



Project Evaluation

IDB and IIC Project Performance: OVE's Review of 2016 Project Completion Reports and Expanded Supervision Reports

IDB and IIC Project Performance: OVE's Review of 2016 Project Completion Reports and Expanded Supervision Reports

Office of Evaluation and Oversight (OVE)

October 2017





This work is distributed under a Creative Commons license <https://creativecommons.org/licenses/by-nc-nd/3.0/us> (CC BY-NC-ND 3.0 US). You are free to share, copy and redistribute the material in any medium or format, Under the following terms:



Attribution - You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.



Non-Commercial - You may not use the material for commercial purposes.



No Derivatives - If you remix, transform, or build upon the material, you may not distribute the modified material.

No additional restrictions - You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

The link provided above includes additional terms and conditions of the license.

© **Inter-American Development Bank, 2017**

Office of Evaluation and Oversight
1350 New York Avenue, N.W.
Washington, D.C. 20577
www.iadb.org/evaluation

RE-520
October 2017

ACRONYMS AND ABBREVIATIONS

ACKNOWLEDGEMENTS

EXECUTIVE SUMMARY

1. INTRODUCTION	1
2. PROJECT EVALUATION FRAMEWORK AND REVIEW PROCESS	5
A. Project evaluation framework.....	5
B. Projects reviewed.....	8
C. The project review process	11
3. RESULTS OF PUBLIC SECTOR PROJECT REVIEWS	13
A. Overall outcomes	13
B. Core criteria.....	14
1. Relevance.....	14
2. Effectiveness	15
3. Efficiency.....	17
4. Sustainability	17
C. Quality of PCRs	18
4. RESULTS OF PRIVATE SECTOR PROJECT REVIEWS	21
A. Overall outcomes	21
B. Core criteria.....	22
1. Relevance.....	22
2. Effectiveness	23
3. Efficiency.....	25
4. Sustainability	26
C. Non-core criteria.....	27
D. Comparison between OVE ratings and Management self-ratings.....	28
E. Quality of XSRs.....	29
5. CONCLUSIONS AND RECOMMENDATIONS.	31

NOTES

Available at www.iadb.org/evaluation

ANNEX I List of NSG operations reviewed: Removed to exclude confidential information in accordance with the “Information Relating to Non-Sovereign Guaranteed Operations” disclosure exception in paragraph 4.1 j. of the Bank’s Access to Information Policy.

ANNEX II [List of SG operations reviewed](#)

ANNEX III [OVE ratings of SG operations](#)

ANNEX IV [Overview of OVE’s assessment of 2016 PCRs](#)

ANNEX V [XSR guidelines for IDBG private sector operations for pilot](#)

ANNEX VI [OVE PCR reviews](#)

ACRONYMS AND ABBREVIATIONS

DEF	Development Effectiveness Framework
DELTA	Development Effectiveness Learning, Tracking, and Assessment Tool
DEM	Development Effectiveness Matrix
ECG	Evaluation Cooperation Group
EOM	Early operating maturity
FI	Financial intermediary
GPS	Good Practice Standards
IDB(G)	Inter-American Development Bank (Group)
IICD	Inter-American Investment Corporation
MDB	Multilateral development bank
NSG	Non-sovereign-guaranteed
OMJ	Opportunity for the Majority
OVE	Office of Evaluation and Oversight
PBP	Programmatic policy-based program
PCR	Project Completion Report
SCF	Structured and Corporate Finance
SG	Sovereign-guaranteed
XSR	Expanded Supervision Report

This document and the supporting project specific reviews were prepared by a team composed of Monika Huppi (team leader), Pablo Alonso, Nayda Avalos, Oliver Azuara, Ernesto Cuestas, Coral Fernandez, María Camila García, Verónica González-Diez, Maya Jansson, Ana María Linares, Melanie Putic, Roland Michelitsch, Juan Felipe Murcia, Johanan Rivera, Jonathan Rose, Patricia Sadeghi, Rafael Seiwald, José Ignacio Sembler and Horst Wattenbach, under the direction of Cheryl Gray, OVE Director.



Review of management self-evaluations by the independent evaluation office is an integral aspect of the project evaluation system across MDBs.

© IDB

Executive Summary

Assessing the results of individual development operations is essential to learning what works and what doesn't and thus to increasing the effectiveness of investments in development. Virtually all international development organizations have thus put in place systems to measure the effectiveness of their support. At the Inter-American Development Bank Group (IDBG), the project evaluation systems for both sovereign guaranteed (SG) and non-sovereign-guaranteed (NSG) operations have three building blocks: (i) an ex-ante assessment of the project's expected ability to report on results at completion; (ii) project monitoring/supervision reports tracking implementation progress and emerging results; and (iii) a management self-evaluation of the results of SG operations at completion (Project Completion Report, or PCR) or of the results of NSG operations when they reach "early operating maturity" (Expanded Supervision Report, or XSR).

Review of management self-evaluations by the independent evaluation office is an integral aspect of the project evaluation system across MDBs. The purpose of this independent review is to ensure the accuracy and credibility of the project performance reporting system. IDBG's partner MDBs use the final validated ratings of the independent evaluation office as the final record of MDB project performance for corporate reporting. IDB has been slow, relative to other MDBs, in developing a credible project self-evaluation and validation system, but for the past eight years IDB management and OVE have sought to develop one, reinforced by mandates included in the Ninth General Capital Increase (IDB-9) in 2010. The IDB-9 agreement stipulates that OVE will validate the results achieved in completed projects, that the latter will be evaluated against the results expected at project approval, and that the performance of completed projects will serve as an indicator in IDB's corporate results framework. On the IIC side, OVE has been validating self-evaluations since 2001.

Evaluation criteria for public and private sector projects financed by MDBs developed separately and were not comparable or consistent with one another. IDBG has been at the forefront of MDBs in bringing those criteria closer together for greater consistency and comparability. Both the IDB and IIC project evaluation methods were recently revised around a common objectives-based evaluation framework. Rather than evaluating project results in the abstract, the objectives-based methodology measures project performance against the project development objectives that were formulated at approval, assessing how relevant these objectives were, to what extent they were achieved, how efficiently project resources were used to achieve them, and how sustainable the achieved results are. A performance rating is assigned to each of these four core criteria, and an overall project outcome rating is derived from them. For NSG operations, additional dimensions – additionality, work quality, and investment outcome – are also evaluated and scored.

2016 was the first year in which both IDB and IIC used these revised frameworks to report on the results achieved by their operations and OVE validated them. This report summarizes the results and findings of OVE’s review of the 30 XSRs completed by IIC in 2016 for NSG operations that reached early operating maturity in 2015, and for the 21 PCRs that IDB completed under the 2014 PCR guidelines by the end of 2016. Given the changes in methodology, the project performance ratings for the cohort of projects evaluated in 2016 are not comparable to those of projects evaluated in prior years.

Seventeen of the 21 SG operations reviewed achieved an overall positive outcome rating. This indicates quite solid performance for this set of SG operations, though because they are only a small subset of the Bank’s SG operations that closed between 2014-2016, no conclusions can be drawn from these operations about the overall performance of IDB’s portfolio of operations that closed over that period or about performance across sectors or countries. The 21 SG operations scored best on relevance and lowest on effectiveness. The high share of projects with a positive relevance rating (20 of 21 operations) indicates that the objectives of the reviewed operations were well aligned with the development challenges in the country and with country realities. About half of the 21 operations demonstrably achieved the majority of the objectives they set out to achieve at approval. Lack of information on results was a factor contributing to a less-than-Satisfactory effectiveness rating for four projects, though in each case it was combined with failure to make substantial progress on some of the objectives for which information on achievement was available. In over one-third of the operations there is a risk that the results will not be sustained.

Twelve of the 30 NSG operations reviewed achieved a positive overall outcome rating, with IIC-approved operations faring somewhat better than those approved by SCF and OMJ. The overall outcome ratings reflect low scores on effectiveness and efficiency. These operations were designed prior to the introduction of the DELTA, and hence less attention was paid to ensuring that clear development objectives were formulated and indicators identified to measure achievement of such objectives. But this only partly explains the relatively low share of operations with a positive outcome rating. Almost

two-thirds of the reviewed operations did not demonstrably achieve the majority of the development objectives they set out to achieve. For 14 of the 19 operations with negative effectiveness ratings, the latter was due to underperformance compared to targets set at approval. In another four projects, the low effectiveness score was due to a combination of low performance on some objectives and lack of information to assess the achievement of others and in one it was primarily due to lack of information. In half of the operations reviewed, sustainability of results was less than satisfactory, while somewhat over one-quarter achieved a positive efficiency rating.

The vast majority of NSG operations reviewed performed well on investment outcome. The high share of operations receiving Satisfactory or higher ratings on investment outcome (26 of 30 operations) suggests that the IDBG did well in structuring and funding operations that contributed positively to its financial success. Yet only 10 of the 26 operations that achieved a positive investment outcome rating also substantially achieved their development objectives, and only seven of them achieved a positive efficiency rating.

About half of the reviewed NSG operations demonstrated additionality. The share of operations with satisfactory financial additionality was somewhat lower than that with nonfinancial additionality. Financial additionality was particularly questionable for FI operations. Six of 16 FI operations provided clear financial additionality, reflecting a tendency to address constrained access to finance with additional liquidity, even if a lack of liquidity was not shown to have been the constraining factor. The quality of IDBG's work on the reviewed operations was modest, particularly at the time of structuring and appraisal, with 12 of the 30 EOM 2015 operations achieving a Satisfactory score on work quality.

Overall, significant progress has been made by the IDBG in establishing a credible and consistent objectives-based self-evaluation system, though further efforts are needed to consolidate these gains. On the NSG side, continued efforts are needed to identify the development objectives and results that each operation seeks to achieve up front, as opposed to merely focusing on business objectives. Results frameworks need to be clearly aligned with the stated objectives to facilitate measuring progress toward achieving these objectives. The new DELTA should help in this regard. The NSG pilot XSR guidelines have proven to be a good basis for the preparation of project self-evaluations and OVE validations, though this exercise has pointed to some areas for further refinement. In general, the XSRs did well in distilling lessons that can help inform design of future operations.

On the SG side, OVE found that PCRs are generally of good quality and candor and distill a range of important lessons that can inform the design of future projects, though there are still challenges that need to be addressed. Stronger efforts are needed to ensure that PCRs are in line with the objectives-based methodology and are consistently delivered on a timely basis. Management's current practice of permitting indicators

and targets to be changed up to the last Project Monitoring Report (or even during the PCR exercise) compromises its ability to assess how effectively an operation has indeed performed. Further clarity is also needed on what type of economic analysis to carry out to assess efficiency for various types of operations.

Validation of project self-evaluations by the independent evaluation office is a core component of a strong project evaluation architecture. A validation system operates most effectively and efficiently when the self-evaluation and the validation apply the same ratings methodology, and thus moving toward a unified rating system used by both management and OVE – particularly in the case of IDB where current differences are more marked – remains an important step in further strengthening the system. Validations may confirm or adjust management’s own project performance ratings, but in either case the validated ratings should be used by both management and the independent evaluation office as the final ratings for purposes of corporate reporting on portfolio performance, as is done in IDBG’s partner MDBs.

In light of these conclusions, OVE recommends the following:

For IIC management:

- (i) Work with OVE to update the pilot XSR guidelines to arrive at a set of guidelines that both parties will use when evaluating and validating project performance, ensuring consistency to the extent possible with IDB’s PCR guidelines. Drawing on the experience gained during this EOM 2015 exercise, issues to be clarified in the revised guidelines include:
 - the types of development objectives to be covered in the effectiveness section (excluding financial objectives which should be covered under the financial performance part of efficiency);
 - how ratings on individual objectives are determined and aggregated into an overall effectiveness rating, ensuring consistency with the SG methodology;
 - how to assess and rate financial and economic performance for FI operations when information on the performance of the relevant sub-portfolio is not available;
 - the factors to consider (in addition to safeguards performance) in assessing sustainability;
 - the weighting and rating scale to use in deriving the overall project outcome rating, ensuring consistency with SG methodology.
- (ii) Use OVE’s final validated ratings for purposes of corporate reporting on portfolio performance.

- (iii) Work with OVE to further define the process and timetable for the annual preparation of XSRs by management and their validation by OVE.

For IDB management:

- (i) Work with OVE to arrive at a set of PCR guidelines that both parties will use when evaluating and validating project performance, ensuring consistency to the extent possible with IIC's XSR guidelines. Specific aspects of the 2016 PCR guidelines that need to be addressed include:
 - the definition of the outcome indicators and targets against which project outcomes are to be assessed. To be consistent with IDB-9 commitments and current XSR guidelines, project outcomes should be assessed against outcome indicators and targets established at approval, which can be changed if an operation is formally restructured with Board approval. If there are delays in an operation reaching eligibility, OVE recommends that updates also be allowed within 60 days of an operation reaching eligibility;
 - the way that the effectiveness of PBP series is assessed. OVE recommends that the program be assessed against end of program targets in each matrix of the series and the overall effectiveness rating for the series be an average of those individual assessments;
 - the type and content of efficiency analysis required under various circumstances;
 - how to include additionality and Bank performance as evaluation criteria, consistent with XSR guidelines;
 - the deadline for completing PCRs. OVE recommends that completion be mandatory within nine months after an operation reaches closed status in Convergence.¹
- (ii) Apply the revised project rating system to all future PCRs that have not yet been shared with Governments for comments, irrespective of whether they are prepared under the 2014 guidelines or later ones. Even PCRs prepared under the 2014 guidelines provide information to do so.
- (iii) Use OVE's final validated ratings for purposes of corporate reporting on portfolio performance.
- (iv) Work with OVE to further define the process and timetable for the annual preparation of PCRs by management and their validation by OVE.

¹ The closed (CO) status is activated when the disbursed resources have been fully justified, unjustified resources have been returned, and undisbursed amounts have been cancelled.



This report summarizes the results and findings of OVE's review of XSRs completed by IIC in 2016 for NSG operations that reached early operating maturity in 2015 and OVE's review of PCRs that IDB completed by the end of 2016 using the revised PCR guidelines of 2014.

1 Introduction

Assessing the results of individual development operations is essential to learning what works and what doesn't and thus to increasing the effectiveness of investments in development. Therefore, international development organizations, including the multilateral development banks (MDBs), have put in place instruments to measure the effectiveness of their investments. The Inter-American Development Bank (IDB) launched its Development Effectiveness Framework (DEF) in 2008, aiming to generate a body of knowledge about what works in meeting the region's development challenges. The commitment to enhance development effectiveness through reporting on the results of Bank operations was reinforced under IDB's ninth capital increase (IDB-9),¹ when the institution pledged to "improv[e] the results measurement framework to identify for every project whether specific and tangible results have been achieved" (CS-3868-1). Until the IDBG's recent merge-out, the DEF applied to sovereign-guaranteed (SG) operations as well as SCF and OMJ operations.

For the Inter-American Investment Corporation (IIC), the more systematic tracking of development results dates back to 2001 (CII/RE-1). The IIC had established a project evaluation system and function in 1999 (CII/GN-141, CII/GN-141-2) and contracted the Office of Evaluation and Oversight (OVE) to support this effort in 2000, and in 2001 it started to self-evaluate its operations following the Good Practice Standards for Private Sector Operations issued by the Evaluation Cooperation Group (ECG). Since the 2016 merge-out of the IDBG's private sector

operations into the IIC, the Development Effectiveness Learning, Tracking, and Assessment tool (DELTA) (CII/PP-169, CII/PP-180) has replaced the two systems that were previously used in SCF/OMJ (DEM) and IIC (Development Impact Assessment Scorecard). Thus, DELTA has become the common system for ex-ante scoring of development results and additionality of operations within the broader framework of the IIC’s “Renewed Vision” for private sector operations (CII/GN-2807-2).

The project performance monitoring systems for SG and non-sovereign-guaranteed (NSG) operations consist of three similar building blocks: (i) an ex-ante assessment of the project’s expected ability to report on results at completion (DEM for SG operations, DELTA for NSG operations); (ii) project monitoring/supervision reports tracking implementation progress and emerging results (Project Monitoring Reports prepared twice yearly for SG operations and Project Supervision Reports prepared annually for NSG operations); and finally, (iii) a management self-evaluation of the results of SG operations at completion (Project Completion Report, or PCR) or of the results of NSG operations when they reach “early operating maturity” (Expanded Supervision Report, or XSR).² PCRs and XSRs are reviewed by OVE, which assigns a project outcome rating to each operation.

Review of project performance ratings by the independent evaluation office is an integral aspect of the project evaluation system across MDBs. IDBG’s partner MDBs – including the World Bank Group, the Asian Development Bank, the African Development Bank, the European Bank for Reconstruction and Development, and others – all have their independent evaluation office review project self-evaluations to validate or, if applicable, override project outcome ratings. The performance ratings reported by these MDBs – and included, where applicable, as indicators in their corporate results frameworks – are the validated ratings emerging from this process. The purpose of this independent review and validation is to ensure the accuracy and credibility of the project performance reporting system, and the approach is enshrined in the Good Practice Standards (GPS) for private and public project evaluations.

While OVE has been validating IIC’s project self-evaluations since 2001 and reporting results to the Board, IDB did not have a similar system in place prior to 2008. But for the past eight years IDB management and OVE have sought to develop one, reinforced by IDB-9 mandates introduced in 2010. IDB’s commitments under IDB-9 stipulate that “OVE will validate achieved results in completed projects” (paragraph 4.7) and includes as an indicator in the IDB-9 Results Framework “the percent of projects with satisfactory rating on development results at completion” (Annex I, indicator 4.2.4). OVE’s validations follow the IDB-9 mandate that completion reports compare actual results achieved with the results expected at project approval (para 6.6). OVE’s validations are based primarily on a review of

the evidence presented in the PCRs for SG operations and in the XSRs for NSG operations, in project supervision reports, and in the project documentation that was approved by the Board.

This report summarizes the results and findings of OVE's review of XSRs completed by IIC in 2016 for NSG operations that reached early operating maturity (EOM) in 2015 and OVE's review of PCRs that IDB completed by the end of 2016 using the revised PCR guidelines of 2014.



2016 was the first year in which the IDBG used a common, objectives-based evaluation framework to report on the results achieved by its SG and NSG operations.

© IDB

2 Project Evaluation Framework and Review Process

A. PROJECT EVALUATION FRAMEWORK

In 2014, IDB launched and piloted new guidelines for the preparation of PCRs. The guidelines built on OECD-DAC criteria and the Good Practice Standards (GPS) for project evaluation issued by the ECG. The system was rolled out in 2015 with the requirement that all SG operations with a DEM (i.e., those approved from 2009 onwards) and that had fully disbursed from August 5, 2014 onwards (the approval date of the guidelines) prepare the PCR using the revised methodology. In 2015, OVE assessed and reported on the PCRs completed during the 2014 pilot (RE-488) and recommended revisions to the system to bring it more in line with the GPS as well as with IDBG's system for assessing NSG operations. IDB updated its PCR guidelines in late 2016 (OP-1242-5) to reflect some of the OVE recommendations. Similarly, in 2015 new pilot guidelines were introduced for the preparation and validation of XSRs for NSG operations. These guidelines for the first time also use an objectives-based methodology, which was piloted on a sample of projects in 2015 and rolled out more broadly in 2016, when it was applied to all XSRs prepared during that year. Given the changes in methodology, the project performance ratings for the cohort of projects evaluated in 2016 are not comparable to those of projects evaluated in prior years.

The objectives-based evaluation methodology takes as the starting point the stated development objectives that an operation aimed to achieve. Thus, rather than evaluating project results in the abstract, the objectives-based methodology measures project performance against the project development objectives that were formulated at approval, assessing how relevant these objectives were, to what extent they were achieved, how efficiently project resources were used to achieve them, and how sustainable the achieved results are. A performance rating is assigned to each of these four core criteria, and an overall project outcome rating is derived from them (Table 2.1). While the overall framework to arrive at a project outcome rating for SG and NSG operations is thus similar, its application differs somewhat, partly because of the different nature of SG and NSG operations, and partly because of still-persisting differences in the guidelines unrelated to characteristics of SG and NSG lending.

TABLE 2.1. CORE PROJECT PERFORMANCE ASSESSMENT CRITERIA

SG operations	CORE CRITERIA	NSG operations
<ul style="list-style-type: none"> • Alignment of project development objectives with country development needs • Alignment with IDBG country strategy • Alignment of project design with country realities • Alignment of project design with project development objectives 	Relevance	<ul style="list-style-type: none"> • Alignment of project objectives with country development needs • Alignment with IDBG country strategy and corporate goals • Alignment of project design with country realities • Alignment of project design with project development objectives
<ul style="list-style-type: none"> • Extent to which project achieved each stated development objective, given project outputs produced 	Effectiveness	<ul style="list-style-type: none"> • Extent to which project achieved each stated development objective, given project outputs produced
<ul style="list-style-type: none"> • Extent to which project benefits exceed project costs or extent to which project benefits were achieved at less than expected or at reasonable costs 	Efficiency	<ul style="list-style-type: none"> • Financial performance: Project contribution to company financial results and extent to which project process and business objectives were achieved • Economic performance: Extent to which project economic benefits exceed costs of capital; project effects on key economic stakeholders
<ul style="list-style-type: none"> • Extent to which environmental, social, technical, financial, economic, and political risks to continuation of project development results have been mitigated and safeguards performance has been satisfactory 	Sustainability	<ul style="list-style-type: none"> • Safeguards performance • Extent of unmitigated risks to continuation of project results
 <p>Overall Project Outcome</p>		

Source: OVE

2016 was the first year in which the IDBG used a common, objectives-based evaluation framework to report on the results achieved by its SG and NSG operations. Over the years, the project evaluation systems for public and private sector operations financed by MDBs developed in parallel, with little comparability and commonality. Recognizing that the contribution to a host country's development is at the core of using MDB's

capital for both SG and NSG operations, several MDBs are currently working to better harmonize the evaluation frameworks for public and private sector operations. Having adopted an objectives-based evaluation framework for SG and NSG operations puts the IDBG at the forefront of multilateral institutions in this regard.

OVE used a four-point rating scale to assess the relevance, effectiveness, efficiency, and sustainability of each SG and NSG operation. It then calculated a weighted average of the four component scores to derive an overall project outcome rating. In line with the revised PCR guidelines of 2016, OVE assigned a weight of 20% each to the scores for relevance, efficiency, and sustainability and a weight of 40% to effectiveness. Because efficiency is not assessed for SG policy-based loans, the overall outcome rating is based on a weighted average of relevance (20%), effectiveness (60%), and sustainability (20%). In line with the ECG GPS, the resulting outcome rating was converted into a six-point scale ranging from Highly Successful to Highly Unsuccessful (see Annex IV). Given the new methodologies applied to evaluate and rate the performance of both, SG and NSG operations, ratings based on the new methodologies cannot be compared to project performance ratings in prior years³.

In addition to the four common core criteria, both the XSR and PCR guidelines stipulate a set of non-core criteria that the project self-evaluations are to cover (Table 2.2). While the non-core criteria of SG operations are not rated, those for NSG operations are rated on a four-point scale and validated by OVE. The non-core criteria for NSG operations include IIC's financial and non-financial additionality, investment outcome, and IIC's work quality during project screening/appraisal/ structuring and during project supervision. SG non-core criteria, which are discussed in the PCR qualitatively, include the contribution to IDB corporate objectives, contributions to country strategy, use of country systems, safeguards performance and project monitoring and evaluation. To the extent that safeguards performance affects the sustainability of project results, OVE considered this dimension in the sustainability rating. The DEF foresees that additionality should also be a criterion that is assessed for SG operations as it is for NSG operations, but thus far this dimension has not been introduced into the PCR framework.

Unlike on the NSG side, Bank performance during project preparation and implementation is not assessed in PCRs. This is not only a considerable deviation from the ECG GPS, but is also contrary to the DEF, which stipulates that Bank performance during the life of the project will be assessed.⁴ IDB is the only major MDB that does not assess and rate Bank performance as part of its project completion assessments. In addition, other multilaterals also rate borrower performance, as specified in the ECG GPS evaluation principles,⁵ an evaluation dimension that is absent from both IDB's and IIC's performance assessments. The lack of assessment of Bank and borrower performance is a notable shortcoming of IDB's current project assessment systems, as these dimensions can help shed significant light on factors within the control of the IDBG and the borrower that have affected project results.

TABLE 2.2. NON-CORE PROJECT EVALUATION CRITERIA

NSG operations		
Additionality	<ul style="list-style-type: none"> Financial additionality: assesses what IIC financial additionality brought to the project compared to what was available in the financial market Nonfinancial additionality: assesses what nonfinancial value proposition IIC brought to the project 	<ul style="list-style-type: none"> Rated on 4-point scale
Investment outcome	<ul style="list-style-type: none"> Assesses the investment's gross profit contribution to IIC 	<ul style="list-style-type: none"> Rated on 4-point scale
IIC work quality	<ul style="list-style-type: none"> Screening, appraisal, and structuring Project supervision and administration 	<ul style="list-style-type: none"> Rated on 4-point scale
SG operations		
Contribution to IDB corporate strategic goals	Qualitative discussion of contributions to (i) poverty reduction and equity, (ii) climate change, renewable energy, and environmental sustainability, and (iii) regional cooperation and integration	<ul style="list-style-type: none"> Not rated
Contribution to Country Strategy objectives	Qualitative assessment of project contributions to the strategic objectives set out in the Country Strategy	<ul style="list-style-type: none"> Not rated
Use of country systems	Project contribution to strengthening country systems	<ul style="list-style-type: none"> Not rated
Safeguards performance	Key safeguards issues and mitigation measures taken	<ul style="list-style-type: none"> Not rated, but safeguards results considered in sustainability rating
Monitoring & evaluation	M&E design, implementation, and utilization	<ul style="list-style-type: none"> Not rated

Source: OVE

B. PROJECTS REVIEWED

On the private sector side, OVE validated the XSRs for all 30 NSG operations that reached early operating maturity in 2015.⁶ Of these 30 operations, 11 had been approved by SCF, five by OMJ, and 14 by IIC. Just over half were financial intermediary (FI) operations, four were energy operations, three were in agriculture, one was in education, and six were in other areas, such as retail and packaging (Table 2.3). The reviewed NSG portfolio amounted to US\$672.8 million in approvals (Figure 2.1), of which 59% went to FI projects and 29% to energy sector operations. The validations covered 16 IDB member countries, eight of which had more than one completed operation (see Annex I for project list).

On the public sector side, OVE reviewed all 21 PCRs prepared under the 2014 guidelines that had been approved by the end of 2016. These included four operations that closed in 2014, 11 that closed in 2015, and six that closed in 2016. This is only a small share (less than 10%) of the 232 SG operations that closed between January 2014 and the end of the first quarter of 2016.⁷ Of the 232 closed operations, 168 were expected to prepare a PCR (as cancelled operations and operations of a programmatic

series that has not been completed do not require a PCR). According to SPD data, a PCR had to be prepared in accordance with the new methodology for only 43 of these, and of those only 17 had actually been delivered by the end of 2016.⁸ An additional four PCRs were delivered for operations that closed after the first quarter of 2016 (Figure 2.2). For operations approved before 2009, the preparation of the PCRs followed the old methodology, and OVE does not review such PCRs.

TABLE 2.3. SECTORAL DISTRIBUTION OF EOM 2015 NSG OPERATIONS

Sector	SCF	OMJ	IIC	Total
Energy	3		1	4
Financial Intermediary	7	4	5	16
Agriculture	1		2	3
Education		1		1
Other			6	6
Total	11	5	14	30

Source: OVE

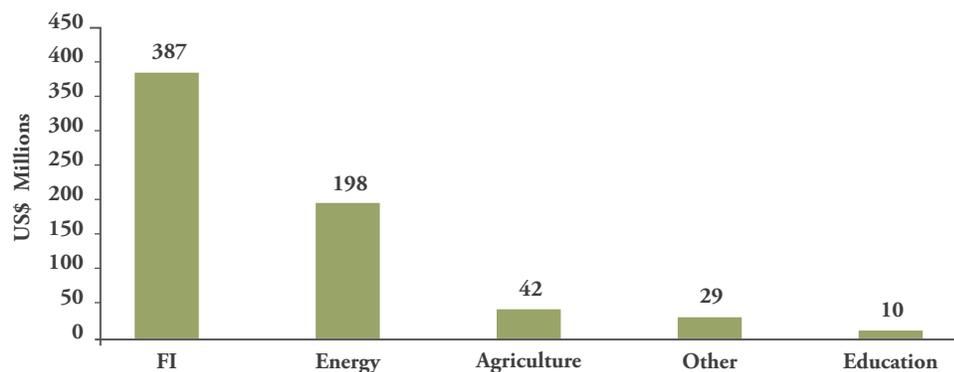


FIGURE 2.1
Approved amounts
of validated NSG
operations by sector

Source: OVE

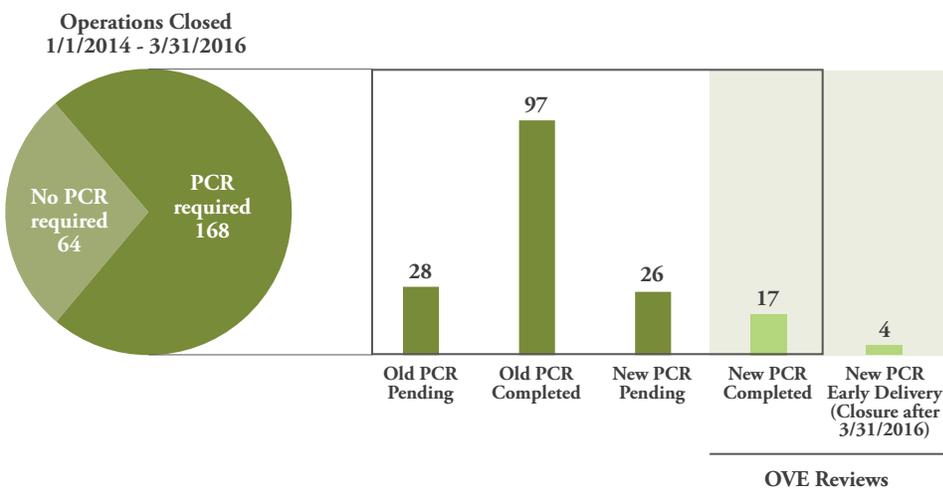


FIGURE 2.2
PCR requirements and
deliveries under old and
new guidelines

Source: OVE

The 21 PCRs covered projects in 11 countries, of which six had at least two operations. The sample included three programmatic policy-based series (PBPs), and the remaining 18 operations were a mix of investment operations (see Annex II). The operations covered 11 sectors and accounted for US\$1.67 billion in approved lending (Table 2.4). Seven operations had been approved each in 2009 and 2010 (during the early phases of the introduction of the DEM), three in 2011, and two each in 2012 and 2014.

TABLE 2.4. SECTORAL DISTRIBUTION OF SG OPERATIONS REVIEWED

Sector	Number of PCRs	Loan amount US\$ million
Competitiveness	1	110.0
Citizen Security	1	5.0
Education	2	387.0
Energy	3	158.5
Environment/Disaster Risk Mgmt	5	106.0
Financial Markets	1	20.0
Health	2	677.5
Public Sector	1	5.0
Science and Technology	1	100.0
Social Protection	2	81.0
Transport	1	17.5
Urban	1	8.4
Total	21	1,675.9

Note: Includes a three-operation PBP in disaster risk management, a two-operation PBP in health, and a three-operation PBP in competitiveness, each of which is covered in a single PCR for the entire program and counted as one PCR above.

Source: OVE

A key purpose of any project rating system is to provide management and other stakeholders a tool to monitor the performance of the portfolio over time, across sectors, and across different lending instruments, allowing the institution to undertake corrective action if performance deteriorates. However, if the system is to be an effective management and monitoring tool, it needs to cover the full portfolio or at least a statistically representative sample of all operations closing or reaching EOM in a given year. While this objective has been achieved for NSG operations, it has not yet been reached on the SG side. Over time, as the share of PCRs prepared according to the old PCR system decreases and new PCRs are prepared using the revised methodology (with some further modifications as outlined later in this report), the SG system has the potential to become a credible portfolio performance monitoring tool, provided PCRs are prepared consistently and on a timely basis.

C. THE PROJECT REVIEW PROCESS

OVE applied the same review process to private and public sector projects. An OVE evaluator was assigned to review each operation and prepare an assessment report. To ensure consistency across project assessments, each assessment was then reviewed by a panel consisting of OVE management and senior staff. The draft assessment reports with project-specific ratings were shared with management for review and comment. OVE then finalized the assessments and ratings, taking into consideration management's comments and additional information provided. After receiving OVE's final ratings, IIC went through another round of revisiting their self-ratings to further align them with OVE's ratings.

There is an important difference between OVE's review of SG and NSG self-evaluations this year. For NSG operations, OVE validated management's ratings in accordance with the jointly agreed pilot XSR guidelines, thus using the same approach as management used in rating each operation. This allows for a comparison between management's own ratings and those derived by OVE. For SG operations, OVE used an approach consistent with the revised 2016 PCR guidelines, while management had applied the method outlined in the 2014 PCR guidelines. As OVE pointed out in its last project performance assessment report (RE-488), the older method has considerable shortcomings and deviates considerably from the GPS for SG project evaluations. Many of these shortcomings have been addressed in the revised 2016 PCR guidelines, but their finalization in late 2016 precluded their application to PCRs prepared in the same year. Given the shortcomings in the 2014 methodology, OVE used the opportunity to test the revised approach outlined in the 2016 guidelines.⁹ This means that the rating scores assigned by OVE to each operation are not directly comparable with those assigned by management for this group of PCRs.



Of the 21 operations reviewed, 17 achieved an overall outcome rating of Partly Successful or higher.

3 Results of Public Sector Project Reviews

A. OVERALL OUTCOMES

The overall outcome rating provides a summary of an operation's performance, taking into consideration how relevant an operation and its objectives were in a given country context, to what extent and how efficiently the project reached its stated development objectives, and how likely achieved results are to be sustained. In keeping with the GPS for project evaluations and IDB's 2016 PCR guidelines, OVE rated project outcomes on a six-point scale ranging from Highly Successful to Highly Unsuccessful. As was explained above, the overall project outcome rating is as a weighted average of the core criteria component ratings.

Of the 21 operations reviewed, 17 achieved an overall outcome rating of Partly Successful or higher (Figure 3.1). This indicates quite solid performance for the 21 operations, though because they are only a small subset of the Bank's SG operations that closed between 2014-2016, no conclusions can be drawn from these operations about the overall performance of IDB's portfolio of operations that closed over that period or about performance across sectors or countries. While no operation achieved the highest overall outcome score of Highly Successful, most were rated at least Partly Successful. One FI operation was rated Unsuccessful. Two of the three PBPs were rated Partly Successful and one Partly Unsuccessful.

FIGURE 3.1
Overall project outcome ratings

Source: OVE

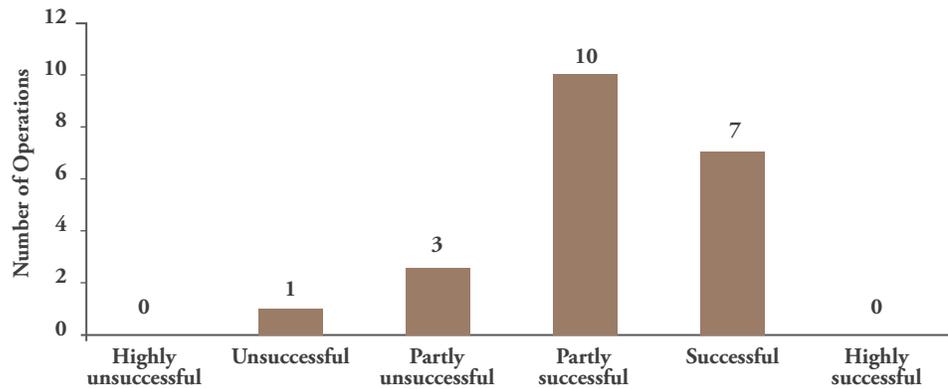
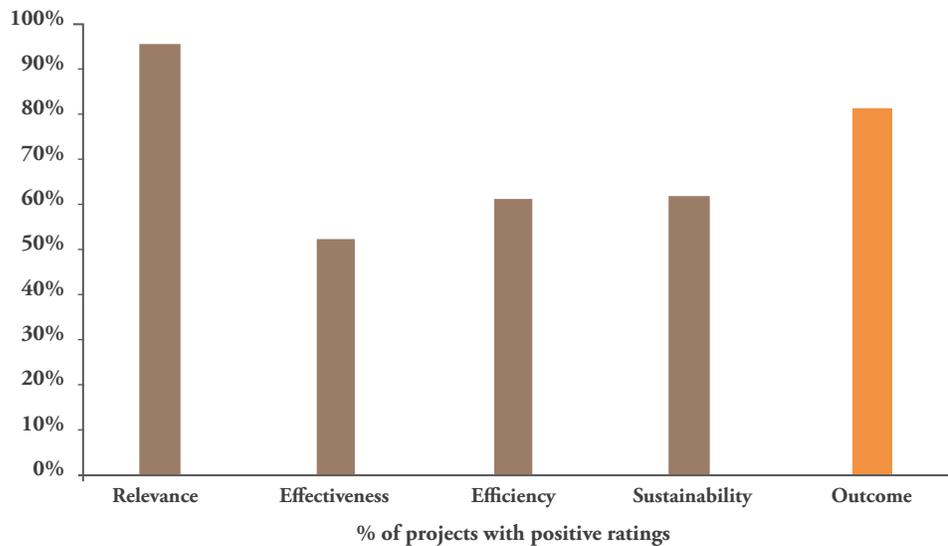


FIGURE 3.2
Core criteria ratings

Note: Relevance, Effectiveness, Efficiency, Sustainability are rated on a four-point scale, Outcome is rated on a six-point scale.
Source: OVE



B. CORE CRITERIA

A look at core criteria ratings that contribute to the overall outcome rating shows that the portfolio scored highest on relevance and lowest on effectiveness. Satisfactory or higher ratings were achieved by 95% of these operations for relevance, about 60% each for efficiency and sustainability, and 52% for effectiveness (Figure 3.2).

1. Relevance

The high share of projects with either an Excellent or Satisfactory relevance rating (20 of 21 operations) indicates that the objectives of the reviewed operations were well aligned with the development challenges in the country and with country realities (Figure 3.3). Only one operation was rated Unsatisfactory on relevance. This operation suffered from considerable design deficiencies and significant delays in reaching eligibility, so that its attempt to provide cash for on-lending in response to the financial crisis was no longer relevant.

2. Effectiveness

In assessing effectiveness, OVE evaluated the achievement of each project development objective as stated in the loan document. It assigned an achievement rating to each objective, and then derived an overall effectiveness rating from the individual objectives ratings (see Annex IV).

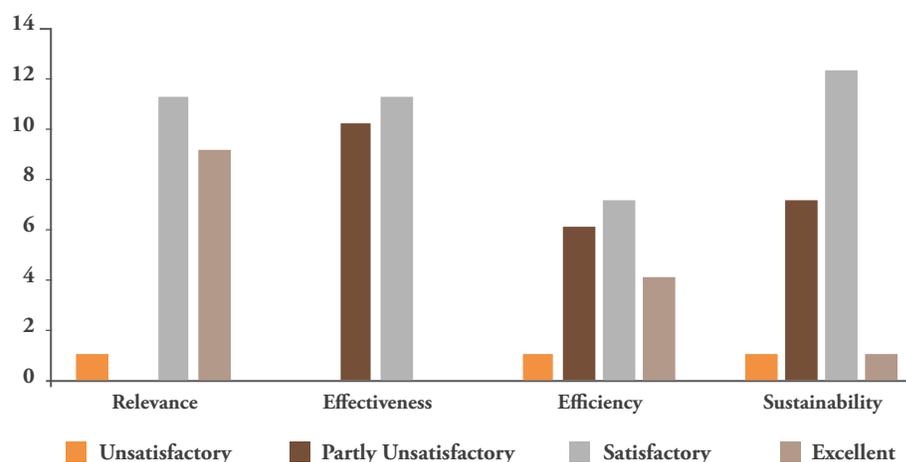


FIGURE 3.3
Distribution of core ratings for SG operations

Source: OVE

Project development objectives are not necessarily synonymous with the individual outcome indicators tracked in a project results matrix. Because project development objectives are often broader than what a single indicator can measure, it may be necessary to consider a range of indicators to determine whether a particular objective has been met. Therefore, to assess the achievement of individual development objectives, OVE mapped results indicators from the project results matrix to individual objectives. For example, one of the objectives of a health sector operation in Honduras was to increase the coverage, quality, and use of essential health services among the poor population. To assess to what extent this objective was achieved, OVE considered a range of indicators from the results matrix, including increase in the number of people covered by a new health services provider model, increase in pre- and postnatal care rates, increase in institutional deliveries, and reduction in chronic child malnutrition.

The achievement rating for each objective in the 21 operations reviewed was based on the performance of the relevant set of indicators against targets set in the loan documents (or amended within 60 days after an operation reached eligibility, if applicable).¹⁰ If the project objectives or the results matrix were changed later as a result of a formal project restructuring approved by the Board, the formally revised matrix was used as a yardstick, but this happened only once (ME-L1086) in this set of 21 projects. For PBPs, if the results matrix changed over the course of the program, OVE evaluated the end-of-program results against the end-of-program targets from each results matrix approved by the Board and then derived an overall effectiveness rating for the program.

OVE's practice of assessing the achievement of objectives against the results matrix contained in the loan document (updated within 60 days after eligibility, if applicable) differs from the approach management uses in its 2014 and 2016 PCR guidelines. Management accepts any changes to the results matrix introduced up to the time of PCR preparation. This means that indicators and targets can be revised until the project is largely completed and evaluated, which makes it much easier for a project to reach its targets and hence achieve a Satisfactory effectiveness rating as long as a result can plausibly be attributed to the project. In OVE's view, this approach undermines effective accountability and hinders learning, as it does not effectively flag operations that failed to achieve what they were expected to achieve when designed. It is also inconsistent with IDB's commitments under IDB-9, which, as noted earlier, stipulate that completion reports will compare results achieved against expected results at entry in the DEM.

OVE noted a significant tendency to introduce changes to project results matrices up to the preparation of the PCR, calling into question the validity of result matrices presented to the Board as part of loan documents at approval. Of the 21 operations reviewed by OVE, 15 had changes to either results indicators or target values for outcomes, and in eight of those cases, changes were made toward the time of PCR preparation. Essentially all had changes to outputs and their targets over the course of the operation. For two of the three PBP's reviewed, the results matrix changed considerably between the first and last operation approved. In some operations, such changes resulted in a substantial reduction of outcome targets compared to what had originally been expected. While adjustments to project outputs through implementation is understandable as projects adjust to changing circumstances, adjustments to objectives and related outcomes through the end of project implementation is questionable for a system that attempts to measure project success based on project objectives.

About half of the 21 operations reviewed demonstrably achieved the majority of their objectives and were rated Satisfactory on effectiveness. Ten projects received an effectiveness rating of Partly Unsatisfactory. Lack of information on results achieved was a factor contributing to a less than Satisfactory effectiveness rating for four of these 10 projects, though in each case it was combined with failure to make substantial progress on some of the objectives for which information on achievement was available. The other six operations were rated less than Satisfactory because they failed to achieve the majority of their objectives. No operation was rated either Excellent (meaning it fully achieved all its targeted outcomes and objectives) or Unsatisfactory (meaning it did not achieve any of its objectives). All three PBP's reviewed were rated Partly Unsatisfactory on effectiveness.

3. Efficiency

The PCR guidelines provide the option of carrying out a cost-benefit analysis, a least-cost analysis, or just an assessment of project cost and time overruns to gauge the efficiency with which projects spent loan funds. However, neither the 2014 nor the 2016 guidelines clearly spell out when a cost-benefit or least-cost analysis is required, leaving it to PCR authors to determine how to complete the efficiency analysis. Of the 18 investment operation PCRs reviewed by OVE, an ex-post cost-benefit analysis had been undertaken for eight and a cost-effectiveness analysis for one. Some of these analyses were based on assumptions that were not clearly spelled out or were not tested for validity in the context of the results actually achieved. Furthermore, the updated PCR guidelines continue to stipulate that ex-post rates of return or cost-benefit ratios should be compared to ex-ante discount rates. As OVE pointed out in its previous report on measuring project performance at the IDB (RE-488), a mere comparison of ex-post with ex-ante rates of return (or ex-ante discount rates used to calculate cost-benefit ratios) can result in misleading assessments.

Of the 18 investment operations reviewed, 11 achieved an efficiency rating of Satisfactory or higher. In five operations, the cost-benefit calculations led to clear conclusions on the efficiency of resource use. In the absence of a valid cost-benefit analysis, and where such an analysis might have been difficult to undertake, OVE considered other factors to assess efficiency: whether the operation was implemented on time, whether outputs were delivered within cost estimates, whether project interventions helped generate substantial program efficiencies, or whether implementation was affected by other inefficiencies. Factors that led to a less than Satisfactory efficiency ratings included a failure to perform an ex-post economic analysis despite an ex-ante analysis combined with considerable implementation delays; cost overruns and underuse of project financed assets due to their degradation prior to project completion; a large number of small activities carried out over dispersed territory leading to implementation difficulties; and teachers not being adequately equipped to make effective use of classroom IT investments due to lack of training and frequent desertion. In the case of a financial intermediary operation the performance of relevant sub-loans deteriorated and the targeted sub-portfolio grew substantially less than the loan amount, suggesting that at least half of the loan amount was not used to effectively grow the portfolio.

4. Sustainability

In over one-third of the SG operations reviewed, there is a risk that the results will not be sustained. The sustainability rating considers risks to the continuation of project-generated benefits, including any unmitigated social and environmental risks arising



from underperformance on safeguards. Of the 21 operations reviewed, 13 achieved a sustainability rating of Satisfactory or Excellent, while eight were rated less than Satisfactory. A lack of guaranteed financial resources to maintain investments was by far the most frequent factor leading to a less than Satisfactory sustainability rating, affecting three of the four environment projects, as well as urban and transport projects and an education project with heavy investments in IT. The sustainability analysis tended to be weaker than other aspects of the PCRs; in some the discussion was limited to risks that affected timely project implementation, rather than risks to sustaining results beyond project completion. On the environmental and social safeguards front, 10 of the 18 investment operations had been classified as Category C projects with no expected environmental and social risks; the remainder were medium-risk Category B projects, on which the PCRs generally reported little beyond stating that safeguards policies had been complied with.

C. QUALITY OF PCRs

Most of the 21 PCRs reviewed were of good quality. OVE assessed the quality of PCRs on a four-point scale, ranging from Excellent to Poor, and rated 13 of them as Good or Excellent. The remaining eight received a Fair rating – that is, they were overall of reasonable quality but had a key shortcoming. In general, PCRs were clear

and candid, analyzed the project authorizing environment and design aspects well, and documented changes undertaken throughout implementation; in addition, many contained valuable lessons that can inform the design of future projects. Where there were shortcomings, they were often related to the economic analysis/efficiency assessments and to the analysis of sustainability, including documentation on safeguards performance. The generally good quality of PCRs suggests that once the methodology is further improved, and if PCRs are consistently produced for all completed operations, they can become a useful tool not only for accountability but also for learning.



The purpose of independent validation of self-reported ratings is to ensure the accuracy and credibility of the project performance reporting system.

© IDB

4 Results of Private Sector Project Reviews

A. OVERALL OUTCOMES

The objectives-based approach to private sector operations puts more emphasis on development results than the previous NSG evaluation methodology. The overall outcome rating for an NSG operation assesses a project's overall contribution to the development of the host country. As for SG operations, OVE derived the overall outcome rating as a weighted average of the core criteria ratings for relevance, effectiveness, efficiency, and sustainability, with relevance, efficiency and sustainability accorded a weight of 20% each and effectiveness 40%.

Of the 30 NSG operations reviewed, 12 were assigned a positive outcome rating by OVE (Figures 4.1 and 4.2). A breakdown by approving private sector window suggests that operations approved by IIC had a somewhat higher success rate (seven of 14 projects rated Partly Successful or higher) than those approved by SCF (four of 11 projects) or OMJ (one of five projects), though these differences must be treated with caution given the small number of projects in each group. The overall outcome ratings were driven by low scores on effectiveness and efficiency, as discussed below. It must be remembered that these operations were designed prior to the introduction of the DELTA and hence less attention was paid to ensuring that clear development objectives were formulated and indicators identified to measure achievement of such objectives. But this only partly explains the relatively low share of operations with a positive overall outcome rating.

FIGURE 4.1
Share of projects with positive overall outcome ratings

Source: OVE

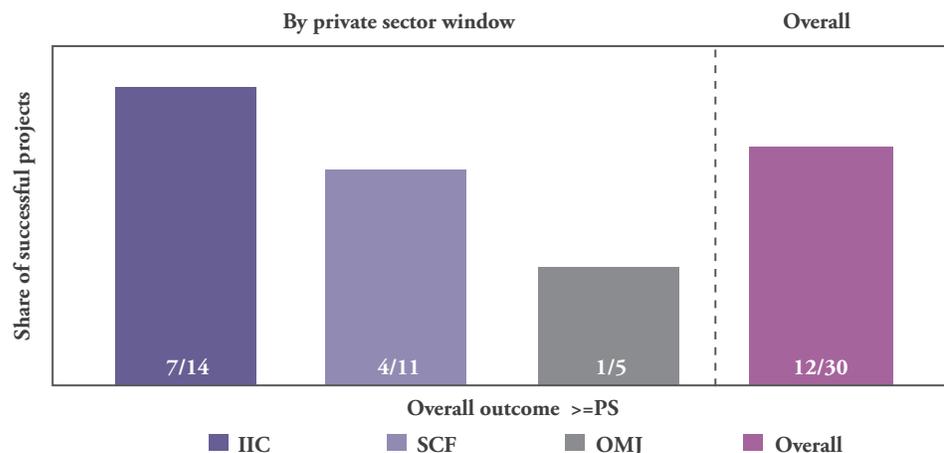
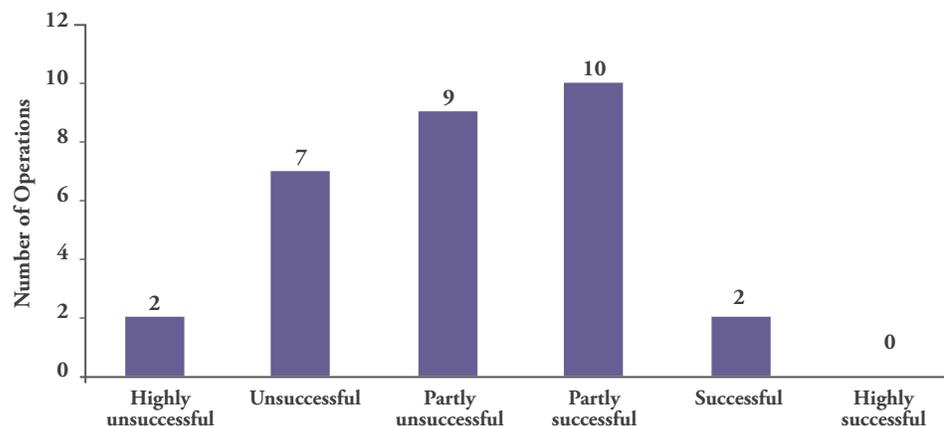


FIGURE 4.2
Frequency distribution of outcome ratings

Source: OVE



B. CORE CRITERIA

1. Relevance

Twenty-two of the 30 validated NSG projects with EOM 2015 were rated as Satisfactory or higher on relevance – that is, these operations aimed to achieve development objectives that were relevant to the country’s development challenges and were aligned with the approving private sector window’s business priorities. All five OMJ-approved operations, eight of 11 SCF-approved operations, and nine of 14 IIC-approved operations achieved a positive relevance rating (Figure 4.3). One IIC-approved operation was rated Unsatisfactory on relevance, as it supported a large producer with monopolistic power, was well aligned with neither the country’s development priorities nor IIC’s business objectives, and did not demonstrate that the company could not have accessed funding elsewhere. Some operations had Partly Unsatisfactory relevance ratings because the lending took place in a distorted policy environment and hence the operation failed to address key factors impeding access to finance. The relevance of two operations deteriorated shortly after their approval when the clients changed their business strategies so that the operations’ objectives and designs were no longer aligned with them. Lack of alignment with country

development needs and priorities and with IIC business objectives also affected the relevance of several operations. One operation financing the transfer of equity between private investors – an activity that is generally excluded from IIC financing – had originally been rejected by the Board but was eventually approved with a waiver, following the addition of a component that was to enhance the client’s environmental management and safety capacity. Yet the client showed little interest in the added activities, which were never implemented, and the loan was eventually prepaid.

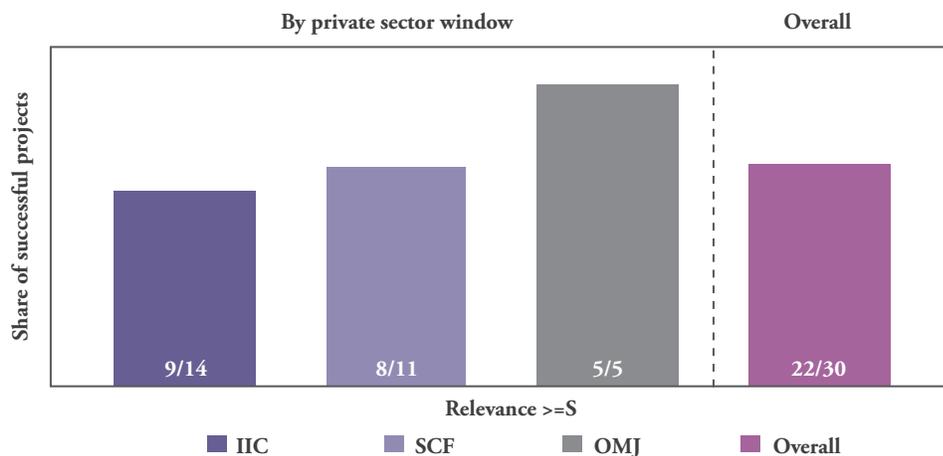


FIGURE 4.3
Share of projects with positive relevance ratings

Source: OVE

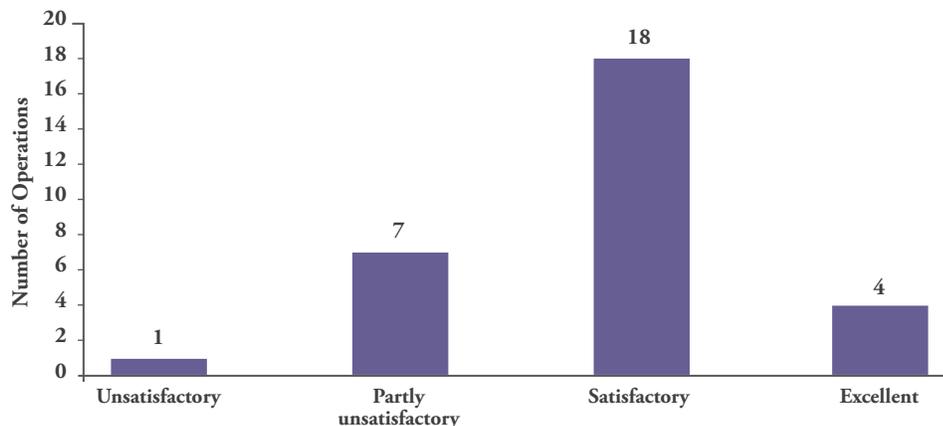


FIGURE 4.4
Frequency distribution of relevance ratings

Source: OVE

2. Effectiveness

Nineteen of the 30 NSG operations reviewed did not demonstrably achieve the majority of their development objectives. Of the 30 operations, 11 achieved a Satisfactory effectiveness rating, including one of the five OMJ-approved, six of the 14 IIC-approved, and four of the 11 SCF-approved operations (Figure 4.5) The private sector initiatives supported by NSG operations are expected to yield broader development impacts than mere benefits to company owners or shareholders. Although the loan documents for all 30 operations stated some expected development outcomes, these were not always matched with indicators to monitor their progress. Across the 30 operations reviewed, 12 had deficiencies in

the definition of indicators or target-setting at approval. For example, six operations targeted an increase in employment, but only three identified a related indicator and set targets. Eleven lacked information on some indicators in the XSRs. The lack of indicators aligned with stated development outcomes may be partly due to the fact that these operations were designed prior to the approval of the DELTA, when more emphasis was placed on measuring business and financial performance than development outcomes. When information related to development objectives could not be found in the XSRs or otherwise, OVE rated the objective as less than Satisfactory. Overall, out of the 19 operations with a less than Satisfactory effectiveness rating, 14 were rated below Satisfactory because they clearly failed to achieve the majority of their objectives, four failed to achieve at least some of their objectives and did not provide information to determine whether the others had been achieved, and for one operation lack of reliable information was the main determinant of the less than Satisfactory effectiveness rating.

FIGURE 4.5
Share of projects with positive effectiveness ratings

Source: OVE

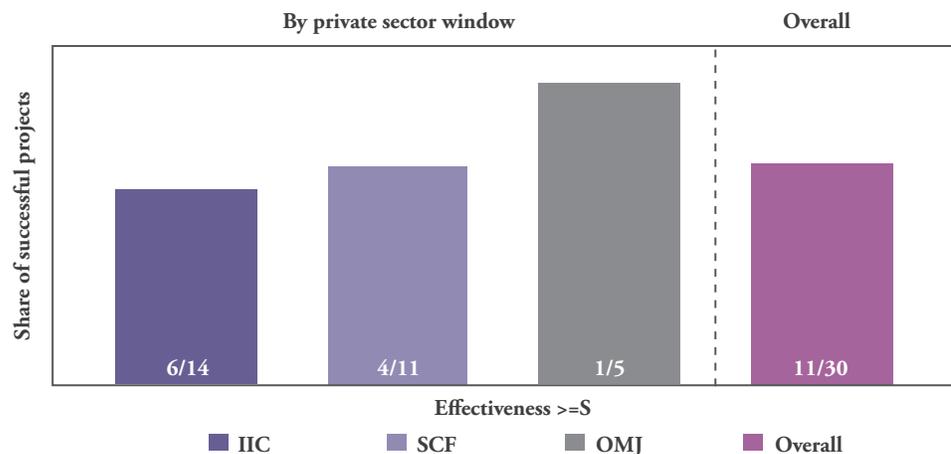
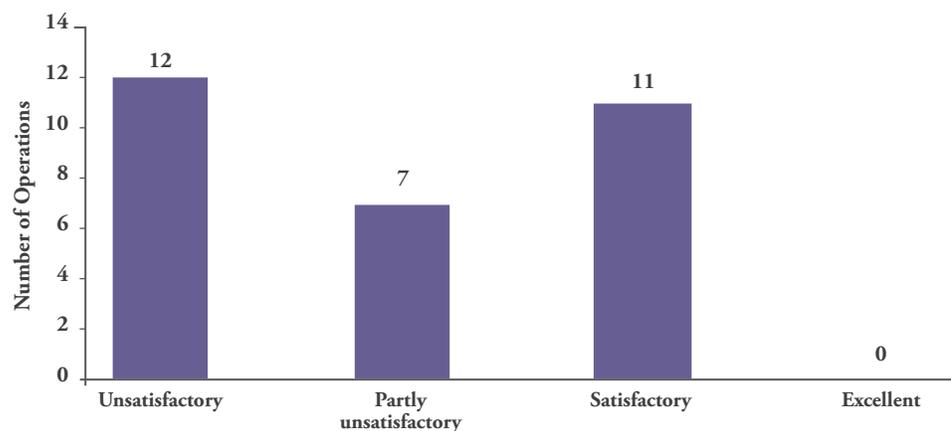


FIGURE 4.6
Frequency distribution of effectiveness ratings

Source: OVE



The 16 FI operations all aimed to improve access to finance for micro, small, or medium enterprises; some targeted female-headed small or medium enterprises, and some also aimed to expand access to mortgage financing for middle- and lower-middle-income households. However, only nine set clear growth targets

for the FI's targeted sub-portfolio, and even fewer included indicators to monitor the performance of that portfolio, while the rest simply set targets for the number of loans or amounts to be lent by the IIC/SCF/OMJ loan. Given the fungibility of money, such target-setting is inappropriate – as OVE pointed out in its recent evaluation of FI operations.¹¹ Only five of these 16 operations were rated Satisfactory, while four were rated Partly Unsatisfactory and seven Unsatisfactory. As a group, the FI operations performed somewhat lower on effectiveness (5 of 16 operations Satisfactory) than non-FI operations (6 of 14 operations Satisfactory), but given the small number of operations, differences need to be viewed with caution.

Most retail and packaging operations focused their goals on sales growth, which by itself provides little indication of development impact. They also added employment generation as an objective, but most of them did not monitor increased employment. The four energy operations aimed to support increases in renewable energy production, and three of them achieved Satisfactory effectiveness. Only one of the three agriculture operations substantially achieved its objectives.

3. Efficiency

Performance on efficiency across the NSG portfolio was low. Seven of the 30 NSG operations achieved a positive efficiency rating (Figures 4.7 and 4.8). Two aspects of NSG operations are assessed under the efficiency criterion: project financial performance and economic performance. Of the 30 NSG operations reviewed, only nine achieved a Satisfactory or higher rating on financial performance and eight on economic performance. Among FI operations, lack of information in XSRs about the performance of the targeted sub-portfolio accounted for half of the 12 operations with a less than Satisfactory efficiency rating. In four others, the performance of the relevant portfolio deteriorated and in the last two other factors were at play. The relatively poor performance of the non-FI portfolio reflects the large share of operations that achieved low financial and economic returns and some that also fell short of other financial or business targets.

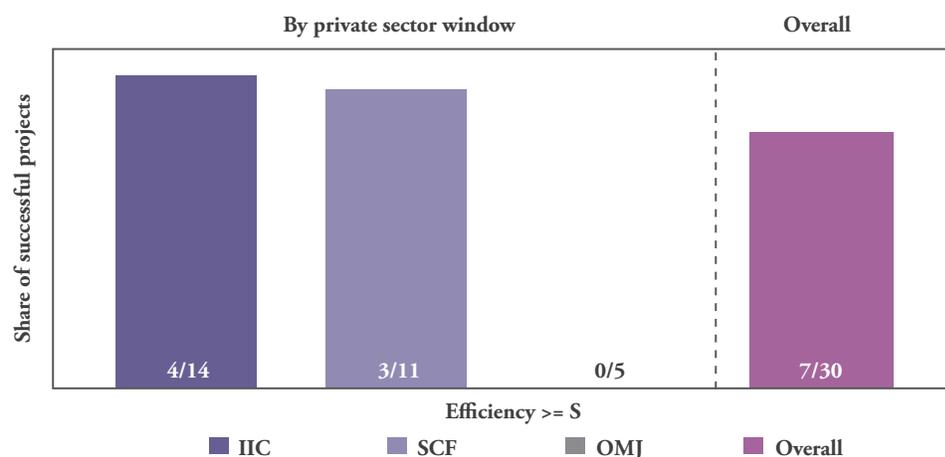
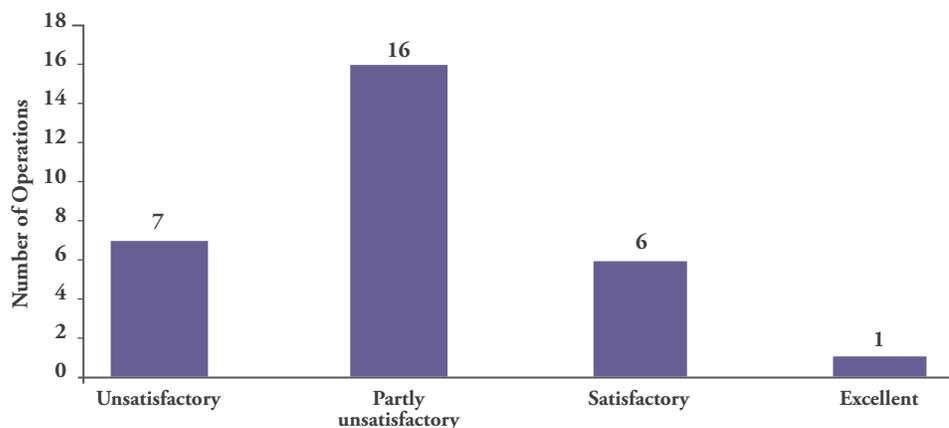


FIGURE 4.7
Share of projects with positive efficiency ratings

Source: OVE

FIGURE 4.8
Frequency
distribution of
efficiency ratings

Source: OVE



4. Sustainability

Overall, only half of the reviewed projects achieved a Satisfactory sustainability rating, with IIC-approved projects scoring better than others (Figure 4.9). The sustainability dimension of NSG operations consists of two sub-ratings: performance on environmental and social safeguards, and sustainability of project results in other dimensions, such as financial and institutional aspects. To receive an overall Satisfactory sustainability rating, an operation needs to achieve a rating of Satisfactory or better in both areas. While almost three-quarters of operations had a Satisfactory rating on environmental and social safeguards, more than a third had a negative rating on other sustainability aspects. This raises concerns about the longer-term benefits of the NSG operations reviewed.

FIGURE 4.9
Share of projects
with positive
sustainability rating

Source: OVE



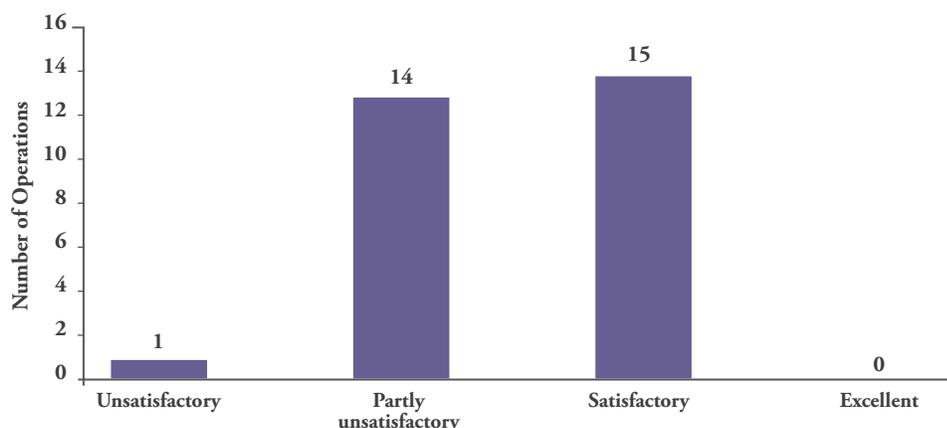


FIGURE 4.10
Frequency
distribution of
sustainability ratings

Source: OVE

C. NON-CORE CRITERIA

The vast majority of NSG operations performed well on investment outcome. The high share of operations (26 of 30) receiving Satisfactory or higher ratings on investment outcome suggests that the IDBG did well in structuring and funding operations that contributed positively to its financial success. Yet only 10 of the 26 operations that achieved a positive investment outcome rating also substantially achieved their development objectives, and only seven of them achieved a positive efficiency rating.

Slightly fewer than half of the reviewed operations demonstrated financial and nonfinancial additionality, and the share of operations demonstrating satisfactory financial additionality was somewhat lower than that demonstrating nonfinancial additionality. Financial additionality was particularly questionable for FI operations: only six of 16 operations provided clear financial additionality, reflecting a tendency to address constrained access to finance with additional liquidity, even if a lack of liquidity was not shown to have been the constraining factor.

The quality of IDBG's work on the reviewed operations was modest, particularly at the time of structuring and appraisal. Only 12 of the 30 EOM 2015 operations demonstrated satisfactory work quality during structuring and appraisal as well as during supervision (Table 4.1). Shortcomings were particularly marked during structuring and appraisal. Failure to clearly analyze and address key factors affecting access to finance and to clearly establish adequate indicators to monitor the evolution and performance of targeted portfolios were key factors affecting work quality at screening and structuring of operations. Less than satisfactory work quality at screening and structuring for non-FI operations was often due to a failure to properly assess risk factors and identify risk mitigation measures.

TABLE 4.1. SHARE OF OPERATIONS RATED SATISFACTORY OR HIGHER ON NON-CORE CRITERIA

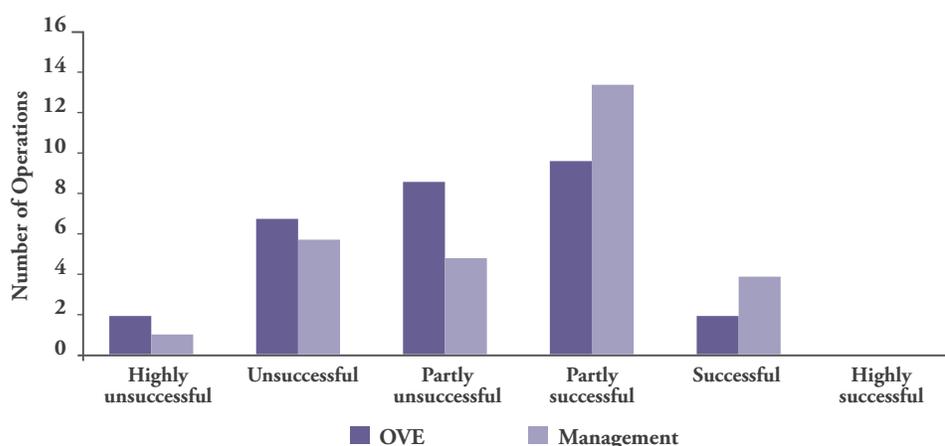
	Approving PS window			Total	Operation type	
	IIC	SCF	OMJ		FI	non-FI
Additionality share =>S	5/14	7/11	2/5	14/30	6/16	8/14
Invest/ outcome share >=S	12/14	9/11	5/5	26/30	15/16	11/14
Work Quality share >=S	4/14	6/11	2/5	12/30	6/16	6/14
# of operations	14	11	5	30	16	14

D. COMPARISON BETWEEN OVE RATINGS AND MANAGEMENT SELF-RATINGS

The purpose of independent validation of self-reported ratings is to ensure the accuracy and credibility of the project performance reporting system. The NSG validation exercise has contributed to refining IIC management’s self-reported outcome ratings for the EOM 2015 NSG operations, resulting in a final disconnect of 17 percentage points for the share of projects with positive outcome ratings (Figure 4.11). While OVE results show that 40% (12 of 30) of operations have achieved an overall positive outcome rating, management’s self-reported ratings point to a higher rate of 57% (17 of 30 operations). To some degree these differences go back to the need for further clarity in the pilot XSR guidelines in several areas. For example, a clearer definition is still needed for what constitutes a development objective and what constitutes a financial objective, as the former are assessed under effectiveness, while financial objectives are assessed under the financial performance section of efficiency. Similarly, the guidelines need to further clarify how to evaluate financial and economic performance of FI operations and how to assess sustainability. In other cases, management argued that external factors affected project performance and hence a project should not have been assessed against originally set targets.

FIGURE 4.11
Comparison between
Management and
OVE Outcome
Ratings

Source: OVE



E. QUALITY OF XSRs

The quality of the XSR was rated Good for 14 of the 30 XSRs reviewed. Thirteen were rated Fair, and three were rated Poor. OVE assessed the quality of XSRs on the same four-point scale as PCRs (see para 3.14). In general, the XSRs did well in distilling valuable lessons that can help inform the design of future operations. A common factor among XSRs with a less-than-Good rating was an inconsistency between the information presented and the self-ratings assigned to the operations. This inconsistency is likely to decline as IIC staff gain experience with the XSR methodology and with the DELTA framework, and as XSR guidelines are further clarified.



Among the MDBs, the IDBG is at the forefront in moving toward a unified framework to evaluate the performance of its SG and NSG operations.

5 Conclusions and Recommendations

Among the MDBs, the IDBG is at the forefront in moving toward a unified framework to evaluate the performance of its SG and NSG operations. The IDBG has made significant progress, on both the SG and NSG sides, toward establishing a strong objectives-based project evaluation system. This report summarizes the results of OVE's first review of SG and NSG operations using this unified framework to assign project outcome ratings.

OVE reviewed the PCRs of 21 SG operations that had been prepared in accordance with revised PCR guidelines and had been completed by end-2016. Because this sample represents less than 10% of the operations that closed between 2014 and March 2016, the findings on the performance of these 21 operations are not sufficient to draw any conclusions on the performance of the overall relevant portfolio. About 80% of the operations reviewed achieved a positive overall project outcome rating, with operations scoring highest on relevance and lowest on project effectiveness. About half of the reviewed operations demonstrably achieved the majority of the development objectives they set out to achieve.

OVE reviewed the XSRs of all 30 NSG operations that had reached early operating maturity in 2015. About 40% of the operations achieved a positive overall outcome rating, with IIC-approved operations faring somewhat better than those approved by SCF and OMJ. The overall outcome ratings reflected low scores on effectiveness and efficiency, with almost two-thirds of the reviewed operations not demonstrably achieving the majority of their development objectives and fewer than one-quarter achieving a Satisfactory efficiency rating. For four of the 30 operations, the lack of achievement of some objectives combined with a lack of information to assess the achievement of others led to a low effectiveness rating and for one operation lack of information was the driving factor behind the low effectiveness rating.

Although progress has been made in developing IDB's and IIC's self-evaluation systems, further efforts are needed to consolidate these gains and ensure that the systems develop into robust and useful tools for measuring and monitoring portfolio performance. On the NSG side, continued efforts are needed to identify the development objectives and results that each operation seeks to achieve, as opposed to merely focusing on business objectives. Results frameworks need to be clearly aligned with the stated objectives to facilitate measuring progress toward achieving them. The new DELTA should help impart this focus.

The NSG pilot XSR guidelines have proven to be a good basis for the preparation of project self-evaluations and OVE validations, though this exercise has pointed to some areas that require further refinement: They need to more clearly define cut-off values for indicators to be considered fully or partly achieved as opposed to not achieved, and to further clarify how to assess sustainability and efficiency, particularly in FI operations. Further clarification is also required on how to derive the overall outcome rating. A core challenge for the preparation of XSRs is to ensure that the self-ratings are consistent with the evidence and analysis presented.

On the SG side, OVE found that PCRs are generally of good quality and candor and distill a range of important lessons that can inform the design of future projects, though there are still challenges that need to be addressed. First, Management's current practice of assessing achievements against indicators and targets established up to the last Project Monitoring Report (or even during the PCR exercise) prevents a candid assessment of how effectively an operation has indeed performed. Second, the guidelines need to make clear that an objectives-based methodology does not mechanically assess against indicators, but steps back and assesses the achievement of clearly stated development objectives against which relevant indicators are mapped. On the project efficiency side, more clarity and guidance are needed about when a PCR is expected to include a cost-benefit analysis. Similarly, more clarity is needed on when a cost-effectiveness analysis is sufficient and what such analyses should entail. In addition, to be consistent with the NSG framework and the SG DEF, PCRs should include additionality and Bank performance as additional dimensions to be assessed. Finally, it is important that PCRs be delivered consistently and on a timely basis. A system in which credible completion reports are prepared for only a small number of operations or are delivered years after project closure does not permit the formation of any conclusions on the quality of IDB's portfolio, regardless of how high the quality of the few PCRs really is.

Validation of project self-evaluations by the independent evaluation office is a core component of a strong project evaluation architecture. A validation system operates most effectively and efficiently when the self-evaluation and the validation apply the same ratings methodology, and thus moving toward a unified rating system used by both management and OVE – particularly in the case of IDB where current differences are more marked – remains an important step in further strengthening the

system. Validations may confirm or adjust management's own project performance ratings, but in either case the validated ratings should be used by both management and the independent evaluation office as the final ratings for purposes of corporate reporting on portfolio performance, as is done in IDB's partner MDBs.

In light of these conclusions, OVE recommends the following:

For IIC management:

- (i) Work with OVE to update the pilot XSR guidelines to arrive at a set of guidelines that both parties will use when evaluating and validating project performance, ensuring consistency to the extent possible with IDB's PCR guidelines. Drawing on the experience gained during this EOM 2015 exercise, issues to be clarified in the revised guidelines include:
 - the types of development objectives to be covered in the effectiveness section (excluding financial objectives which should be covered under the financial performance part of efficiency);
 - how ratings on individual objectives are determined and aggregated into an overall effectiveness rating, ensuring consistency with the SG methodology;
 - how to assess and rate financial and economic performance for FI operations when information on the performance of the relevant sub-portfolio is not available;
 - the factors to consider (in addition to safeguards performance) in assessing sustainability;
 - the weighting and rating scale to use in deriving the overall project outcome rating, ensuring consistency with SG methodology.
- (ii) Use OVE's final validated ratings for purposes of corporate reporting on portfolio performance.
- (iii) Work with OVE to further define the process and timetable for the annual preparation of XSRs by management and their validation by OVE.

For IDB management:

- (i) Work with OVE to arrive at a set of PCR guidelines that both parties will use when evaluating and validating project performance, ensuring consistency to the extent possible with IIC's XSR guidelines. Specific aspects of the 2016 PCR guidelines that need to be addressed include:



- the definition of the outcome indicators and targets against which project outcomes are to be assessed. To be consistent with IDB-9 commitments and current XSR guidelines, project outcomes should be assessed against outcome indicators and targets established at approval, which can be changed if an operation is formally restructured with Board approval. If there are delays in an operation reaching eligibility, OVE recommends that updates also be allowed within 60 days of an operation reaching eligibility;
- the way that the effectiveness of PBP series is assessed. OVE recommends that the program be assessed against end of program targets in each matrix of the series and the overall effectiveness rating for the series be an average of those individual assessments;
- the type and content of efficiency analysis required under various circumstances;
- how to include additionality and Bank performance as evaluation criteria, consistent with XSR guidelines;

- the deadline for completing PCRs. OVE recommends that completion be mandatory within nine months after an operation reaches closed status in Convergence.¹²
- (ii)** Apply the revised project rating system to all future PCRs that have not yet been shared with Governments for comments, irrespective of whether they are prepared under the 2014 guidelines or later ones. Even PCRs prepared under the 2014 guidelines provide information to do so.
- (iii)** Use OVE's final validated ratings for purposes of corporate reporting on portfolio performance.
- (iv)** Work with OVE to further define the process and timetable for the annual preparation of PCRs by management and their validation by OVE.

- ¹ Report on the Ninth General Increase in the Resources of the Inter-American Development Bank, AB-2764.
- ² The exact definition of early operating maturity varies by project type, but implies that IDBG has made its financial material disbursement and received at least one set of audited financial statements covering 12-36 months of operating revenues post-disbursement/project completion, with the period depending on the type of project. See Annex 2 of the IDBG XSR guidelines for further details.
- ³ For a detailed review of the differences between the prior methodologies and the objectives based methodology see OVE, Measuring Project Performance at the IDB: Recent Development in the Project Completion Report and the Expanded Project Supervision Report Systems, RE-488, May 2015.
- ⁴ IDB Results Framework 2012-2015, paragraph 6.6.
- ⁵ “Evaluation Principle H: Borrower Performance assesses the adequacy of the Borrower’s assumption of ownership and responsibility during all project phases.” ECG Big Book on Evaluation Good Practice Standards, November 2002, p. 31. Ratings should be symmetric and carried out on a four- or six-point scale, as specified in Operational Practice 20.2; *ibid.*, p. 32.
- ⁶ In addition to the XSRs for NSG operations with EOM 2015, the XSRs reviewed also included two operations that had reached EOM earlier but for which IIC had only been prepared in 2016, as they had been succeeded by a follow-on operation that reached EOM in 2015. In keeping with the GPS for NSG project evaluations, which prescribe that the results of project ratings should be reported by cohort of projects that reach EOM in the same year, this report will present the validation results of only the 30 operations that reached EOM in 2015.
- ⁷ The March 2016 project closing date cut-off is used to compare the number of PCRs reviewed to the universe of projects closed, as it typically takes about nine months to complete a PCR after project closure, hence all projects closed by March 2016 could have been expected to deliver the PCR by December 2016. Four PCRs were delivered “early” for operations that closed after March 2016.
- ⁸ The 2014 PCR guidelines apply only to operations that were approved after January 1, 2009 with a DEM, and had fully disbursed from August 5, 2014, onwards. PCRs for programmatic operations are prepared only when the last operation in the program closes. No PCRs are required for cancelled operations that disbursed less than 30%. For 20 of the 43 PCRs that were to be prepared under the 2014 guidelines, Sector Managers had granted an extension for the preparation of the PCRs. For operations approved before 2009, the preparation of the PCRs follows the old methodology; OVE does not review such PCRs.
- ⁹ While OVE’s approach used here is largely consistent with the 2016 PCR guidelines, there are some differences, which will be outlined in the effectiveness section of this report.
- ¹⁰ The 60-day cutoff date for updating the results matrix against which OVE evaluated results achieved is consistent with the deadline by which project teams are expected to update the results matrix when a project starts to disburse loan funds.
- ¹¹ Evaluation of IDB Group’s Work through Financial Intermediaries, 2016, RE-486-2, CII/RE-18.
- ¹² The closed (CO) status is activated when the disbursed resources have been fully justified, unjustified resources have been returned, and undisbursed amounts have been cancelled.

ANNEX II - LIST OF SG OPERATIONS REVIEWED

Operation Number	Operation Name	Sector	Country	Operation Type	Appr. Year
AR-L1073	CCLIP: Program of Technological Innovation	Science and Tech	AR	PDL	2009
BL-L1014	Community Action for Public Safety	Public sector	BL	ESP	2010
BO-L1047	Transparency and Anticorruption Program	Public sector	BO	ESP	2009
BO-L1053	Environmental Management of Misicuni Watershed	Environment	BO	ESP	2010
BR-L1171	Catanduva Integrated Urban Development Program	Urban	BR	GCR	2009
CO-L1141	Support to the Health Sector Reform II	Health	CO	PBP	2014
CO-X1004	Protecting Biodiversity in the Southwestern Caribbean Sea	Environment	CO	IGR	2009
EC-L1040	Modernization of Pumping Stations on the Esmeraldas-Quito Multiproduct Pipeline	Energy	EC	ESP	2010
EC-L1070	Support for the Transmission Program	Energy	EC	ESP	2010
EC-L1087	Electrification Program for rural and marginal urban areas of Ecuador	Energy	EC	ESP	2011
HA-L1086	Emergency Road Rehabilitation Program in Response to Hurricane Sandy	Transport	HA	IRF	2012
HO-L1059	Strengthen Decentralized Management and Supply of Health Services in Honduras	Health	HO	ESP	2010
HO-L1062	Primary Education and Technology Integration Program	Education	HO	ESP	2011
HO-L1071	Support to the Social Protection Network Program	Social Protection	HO	ESP	2012
ME-L1086	Dignified Schools Program (PED)	Education	ME	ESP	2011
NI-L1010	Storm water Drainage and Development Management Sub Watershed III Managua	Environment	NI	ESP	2009
NI-L1046	Global Multisector Credit Program	FI	NI	GCR	2009
PE-L1070	Water Resources Management Modernization Program	Environment	PE	ESP	2009
PE-L1099	Program for Improving Productivity and Competitiveness	Competitive-ness	PE	PBP	2010
PE-L1100	Support for Social Sector Reforms Program	Social Protection	PE	TCR	2010
PE-L1138	Program to Reduce the Vulnerability of the State to Disasters III	Disaster Risk	PE	PBP	2014

ESP = Specific Investment Operation; GCR = Global Credit Operation; IRF = Immediate Response Facility for Emergencies; IGR = Investment Grant Operation; PDL = Performance-Driven Loan; PBP = Programmatic Policy-Based Program; TCR = Technical Cooperation Loan.
Source: OVE

ANNEX III - OVE RATINGS OF SG OPERATIONS

Operation Number	Operation Name	Appr. Year	Overall
AR-L1073	CCLIP: Program of Technological Innovation	2009	Successful
BL-L1014	Community Action for Public Safety	2010	Partly Successful
BO-L1047	Transparency and Anticorruption Program	2009	Partly Successful
BO-L1053	Environmental Management of Misicuni Watershed	2010	Partly Successful
BR-L1171	Catanduva Integrated Urban Development Program	2009	Partly Unsuccessful
CO-L1141	Support to the Health Sector Reform II	2014	Partly Unsuccessful
CO-X1004	Protecting Biodiversity in the Southwestern Caribbean Sea	2009	Partly Unsuccessful
EC-L1040	Modernization of Pumping Stations on the Esmeraldas-Quito Multiproduct Pipeline	2010	Successful
EC-L1070	Support for the Transmission Program	2010	Successful
EC-L1087	Electrification Program for rural and marginal urban areas of Ecuador	2011	Successful
HA-L1086	Emergency Road Rehabilitation Program in Response to Hurricane Sandy	2012	Partly Successful
HO-L1059	Strengthen Decentralized Management and Supply of Health Services in Honduras	2010	Successful
HO-L1062	Primary Education and Technology Integration Program	2011	Partly Successful
HO-L1071	Support to the Social Protection Network Program	2012	Partly Successful
ME-L1086	Dignified Schools Program (PED)	2011	Successful
NI-L1010	Storm water Drainage and Development Management Sub Watershed III Managua	2009	Successful
NI-L1046	Global Multisector Credit Program	2009	Unsuccessful
PE-L1070	Water Resources Management Modernization Program	2009	Partly Successful
PE-L1099	Program for Improving Productivity and Competitiveness	2010	Partly Successful
PE-L1100	Support for Social Sector Reforms Program	2010	Partly Successful
PE-L1138	Program to Reduce the Vulnerability of the State to Disasters III	2014	Partly Successful

Source: OVE.

Relevance	Effectiveness	Efficiency	Sustainability	Quality of PCR
Excellent	Satisfactory	Satisfactory	Excellent	Excellent
Satisfactory	Partly Unsatisfactory	Satisfactory	Satisfactory	Good
Satisfactory	Partly Unsatisfactory	Excellent	Satisfactory	Good
Satisfactory	Satisfactory	Partly Unsatisfactory	Partly Unsatisfactory	Fair
Satisfactory	Partly Unsatisfactory	Partly Unsatisfactory	Partly Unsatisfactory	Fair
Satisfactory	Partly Unsatisfactory	NA	Satisfactory	Good
Satisfactory	Partly Unsatisfactory	Partly Unsatisfactory	Unsatisfactory	Good
Excellent	Satisfactory	Satisfactory	Satisfactory	Good
Excellent	Satisfactory	Excellent	Satisfactory	Good
Excellent	Satisfactory	Excellent	Satisfactory	Good
Satisfactory	Satisfactory	Satisfactory	Partly Unsatisfactory	Fair
Excellent	Satisfactory	Satisfactory	Partly Unsatisfactory	Fair
Satisfactory	Satisfactory	Partly Unsatisfactory	Partly Unsatisfactory	Good
Excellent	Partly Unsatisfactory	Satisfactory	Satisfactory	Fair
Satisfactory	Satisfactory	Satisfactory	Satisfactory	Good
Excellent	Satisfactory	Partly Unsatisfactory	Satisfactory	Good
Unsatisfactory	Partly Unsatisfactory	Unsatisfactory	Partly Unsatisfactory	Fair
Satisfactory	Satisfactory	Partly Unsatisfactory	Partly Unsatisfactory	Fair
Excellent	Partly Unsatisfactory	NA	Satisfactory	Good
Satisfactory	Partly Unsatisfactory	Excellent	Satisfactory	Good
Excellent	Partly Unsatisfactory	NA	Satisfactory	Fair