect of credit policies on microenterprises and SMEs in Brazil, Chile, Paraguay and Peru. Pp. 104-105.</p>	<p>1) Colombia: Evaluating the Impact of Credit. The evaluation will analyze the effectiveness of Bancoldex in expanding SME access to medium- and long-term financing and thereby supporting a more dynamic and sounder financial and economic performance by beneficiary firms. Pp.120-121</p>	<p>1) Ecuador: Global Credit for Microfinance Support Program (GMSP). The evaluation design will make it possible to measure the average impact of having a loan on microenterprises that would take a loan were it offered to them. Pp. 76-77.</p>

2008-09 DEO	2010 DEO	2011 DEO
<p>2) Trinidad and Tobago: Citizen Security. The evaluation will focus on studying the links between program components in order to disentangle the impact of single components from the impact of the program as a whole. Pp. 105-108.</p>	<p>2) Colombia: Empresas Públicas de Medellín. The evaluation will attempt to measure the impact of the program on: (1) the provision of credit to beneficiaries that were not “bankable” prior to the program; (2) access to other financial services provided through the formal banking system; (3) household purchases of new and more energy-efficient durable goods, and (4) house improvements. P. 122</p> <p>3) El Salvador, Belize, Costa Rica: Reducing crime in urban areas in El Salvador, Community Action for Public Safety Program in Belize, Reducing Violent Crime in Costa Rica. OVE’s recommendations and the Bank’s increasing emphasis on measuring results and learning from its projects led to a push to design rigorous evaluations for these three projects from the very initial stages. Pp. 133-136.</p>	
COMPETITIVE REGIONAL AND GLOBAL INTERNATIONAL INTEGRATION		
<p>1) Argentina, Paraguay, Uruguay: Fostering science and technology adoption. The SDV and CMF divisions have planned a joint work program on cluster and BE policies that includes: (i) methodological guidelines to support the adoption of state-of-the-art impact evaluation techniques; (ii) training to increase the ability of Bank specialists and government counterparts to define rigorous impact evaluation plans; (iii) pilot impact evaluations of IDB projects in these areas of intervention. Pp.122-127.</p>	<p>1) Colombia: Development of Human Capital and the Colombian System of Science, Technology and Innovation. The impact evaluations will be focused in three key areas of intervention: building human capital and promoting research in the sciences, elementary- and middle-school science education, and innovation by Colombian firms. Pp. 152-153</p>	<p>1) Uruguay: Support to global export services. The evaluation design will incorporate three different approaches to the measurement of impacts in service exports and employment for supported firms, the impacts of the creation of the registry of skilled service sector workers, and the impact of the impact of actions designed to improve the efficiency of the logistics sectors. Pp. 88-90.</p>

2008-09 DEO	2010 DEO	2011 DEO
	<p>2) Argentina: Impact Evaluations of the Development and Strengthening of Scientific, Technological and Innovation Capacities. The impact evaluations will be focused in three key areas of intervention: increasing innovation by Argentine firms and institutions, building human capital, and promoting research in the sciences. Pp. 152-153</p>	
PROTECTING THE ENVIRONMENT, RESPONDING TO CLIMATE CHANGE, PROMOTING RENEWABLE ENERGY, AND ENSURING FOOD SECURITY		
	<p>1) Dominican Republic: Agricultural Technology Transfer (PATCA). The evaluation will provide answers to questions about how to promote productivity improvements in agriculture. The PATCA evaluation will be based on random assignment and will explore desirable spillover effects. Pp. 166-168</p> <p>2) Mexico: Farmers Direct Support Program PROCAMPO. The evaluation will first identify the impact of PROCAMPO on productive investment and spending in agriculture and secondly it will identify how this can be translated into higher productivity and thus higher income and competitiveness. Pp. 168-171</p> <p>3) Nicaragua: Environmental Program for Disaster Risk and Climate Change Management. The evaluation will include an indication of whether the program improved farmers' well-being, enhanced resource conservation on farms and generally reduced the vulnerability of priority watersheds. Pp. 176-178</p>	<p>1) Dominican Republic: Promoting food health and safety. The program evaluation plan will focus primarily on the adoption of best practices among agricultural producers and the concomitant effects of implementing best practices, as well as the effects of the livestock health component of the program. Pp. 95-96.</p> <p>2) Paraguay: Land formalization. The program design includes an impact evaluation plan that uses an experimental approach to assess the effectiveness of land formalization in achieving its development objectives. Pp. 96-98.</p> <p>3) Uruguay: Promoting nautical tourism. The new tourism program includes a carefully designed ex-post impact evaluation strategy that should shed some light on the multiplier effects of nautical tourism and its impact on poverty. Pp. 98-99.</p> <p>4) Panama and Trinidad and Tobago: Climate change policy-based loans. At the end of the project, the team will evaluate whether changes in government policies and regulations did lead to the improvement of coastal defenses.</p>

2008-09 DEO

2010 DEO

2011 DEO

Pp. 99-101.

THE IDB'S WORK IN LACS LESS DEVELOPED COUNTRIES

1) Guyana: The Second Low Income Settlement Program. The evaluation of the project, designed jointly with the government of Guyana, will focus on the level of occupancy achieved in existing sites vis-à-vis the level of infrastructure provided, the level of ownership of lots by low income beneficiaries, community participation and involvement, and the uptake of lots in new schemes. Pp. 151-152.

Note: This list excludes the evaluations that were classified as ongoing at the time of publication of the DEO.

Table 2. Information on timeline of the promised evaluations provided in the DEOs

DEO 2008/2009: Project	Dates to conduct the evaluation	Dates to release the results
SOCIAL POLICY FOR EQUITY AND PRODUCTIVITY		
México, Colombia, El Salvador, Ecuador: Social protection through conditional cash transfers.	No	No
INFRASTRUCTURE FOR COMPETITIVENESS AND SOCIAL WELFARE		
LAC: Provision of Sustainable, reliable and affordable Energy.	No	No
INSTITUTIONS FOR GROWTH AND SOCIAL WELFARE		
Brazil, Chile, Paraguay, Peru: Access to credit and Small and Medium Enterprise Productivity.	No	No
Trinidad and Tobago: Citizen Security.	No	No
COMPETITIVE REGIONAL AND GLOBAL INTERNATIONAL INTEGRATION		
Argentina, Paraguay, Uruguay: Fostering science and technology adoption.	No	No
THE IDB'S WORK IN LACS LESS DEVELOPED COUNTRIES		
Guyana: The Second Low Income Settlement Program.	No	No
2010 : Project	Dates to conduct the evaluation	Dates to release the results
SOCIAL POLICY FOR EQUITY AND PRODUCTIVITY		
Guatemala (SM2015): Improved Access and Quality of Health and Nutrition Services.	Yes	No
Honduras (SM2015): Program to Strengthen Decentralized Management and Supply of Health Services.	Yes	No
Colombia, Nicaragua, Jamaica and Guyana: The Bank's work on Early Childhood Development (ECD).	No	No

Bolivia: Evaluating the Employment Programs using a Random Encouragement Design.	No	No
LAC: Building the Capacity of State and Local Actors to Respond to Violence Against Women.	No	No
Colombia: <i>Sol y Luna (SyL)</i>, Curbing Teenage Pregnancy in Latin America.	No	No
Colombia: Music and Life: Music Schools for Youth.	No	No
LAC: Promoting Women's Political Empowerment (PROLEAD).	Yes	No
INFRASTRUCTURE FOR COMPETITIVENESS AND SOCIAL WELFARE		
Paraguay. Water and Sanitation, Measuring their Impact on the Health of Rural Populations.	No	No
El Salvador: Learning from Rural Roads for Development.	No	No
INSTITUTIONS FOR GROWTH AND SOCIAL WELFARE		
Colombia: Evaluating the Impact of Credit.	Yes	Yes
Colombia: Empresas Públicas de Medellín.	No	No
El Salvador, Belize, Costa Rica: Reducing crime in urban areas in El Salvador, Community Action for Public Safety Program in Belize, Reducing Violent Crime in Costa Rica.	No	No
COMPETITIVE REGIONAL AND GLOBAL INTERNATIONAL INTREGRATION		
Colombia: Development of Human Capital and the Colombian System of Science, Technology and Innovation.	No	No
Argentina: Impact Evaluations of the Development and Strengthening of Scientific, Technological and Innovation Capacities.	No	No

PROTECTING THE ENVIRONMENT, RESPONDING TO CLIMATE CHANGE, PROMOTING RENEWABLE ENERGY, AND ENSURING FOOD SECURITY

Dominican Republic: Agricultural Technology Transfer (PATCA).	Yes	No
Mexico: Farmers Direct Support Program PROCAMPO.	Yes	No
Nicaragua: Environmental Program for Disaster Risk and Climate Change Management.	Yes	No

2011: Project	Dates to conduct the evaluation	Dates to release the results
SOCIAL POLICY FOR EQUITY AND PRODUCTIVITY		

Mexico: Water, Sanitation, and Hygiene in Basic Education (PASHEB).	No	No
Haiti: lessons from Katrina for a new and improved Haitian education system.	No	No
El Salvador (SM2015): Using results-based financing to impact the health outcomes of the poorest.	No	No
Nicaragua (SM2015) Are Commitments Stronger than Cash?	No	No
Improving efficiency and effectiveness of Social Safety Nets in Trinidad and Tobago	No	No
El Salvador: <i>Ciudad Mujer</i>, Empowering Women to Promote Development.	No	No
Dominican Republic: How Evaluation Findings on Socio-emotional Skills have Improved the Youth Training Program.	No	No

INFRASTRUCTURE FOR COMPETITIVENESS AND SOCIAL WELFARE

Bolivia: Water and Sanitation for Small Towns and Rural Communities.	No	No
-----------------------------------------------------------------------------	----	----

INSTITUTIONS FOR GROWTH AND SOCIAL WELFARE

Ecuador: Global Credit for Microfinance Support Program (GMSP).	No	No
------------------------------------------------------------------------	----	----

COMPETITIVE REGIONAL AND GLOBAL INTERNATIONAL INTEGRATION

Uruguay: Support to global export services.	No	No
----------------------------------------------------	----	----

PROTECTING THE ENVIRONMENT, RESPONDING TO CLIMATE CHANGE, PROMOTING RENEWABLE ENERGY, AND ENSURING FOOD SECURITY

Dominican Republic: Promoting food health and safety.	No	No
--------------------------------------------------------------	----	----

Paraguay: Land formalization.	No	No
--------------------------------------	----	----

Uruguay: Promoting nautical tourism.	No	No
---------------------------------------------	----	----

Panama and Trinidad and Tobago: Climate change policy-based loans.	No	No
---------------------------------------------------------------------------	----	----

ANNEX D. COMPARISON BETWEEN THE EFFECTIVENESS OVERVIEW OF DIFFERENT DEVELOPMENT BANKS

Table 1.D. The Effectiveness Overview Report of Different Development Banks

	ADB - DEfR	AfDB - ADER	WB - WB4R	WB - RAP
Annual Publication.	The DEfR is being published annually since 2008.	The African Development Bank published the first edition of the ADER in 2011.	The World Bank published the first edition of the WB4R in 2011, which constitutes a companion for the new World Bank Corporate Scorecard.	In 2010, the Independent Evaluation Group (IEG) introduced an integrated annual report (RAP) to present the results and performance for the three World Bank Group institutions (IBRD/IDA, IFC and MIGA).*
Achievement of development results.	In its 2011 edition, The DEfR reports on ADB's contribution development outcomes for completed sovereign (76) and non-sovereign (14) projects approved between 1997-2010. It breaks down the results by core sectors, other sectors, countries, and individual projects by country. It presents a brief analysis on the main trends of these ratings over the last years.	The ADER 2011 reported preliminary results on the outputs for the Core Sector indicators as the evaluation tools were still in an early stage of implementation. The 2012 edition included outputs from all closed investment operations for which a Project Completion Report was produced in the 2009-2011 period, excluding Budget Support Operations and studies, with a total of 159 Public Sector Operations and 41 Private Sector Operations.	The WB4R reports the development outcome ratings from the Independent Evaluation Group of the World Bank for IBRD and IDA operations. It presents a brief discussion on the trend of the ratings.	The RAP reports on the effectiveness to achieve development outcomes of the World Bank, IFC and MIGA interventions. The ratings correspond to WB and IFC interventions completed and reviewed by IEG in 2008-11, and from 2009 to 2011 for MIGA operations. It includes a discussion on the factors that have impact on these ratings.
Institutional priorities.	The DEfR 2011 includes an analysis of the compliance of the financing Strategy 2020 Priorities. It also reports on the compliance of the operations with the five core specializations that best support	The ADER presents a compendium of the work of the AfDB in the nine priority areas of the Bank and the expected results of such interventions in the region. It presents information on the actual	The 2011 WB4R provides data on results supported by the WB spanning the development agenda for countries where the Bank is active in six areas of Bank engagement. It also reports on	The WB introduced four core development goals as the conceptual framework for the evaluation, mapping activities and associating rating to the proposed structure. It is recognized

	ADB - DEfR	AfDB - ADER	WB - WB4R	WB - RAP
	its agenda.	values of output compared to the Medium-Term Strategy 2008-2012, although it is not exhaustive in the analysis of lending rates or particular interventions in these areas.	the track of the countries to achieve the Development Millennium Goals by region by 2015.	that lending amounts in Bank sectors and associated rating to this structure may not match with the standard thematic and sector themes in the Bank's database.
Evaluability and quality at entry.	The quality at entry of the projects is reported by country partnership strategies, sovereign projects and non-sovereign projects. This exercise is conducted biennial and it includes an analysis of the improvements in this area.	In 2011, the AfDB presented the methodology of the new quality-at-entry assessment tool, providing preliminary results for a selected sample of projects. In the 2012 edition, the report included rates for 769 operations approved until the end of 2011 (active portfolio). However, AfDB 2012 is not exhaustive in explaining changes in the scores, as this tool has been implemented recently.	The results of Quality Assessments of the Lending Portfolios conducted by IEG in fiscal years 2009 and 2011 are mentioned. These reviews analyze the quality of design for a sample of operations at the midpoint of their implementation. Although the report mentions that there has been an improvement of such ratings, it is not extensive on the analysis of this topic.	The RAP reports quality at entry ratings for the three World Bank Group institutions. The report is exhaustive in explaining the trends of quality at entry rating across sector and regions, providing factors accounting for changes in quality at entry ratings.
Economic Analysis	The DEfR does not provide information or analysis on the share of projects that include an ex-ante analysis.	The ADER does not provide information or analysis on the share of projects that include an ex-ante analysis.	The WB4R does not provide information or analysis on the share of projects that include an ex-ante analysis.	The RAP includes a short analysis of the trend of ERR of Bank-supported projects in past decades. It also presents the main figures of the ERR for the projects the ERR has been calculated.
Impact Evaluations.	As it is structured, the DEfR does not include the analysis or results of ex-post impact evaluations for past or current ADB interventions.	The ADER does not include analysis or results of ex-post evaluation for past or current AfDB interventions.	The WB4R illustrates its analysis with some examples of impact evaluations. However, there is not selection criteria for the examples mentioned.	The RAP presents the main findings from the ex post reviews of selected projects of the three World Bank institutions. The structure of the RAP does not include an

	ADB - DEfR	AfDB - ADER	WB - WB4R	WB - RAP
				analysis of impact evaluations, as these are published in separate reports by the IEG.
Recommendations of the Independent Evaluation Unit.	The DEfR provides a list with the findings and recommendations of the implementation working group, which is accompanied with a report of the priority actions to be taken by the ADB in the following year..	The AfDB effectiveness review includes a list of actions points to take by the Bank in the future. However, these actions do not constitute findings or recommendations from the Evaluation Group of the Bank, but the result of an internal consultation process with the stakeholder departments that have participation in the Report.	The WB4R follow up on specific recommendations from the EIG. Particularly, it reports on the actions taken to strength the results framework for Economic and sector work and Technical Assistance.	The IEG tracks its own recommendations and the actions taken by Management. RAP reports on the share of adopted recommendations by the three World Bank Group institutions. It includes the specific recommendation that each of the institutions have followed over the past years and the remaining challenges.
Progress on Results Framework targets.	The DEfR reports on the progress to achieve ADB targets on outputs, financing for strategy 2020 priorities, quality at entrance, portfolio performance, finance transfer and mobilization, and business processes and practices.	Progress in specific indicators of the strategic sectors and how well AfDB manages its operations in five areas: portfolio performance, quality at entry, Paris declaration indicators of effective aid, knowledge management, gender mainstreaming and climate change.	The WB4R reports on the progress of the WB following the results of the Paris Declaration Surveys in different aspects such as Aid Effectiveness, Procurement Systems and Fiscal Management Systems.	The structure of the RAP does not include the report on progress of targets established by the World Bank institution in any dimension.

**In contrast with the rest of Development Agencies, the Results and Performance Report of the World Bank is conducted by the Evaluation Group of the Institution. RAP includes a section with the Management comments.*

Management Comments

• • • • • • • • • •

Mid-Term Evaluation of IDB-9 Commitments
Background Paper: Development Effectiveness Framework and
Development Effectiveness Overview
Management Response

I. INTRODUCTION

- 1.1 Management welcomes the background paper on the Development Effectiveness Framework (DEF) and the Development Effectiveness Overview (DEO) as part of OVE's mid-term evaluation of the IDB-9 commitments and would like to thank OVE for the constructive dialogue during its preparation. The paper's suggestions will contribute to the Bank's efforts to more effectively implement the DEF instruments and report on the Bank's effectiveness.
- 1.2 Management provided detailed comments to OVE on an earlier draft of this paper and is pleased to see that many of its suggestions were incorporated in the final version.

II. OVERALL FINDINGS AND SUGGESTIONS

- 2.1 Management agrees overall with the analysis and suggestions for the DEF. Management considers that OVE's comments about the Development Effectiveness Matrix (DEM), Operations Policy Committee (OPC), Project Monitoring Report (PMR) and the Project Completion Report (PCR) are relevant, concrete and constructive and that they will contribute to refining the DEF framework.
- 2.2 As the paper explains, the DEM for Non Sovereign Guaranteed operations (NSG DEM) was initially designed to be a development results assessment tool. Reflecting the logic of private sector operations and best practices among multilateral development banks (MDBs), the NSG DEM was based on the Good Practice Standards of the Evaluation Cooperation Group (ECG-GPS) and the stakeholder framework for private sector operations and therefore did not incorporate all evaluability elements. It is worth noting that NSG operations applied the DEM before it was applied to Sovereign Guaranteed (SG) loan operations.
- 2.3 OVE's assessment of the OPC is based on the attention given to evaluability as reflected in the minutes of the OPC meetings. While these minutes contain valuable information, they only concern those projects that are required to go to OPC for consideration. An analysis of written comments by the EVP's office or other members of the OPC during the non-objection process would provide a more complete view. Additionally, the paper notes that the number of discussions on evaluability decreased after the approval of the five point threshold for the DEM score. Management considers this to be consistent with the increase in overall evaluability levels in this same period since, as project evaluability improves, it can be expected that high-level discussions on the subject would decrease.
- 2.4 With respect to PCRs, Management would like to highlight that in addition to the results of a survey of Bank staff conducted by OVE in 2011, the paper relied heavily on OVE's previous evaluation of PCRs (on which Management has commented separately) and

therefore the data used to assess the quality of the PCR system is predominantly based on the assessment of a sample of PCRs approved in 2008-2009 and the review of 79 PCRs approved in 2010. As the paper recognizes, PCRs approved in 2008 and 2009 had not benefited from the DEM revision process at project preparation. Management would like to add that neither did these PCRs benefit from (i) efforts to promote project impact evaluations; (ii) the requirement that projects include calculations of ERR, or (iii) efforts to strengthen the PMR process. Consequently, Management believes that the analysis of the PCR system could have benefited from having assessed PCRs approved after IDB-9.

- 2.5 With respect to the DEO, while Management acknowledges that, to date, reporting on effectiveness in the DEO has been restricted by the incomplete implementation of the DEF tools, Management also believes that a single report should not be expected to provide an exhaustive and in-depth analysis of every aspect of effectiveness. As an overview report, the DEO should be complementary to other tools for accountability and provide information in an accessible format. In this respect, Management differs in its interpretation of the IDB-9 mandates regarding the scope of the DEO.
- 2.6 Similarly, the selection of projects discussed in the DEO was never meant to be statistically representative of all Bank projects, sectors or priorities, but rather illustrative of the work the Bank does and its progress towards being a more results-based institution. The first part of the DEO already includes some statistics, which will be supplemented with data from the evaluation system and the PCRs as these begin to generate reliable development effectiveness data as was conceived in the DEF.

III. LOOKING FORWARD

- 3.1 Much of the Bank's work in further implementing the DEF instruments and improving reporting on the Bank's effectiveness is already underway and is consistent with the suggestions presented in the paper. Going forward, Management will give careful consideration to OVE's other suggestions:
- i. Recognizing the Bank's commitment to improve the results measurement framework, and based on OVE's recommendations to improve project evaluability for NSG projects, Management is working to revise the NSG DEM, including the new evaluability tool and Results Matrix;
 - ii. Management would like to consider the practical application of OVE's suggestion to include evaluation plans in loan contracts as each time adjustments to a project's evaluation plan are needed, the contract would have to be amended;
 - iii. On the incorporation of OVE's recommendations, Management considers that a good tracking system will support an effective incorporation of lessons learned. Such a tracking system is being developed jointly by OVE and SPD and should be fully operating in 2013;
 - iv. Management expects that the DEF will be entirely implemented by mid-2014, at which point the DEO will be able to report more fully on the Bank's effectiveness.

ANNEX I: FACTUAL CORRECTIONS

- 1.1 It should be clarified that the PMR and the PCR are only relevant for SG operations. NSG operations use the PSR (Project Supervision Report) for the annual monitoring report and the XPSR (Expanded Project Supervision Report) for more extensive self-evaluation report.
- 1.2 Table I: Following the modification of the DEM required by the Board in 2010, the DEM score takes into account only the requirements of the evaluability section. Sections on alignment and additionality are included in the summary, but are not considered for the score.
- 1.3 Paragraph 2.9: The NSG DEM does not limit the number of indicators to be tracked. Project teams are encouraged to use at least four indicators, but are free to include more indicators as applicable.