



Project Evaluation

# **OVE's Review of Project Completion Reports (PCRs) and Expanded Supervision Reports (XSRs): The 2022 Validation Cycle**

**Copyright © 2022 Inter-American Development Bank. This work is licensed under a Creative Commons IGO 3.0 Attribution - NonCommercial - NoDerivatives (CC-IGO BY-NC-ND 3.0 IGO) license**

(<http://creativecommons.org/licenses/by-nc-nd/3.0/igo/legalcode>) and may be reproduced with attribution to the IDB and for any non-commercial purpose. No derivative work is allowed.

Any dispute related to the use of the works of the IDB that cannot be settled amicably shall be submitted to arbitration pursuant to the UNCITRAL rules. The use of the IDB's name for any purpose other than for attribution, and the use of IDB's logo shall be subject to a separate written license agreement between the IDB and the user and is not authorized as part of this CC-IGO license.

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

The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent.



**Inter-American Development Bank, 2022**

Office of Evaluation and Oversight  
1350 New York Avenue, N.W.  
Washington, D.C. 20577  
[www.iadb.org/evaluation](http://www.iadb.org/evaluation)

**RE-565  
October 2022**

**Project Evaluation**

**OVE's Review of  
Project Completion  
Reports  
and Expanded  
Supervision  
Reports:  
The 2022  
Validation Cycle**

**Office of Evaluation and Oversight**





# Contents

<b>Acknowledgements</b> .....	<b>v</b>
<b>Acronyms and Abbreviations</b> .....	<b>v</b>
<b>Executive Summary</b> .....	<b>vii</b>
<b>Introduction</b> .....	<b>01</b>
A. The IDB Group’s evaluation methodology and the role of OVE .....	01
B. Scope of this report .....	02
<b>Results of Operations Validated in 2022</b> .....	<b>04</b>
A. Public sector projects reviewed in 2022.....	05
B. Private sector projects reviewed in 2022 .....	13
C. Review of Management report on prepaid and repaid NSG projects between 2016 and 2019.....	20
<b>Performance Over Time</b> .....	<b>26</b>
A. Performance of public sector operations, 2021–2022 .....	27
B. Performance of private sector operations, 2018–2022.....	28
<b>Quality of Self-Assessments</b> .....	<b>32</b>
A. Quality of PCRs validated in 2022.....	33
B. Quality of XSRs validated in 2022.....	37
<b>Conclusions and Recommendations</b> .....	<b>40</b>
<b>References</b> .....	<b>46</b>

<b>Annex I:</b>	<a href="#">IDB Group’s project evaluation framework</a>
<b>Annex II:</b>	<a href="#">SG operations reviewed</a>
<b>Annex III:</b>	<a href="#">OVE’s ratings of SG operations and core criteria</a>
<b>Annex IV:</b>	<a href="#">OVE’s validation notes on SG operations</a>
<b>Annex V:</b>	<a href="#">Rating criteria for quality of self-assessments</a>
<b>Annex VI:</b>	<a href="#">NSG operations reviewed with full XSRs and with Mini-XSRs</a>
<b>Annex VII:</b>	<a href="#">OVE’s aggregate ratings of NSG operations with full XSRs and with Mini-XSRs</a>
<b>Annex VIII:</b>	<a href="#">Mini-XSR methodology</a>

## [Response by IDB and IDB Invest Management](#)

# Acknowledgements

This report was prepared by Chiaki Yamamoto (team leader), Luisa Riveros, Oliver Peña-Habib, Stefania De Santis, and Julie King, under the direction of Ivory Yong, OVE Director, and Monika Huppi, Principal Advisor.

## Acronyms and Abbreviations

<b>CBA</b>	Cost-benefit analysis
<b>CEA</b>	Cost-effectiveness analysis
<b>CO</b>	Operational closure
<b>DEO</b>	Development Effectiveness Overview
<b>E&amp;S</b>	Environmental and social
<b>EOM</b>	Early operating maturity
<b>FI</b>	Financial institution
<b>GHG</b>	Greenhouse gas
<b>I&amp;E</b>	Infrastructure and energy
<b>IDB</b>	Inter-American Development Bank
<b>M&amp;E</b>	Monitoring and evaluation
<b>MSMEs</b>	Micro, small, and medium enterprises
<b>NPL</b>	Nonperforming loan
<b>NSG</b>	Non-sovereign-guaranteed
<b>OVE</b>	Office of Evaluation and Oversight
<b>PBP</b>	Programmatic policy-based loan
<b>PCR</b>	Project Completion Report
<b>SG</b>	Sovereign-guaranteed
<b>XSR</b>	Expanded Supervision Report



# Executive Summary

This report summarizes the results of the Office of Evaluation and Oversight's (OVE's) annual validation of Project Completion Reports (PCRs) and Expanded Supervision Reports (XSRs). PCRs are self-assessments of project performance by the Inter-American Development Bank (IDB); XSRs are their equivalent for IDB Invest. Both PCRs and XSRs are intended to serve as accountability and learning tools.

The IDB Group's self-assessment methodology is objectives-based. Project performance is measured against four core criteria: (i) relevance of the project to the country's needs and realities and to the Bank's strategic objectives, (ii) effectiveness of the project in achieving the objectives stated at project approval, (iii) the efficiency with which project resources were used, and (iv) the sustainability of the results achieved. Each core criterion is rated on a four-point scale ranging from "unsatisfactory" to "excellent." Based on the assessment of the core criteria, each project receives an overall outcome rating, using a six-point scale ranging from "highly unsuccessful" to "highly successful." In addition to the core criteria, the PCR and XSR Guidelines require the inclusion of noncore criteria, which are rated (on the same four-point scale as the core criteria) but do not count toward a project's overall outcome rating. For simplicity, the report groups ratings into positive and negative categories. A positive rating is any one of the top three ratings on the six-point scale used for the overall project development outcome rating and any one of the top two ratings on the four-point scale used for core and noncore criteria and for the quality of the PCR or XSR.

OVE supports accountability and learning through the validation of completed PCRs and XSRs. Its independent validation of all PCRs and XSRs prepared by Management contributes to the credibility of the IDB Group's project performance reporting system. Management assigns performance ratings to each completed project or program in the PCRs and to each operation reaching early operating maturity (EOM) in the XSRs, based on the relevant guidelines. OVE's validation assesses whether these self-assessments are substantiated by evidence and are prepared in accordance with the respective guidelines. Based on this assessment, OVE assigns a final overall project outcome rating. This is the sixth consecutive year in which OVE has validated these self-assessments.



During the 2022 validation cycle, OVE reviewed and validated the PCRs of all 62 sovereign-guaranteed (SG) operations and 47 XSRs for non-sovereign-guaranteed (NSG) operations prepared by Management. These operations represented US\$6.8 billion in SG approvals and US\$2.6 billion in NSG approvals. In addition, this year IDB Invest Management submitted to OVE a one-off report covering an additional 24 NSG operations that were repaid or prepaid between 2016 and 2019 and for which XSRs had never been prepared. The present report also discusses findings that emerged from OVE's review of this one-off report.

About half (53%) of the 62 validated SG operations achieved a positive overall outcome rating. The 53% positive rate falls short of the target of 70% set in the 2020–2023 Corporate Results Framework. As in previous cycles, relevance was the highest-rated core criterion among SG operations (with 76% rated positive). On effectiveness, 27% of SG operations achieved positive ratings. Negative effectiveness ratings were due to a combination of two factors—underachievement and poor monitoring and evaluation (M&E). OVE found that 43 of the 45 operations with negative effectiveness ratings were at least in part affected by underachievement, while 36 suffered from poor M&E, with most operations having been affected by both (34 of the 45). Multiple factors contributed to underachievement, though many PCRs did not sufficiently explain why outcomes were not achieved. With 59% of SG operations achieving a positive efficiency rating, the remainder were negative mostly due to the deficient quality of the ex post economic analyses, which did not allow to come to a clear conclusion on the efficiency of projects' resource use. On sustainability, 63% achieved positive ratings. Bank performance was rated positive for 55% of the operations, while the rest were affected by a combination of factors, most notably an overly ambitious project design coupled with poor M&E. Borrower performance was rated positive for 58% of operations.

Out of 47 NSG operations validated this cycle, 29 achieved positive overall outcome ratings. Aside from 18 validated operations with negative overall outcome ratings, 4 operations with EOM in 2015 or later and approved amounts of over US\$1 million exited the portfolio in 2020 from Special Assets and were assigned a negative overall outcome rating by OVE. Combined, the share of operations with a positive overall outcome rating is 57%, short of the target of 65% set in the 2020–2023 Corporate Results Framework. Among validated operations, relevance continues to be the highest-rated core criterion, with 87% of the operations achieving positive ratings. About half of the operations validated this year achieved positive effectiveness ratings; the other half were negatively rated due to failure to achieve the targeted results, as well as poor M&E that made it impossible to evaluate whether the objectives were achieved or whether results were attributable to the operation. Out of all NSG operations, 55% had

positive efficiency ratings and 62% achieved positive sustainability ratings. Among the noncore criteria, work quality remained the lowest-rated, considered “satisfactory” for a little less than half of operations. Positive investment outcomes were achieved by 85% of NSG operations, and another 85% demonstrated additionality.

Of 24 operations assessed through Mini-XSRs, 6 had a “satisfactory” overall outcome rating. In July 2022, IDB Invest Management submitted a one-off report on 24 operations repaid or prepaid between 2016 and 2019 for which XSRs had never been prepared, to ensure that all operations had been accounted for. Management stated that collecting data necessary for full analysis required for XSRs is difficult when significant time has passed since the repayment or prepayment and IDB Invest’s disengagement with the client. Accordingly, as an exceptional measure, OVE and Management agreed on a specific and simplified methodology—the Mini-XSR, which differs from the full XSR methodology. On the Mini-XSR, 6 of the 24 repaid or prepaid operations had a “satisfactory” overall outcome rating, 12 were rated “unsatisfactory,” and there was insufficient data to assign ratings for 6 operations. Half of the operations achieved an “unsatisfactory” investment outcome, while work quality was rated “unsatisfactory” in 20 out of 24 operations. Lessons learned related to project appraisal and supervision were in line with past evaluations and good practices; however, the report should have delved more deeply into lessons from prepaid operations, including further reflections on structuring and the possibility of prepayment penalties, as well as an overarching lesson associated with situations when a long time elapses between approval, commitment, and disbursement.

Trends over time show that overall outcome ratings have modestly improved for NSG operations, while no comparable time series data exist for SG operations. Comparability of performance data over time is essential for decision making. At the IDB Group, comparable project performance data are available for NSG operations over the past five years, given that the same evaluation methodology has been used since 2018. In contrast, for SG operations, repeated changes to PCR Guidelines have prevented the creation of time series data that are comparable over time. As a result, performance data for SG operations are comparable for only the last two validation cycles. Overall outcome ratings for SG operations were similar in 2021 and 2022, as gains in sustainability were offset by losses in the remaining criteria. NSG overall outcome ratings have improved only modestly in recent years, due to an unsteady trend in effectiveness despite clear improvements in all other core criteria.

Among SG operations, differences between OVE’s and Management’s ratings continued to widen. During the 2022 validation cycle, OVE rated the overall outcome positively in 33 (53%) of validated SG operations, but Management did so in 50 (81%) of them. Out of these

17 downgrades, all but 4 were affected by downgrades in more than one core criterion. Relevance, effectiveness, and efficiency each were downgraded in 10 cases, and sustainability in 9. There was no upgrade of overall outcome ratings from negative to positive. Fewer than half of the PCRs were of “satisfactory” quality. Over half of the PCRs were missing key information necessary to assess performance, which required Management to submit additional information and OVE to revisit evidence and reassess the relevant ratings.

Among validated NSG operations, OVE’s and Management’s overall outcome and core criteria ratings were very close. While OVE rated 62% of the validated NSG operations’ overall outcomes positively, Management rated 60% positively. OVE downgraded the outcome rating of one operation from positive to negative and upgraded two operations from negative to positive. The overall quality of XSRs was high, with over 80% achieving a positive quality rating. The majority of the nine XSRs rated “fair” had combinations of issues, including lack of or inaccuracies in key analyses, and insufficient or inaccurate information. In six cases, Management provided additional information during the feedback process that resulted in changes in core criteria ratings.

Considering these findings, OVE recommends the following.

**A. For IDB Management:**

1. **Strengthen PCR quality.** Over half of the PCRs reviewed by OVE this year had considerable shortcomings, and many of them lacked the information necessary to make an adequate assessment of project performance. Incompleteness in key information results in significant inefficiencies in the review process. It requires Management to respond to draft validations that were based on incomplete information. OVE then needs to assess additional information that Management submits subsequently and revise the validations, both of which take considerable time. In addition, while it is essential for IDB to systematically analyze why its operations may not achieve the targeted objectives, OVE found that in about half of PCRs for operations with negative effectiveness ratings, no or insufficient explanation was provided for underperformance on at least one outcome indicator. Quality of economic analysis also continues to be a key factor that drives down the quality of PCRs, even after Management developed sector-specific guidance notes on the preparation of cost-benefit analyses (CBAs) and cost-effectiveness analyses (CEAs) following OVE’s 2018 recommendation. The following actions are necessary for strengthening PCR quality:

- i. Strengthen Management's internal quality assurance to ensure that PCRs submitted to OVE include all key information needed to substantiate all ratings and are prepared in accordance with the PCR Guidelines.
- ii. Strengthen institutional learning by assessing and documenting in PCRs the reasons why operations underachieve.
- iii. Adequately implement the CBA and CEA guidance and include the underlying analysis in an Excel file for validation.
- iv. Conduct PCR training jointly with OVE.

**B. For IDB Invest Management:**

- 1. Complete another one-off report to fully catch up on operations repaid or prepaid between 2015 and 2019 without XSRs, focusing on lessons learned, and ensure that full XSRs are prepared for operations repaid or prepaid in 2020 and onward.** Management made significant progress in strengthening the system to identify projects subject to XSR each year, including updating the historical list of operations. OVE supports Management's proposal to prepare another one-off report in 2023 to complete catch-up reporting for the operations that reached EOM between 2015 and 2019. Given that the report will exclusively review the performance of repaid and prepaid operations, lessons on prepayments should include further reflections on structuring and the possibility of prepayment penalties.
- 2. Establish a system to ensure that all information necessary to prepare an XSR is collected at the time when a client indicates intention to prepay or repay a loan.** Given the difficulty in accounting for development results when substantial time has elapsed after repayment or prepayment and IDB Invest's disengagement with the client, relevant data should be collected systematically at the time at which the client indicates its intention to repay or prepay.
- 3. Set a time point after which reappraisal is mandatory when a long time has elapsed between approval, commitment, and disbursement; focus reappraisal on any material adverse development in the intervening period, as well as an update to the baselines for the relevant development results indicators.** Considering the findings of the one-off report that prepaid operations had a tendency to take longer to first disbursement, and that prepayments followed changes in market competitiveness and the clients' financial conditions, it is important to establish a process to ensure continued

alignment of project design with market conditions and client focus, thus improving the likelihood of operational success and smoother data collection.



**01**

Introduction

1.1 This report summarizes the results of the Office of Evaluation and Oversight's (OVE's) annual validation of Project Completion Reports (PCRs) and Expanded Supervision Reports (XSRs). PCRs are self-assessments of project performance by the Inter-American Development Bank (IDB); XSRs are their equivalent for IDB Invest. Both PCRs and XSRs are intended to serve as accountability and learning tools. As both PCR and XSR Guidelines<sup>1</sup> spell out, the accountability goal addresses the need for the IDB Group to ensure that project resources were used to achieve the objectives for which financing was granted. Guidelines also state that “the learning goal aims to replicate successes and avoid mistakes in the future by providing lessons to guide the execution of ongoing projects and the design of future ones.”

## **A. The IDB Group's evaluation methodology and the role of OVE**

1.2 The IDB Group's self-assessment methodology is objectives-based. Project performance is measured against four core criteria: (i) relevance of the project to the country's needs and realities and to the Bank's strategic objectives, (ii) effectiveness of the project in achieving the objectives stated at project approval, (iii) the efficiency with which project resources were used, and (iv) the sustainability of the results achieved. Each core criterion is rated on a four-point scale ranging from “unsatisfactory” to “excellent.” Based on the assessment of the core criteria, each project receives an overall outcome rating using a six-point scale ranging from “highly unsuccessful” to “highly successful.” In addition to the core criteria, the PCR and XSR Guidelines require the inclusion of noncore criteria, which are rated (on the same four-point scale as the core criteria) but do not count toward a project's overall outcome rating. Among these noncore criteria, PCRs include the performance of the Bank and its counterparts (borrowers), while XSRs include the financial and nonfinancial additionality of IDB Invest, the investment outcome, and IDB Invest's work quality. Annex I provides a full description of IDB's project evaluation framework.

1.3 OVE supports accountability and learning through the validation of completed PCRs and XSRs. OVE contributes to the credibility of the IDB Group's project performance reporting system through the independent validation of all PCRs and XSRs prepared by Management. Management assigns performance ratings to each completed project or program in the PCRs and to each operation reaching early operating maturity (EOM) in

---

<sup>1</sup> See PCR Guidelines, 2020 update (document [OP-1696-6](#)) and [2018 XSR Guidelines for IDB Group Private Sector Operations](#).

the XSRs, based on the relevant guidelines. OVE’s validation assesses whether these self-assessments are substantiated by evidence and are prepared in accordance with the respective guidelines. Based on this assessment, OVE assigns a final project overall outcome rating. These final ratings are reported by IDB Group Management in its Development Effectiveness Overview (DEO) report. This is the sixth consecutive year in which OVE has validated these self-assessments in the context of the IDB Group’s objectives-based project evaluation framework for public and private sector operations.

## **B. Scope of this report**

- 1.4 This report presents the results of the 2022 validation cycle, undertaken by OVE between September 2021 and May 2022. OVE reviewed and validated the PCRs of all 62 sovereign-guaranteed (SG) operations, as well as all 47 XSRs for non-sovereign-guaranteed (NSG) operations, prepared by Management. These operations represented US\$6.8 billion in SG approvals and US\$2.6 billion in NSG approvals. Annexes II and VI provide, respectively, the lists of SG and NSG operations reviewed, and Annex IV provides OVE’s validation notes for each SG operation. Most PCRs and XSRs assessed a single project, while some assessed a program covering a group of related projects with shared objectives. Therefore, throughout this report, the term *operation* refers to both projects and programs.<sup>2</sup>
- 1.5 The 2022 validation cycle also included OVE’s review of IDB Invest’s one-off report that accounted for the performance and documented lessons learned from NSG operations repaid or prepaid between 2016 and 2019 without prior XSRs. As discussed in OVE’s 2021 Validation Report,<sup>3</sup> Management had identified a group of operations that had been prepaid or repaid between 2016 and 2019, for which XSRs had never been prepared. During the 2022 validation cycle, Management prepared XSRs for 3 of those operations, about which IDB Invest continues to have engagement with the client. For 24 operations, on which IDB Invest no longer has engagement with the client, Management prepared “Mini-XSRs” and the one-off report that summarizes the findings from them. Annex VI provides the list of NSG operations reviewed in the Mini-XSRs.

---

2 The 62 SG operations reviewed covered 51 projects and 11 programs (73 projects in total). The 47 NSG operations reviewed covered 39 projects and 8 groups of projects with the same client assessed through joint XSRs (55 projects in total).

3 Document [RE-565-1](#), para 2.16.



1.6 The present report is organized as follows. Section II presents the results of OVE’s validations for SG and NSG operations in 2022, including the results of prepaid or repaid projects between 2016 and 2019 that were included in Management’s one-off report. Section III shows performance over time, and Section IV addresses the quality of the self-assessments. Section V presents conclusions and recommendations.



# 02

Results of  
Operations  
Validated in 2022

- 2.1 This section presents the results of the 2022 validation cycle. It includes the highlights of this year’s validations, overall project development outcome and core criteria ratings, and a summary of noncore criteria ratings. It ends with a review of IDB Invest’s one-off report on operations with Mini-XSRs.
- 2.2 For simplicity, the report groups ratings into positive and negative categories. A positive rating is one within the top three ratings on the six-point scale used for the overall project development outcome rating (i.e., “partly successful” or higher) or within the top two ratings on the four-point scale used for core and noncore criteria, and for rating the quality of a PCR or XSR (i.e., “satisfactory” or “excellent” for core and noncore criteria, “good” or “excellent” for the quality of the PCR or XSR). A negative rating falls within the ratings of “partly unsuccessful” or below for the overall project development outcome rating of the operation, “partly unsatisfactory” and “unsatisfactory” for the core and noncore criteria, and “fair” and “poor” for the quality of a PCR or XSR. Table 2.1 summarizes the rating scales.

**Table 2.1. Positive and negative rating categories**

Measure	Positive	Negative
Overall project development outcome	6 Highly successful	3 Partly unsuccessful
	5 Successful	2 Unsuccessful
	4 Partly successful	1 Highly unsuccessful
Core criteria (relevance, effectiveness, efficiency, sustainability)	4 Excellent	2 Partly unsatisfactory
	3 Satisfactory	1 Unsatisfactory
Noncore criteria (PCR: Bank performance, borrower performance; XSR: investment outcome, work quality, additionality)	4 Excellent	2 Partly unsatisfactory
	3 Satisfactory	1 Unsatisfactory
PCR/XSR quality	4 Excellent	2 Fair
	3 Good	1 Poor

Source: PCR Guidelines, 2020 update (document [OP-1696-6](#)) and 2018 XSR Guidelines.

## A. Public sector projects reviewed in 2022

- 2.3 OVE reviewed PCRs for 62 operations, of which 50 closed in 2020, 10 in 2019, and 2 in 2018. The PCRs for 59 of these operations were due this cycle, while the ones for the remaining 3 operations were past due from the previous cycle.<sup>4</sup> These PCRs covered 56

4 PCR Guidelines (document [OP-1696-6](#), Annex I) stipulate that PCRs for investment operations need to be submitted to OVE 6 months after operational closure (CO), while those for policy-based loans are expected to be submitted 30 months after CO. For programmatic operations, such as programmatic policy-based loans (PBPs) and conditional credit lines for investment projects financing dependent operations,

investment operations and six programmatic policy-based loan (PBP) series. Among the six PBP series, four were truncated after the first loan and two were completed series of two loans each. The 62 operations represented US\$6.8 billion in total approved amount. Table 2.2 breaks these operations down by sector.

**Table 2.2. SG operations by sector, 2022**

Department	SG operations validated	Approved amount (US\$ million)
Institutions for Development Sector (IFD)	22	3,645
Infrastructure and Environment Sector (INE)	15	1,312
Social Sector (SCL)	8	926
Climate Change and Sustainable Development Sector (CSD)	15	908
Integration and Trade Sector (INT)	2	34
<b>Total SG operations</b>	<b>62</b>	<b>6,825</b>

Source: OVE, based on data from IDB (2022).

Note: Operations approved as a program that share the same objectives and results matrix are evaluated in one PCR and thus counted as a single operation in this table.

## 1. Overall project development outcome ratings

2.4 Of the 62 validated SG operations, 33 (53%) achieved a positive overall outcome rating. The overall outcome rating is a weighted average of the four core criteria, where effectiveness has the highest weight.<sup>5</sup> This 53% positive rate falls short of the target of 70% set in the 2020–2023 Corporate Results Framework. The ratings of 55 operations (89%) fell in the middle four categories, “successful,” “partly successful,” “partly unsuccessful,” and “unsuccessful,” while only 2 (3%) were rated “highly successful” and 5 (8%) were rated “highly unsuccessful” (Figure 2.1).

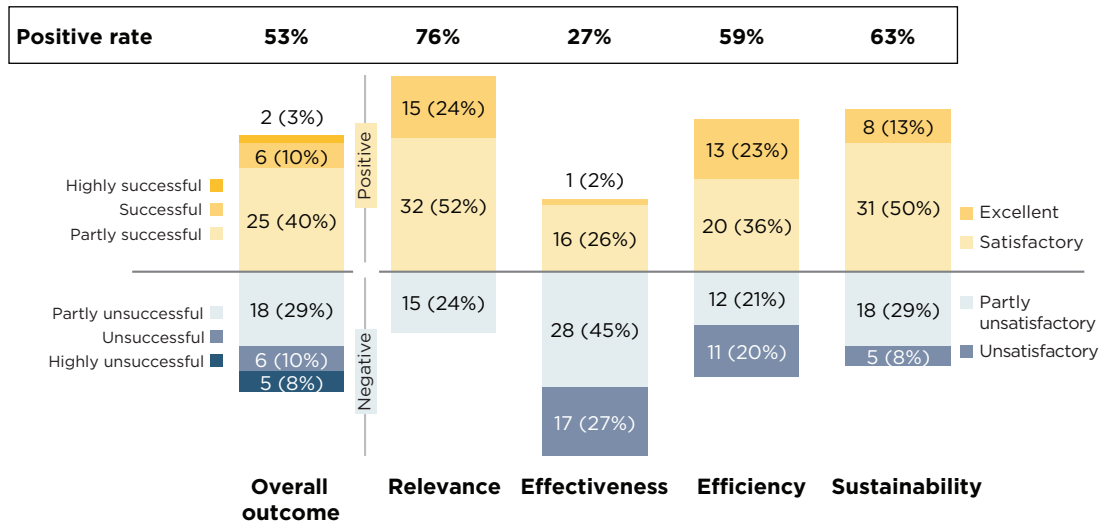
---

a single PCR is prepared for the entire series. The 50 operations closed in 2020 submitted their PCRs on time. Of the 10 operations that closed in 2019, 4 were part of a PBP series and 3 were part of a series of investment loans, so they met time expectations; and the remaining 3 were late submissions that should have been turned in last year (BO-L1093, HA-L1059, and PE-L1031). The 2 operations closed in the second half of 2018 were truncated PBPs, so their reports were submitted in line with the expected time frame.

5 For investment operations, effectiveness is weighted at 40%, while relevance, efficiency, and sustainability weigh 20% each. For policy-based loans, efficiency is not rated, so effectiveness is weighted at 60%, while relevance and sustainability weigh 20% each. See PCR Guidelines for further details.

**Figure 2.1**  
SG operations' ratings, 2022

Source: OVE.



Source: OVE.

Note:  $N = 62$  validated SG operations. Figures for efficiency add up to only 56 investment operations because the remaining 6 are PBP series, for which efficiency is not rated. See Annex III for detailed ratings by operation.

## 2. Project performance: Core criteria ratings

2.5 As in previous cycles, relevance was the highest-rated core criterion among SG operations. Out of the 62 validated SG operations, 47 (76%) were rated “satisfactory” or “excellent” on relevance. The remaining 15 operations achieved negative ratings due to a combination of not mutually exclusive factors. These include project designs that did not take into account country realities at approval (12), changes that weakened the original vertical logic during implementation (9), and finally, misalignment of objectives with IDB strategies either at design or during implementation without adequate justification of relevance to country development challenges (5).

2.6 Of the SG operations, 17 (27%) achieved positive effectiveness ratings, while 45 (73%) were rated negatively. Negative effectiveness ratings were due to a combination of two factors—underachievement and poor monitoring and evaluation (M&E). Underachievement means that there is evidence that the operation did not achieve most of its specific objectives. Underachievement can be further broken down by its key determinants into output shortfalls and outcome shortfalls, depending on where the chain breaks along the operation’s vertical logic. Output shortfalls occur when the project does not produce the planned outputs (or not to their full extent), and as a consequence the expected outcomes are not achieved. In turn, outcome shortfalls occur when the project produces the planned outputs but, in spite of this, the expected outcomes are not achieved, pointing to a broken link between the outputs and the outcomes that the

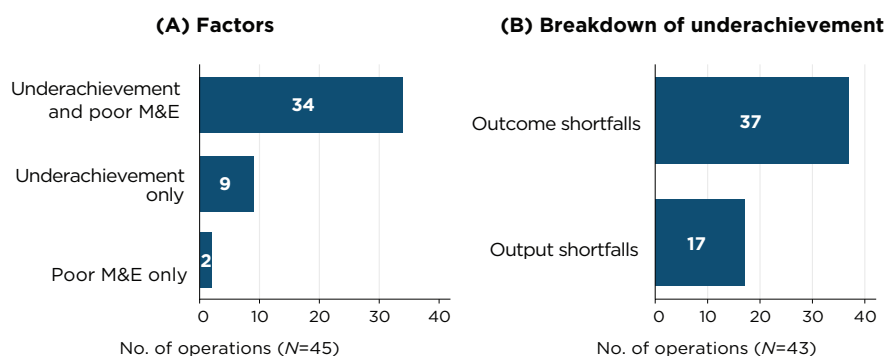
project had envisioned. The second factor, poor M&E, means that there is not sufficient evidence to demonstrate achievement of the project objectives due to measurement issues.

2.7 Underachievement was the key determinant of low effectiveness ratings, although most operations were affected by a combination of underachievement and poor M&E. OVE found that 43 of the 45 operations (96%) with negative effectiveness ratings were at least in part affected by underachievement, while 36 (80%) suffered from poor M&E, with most operations (34 of the 45) having been affected by both (Figure 2.2, graph A). Overall, though, while poor M&E was an important factor, it was second only to underachievement. Of the 43 operations affected by underachievement, 37 experienced outcome shortfalls even though they delivered the outputs, and 17 suffered from output shortfalls that then led to outcome shortfalls (Figure 2.2, graph B). Out of the 17 operations with output shortfalls, 10 had cancellations for a median amount of US\$8.4 million and a median of 26% of the original approved amount. In line with findings from previous OVE validation reports, large partial cancellations were strongly associated with the output shortfalls that led to underachievement.<sup>6</sup>

**Figure 2.2**

**Factors behind negative effectiveness ratings, SG operations**

Source: OVE's categorization, based on PCR contents and OVE's validation notes.



Note: In graph B, numbers do not add up to 43 because shortfalls are not mutually exclusive (26 operations faced exclusively outcome shortfalls, 6 exclusively output shortfalls, and 11 both types of shortfalls).

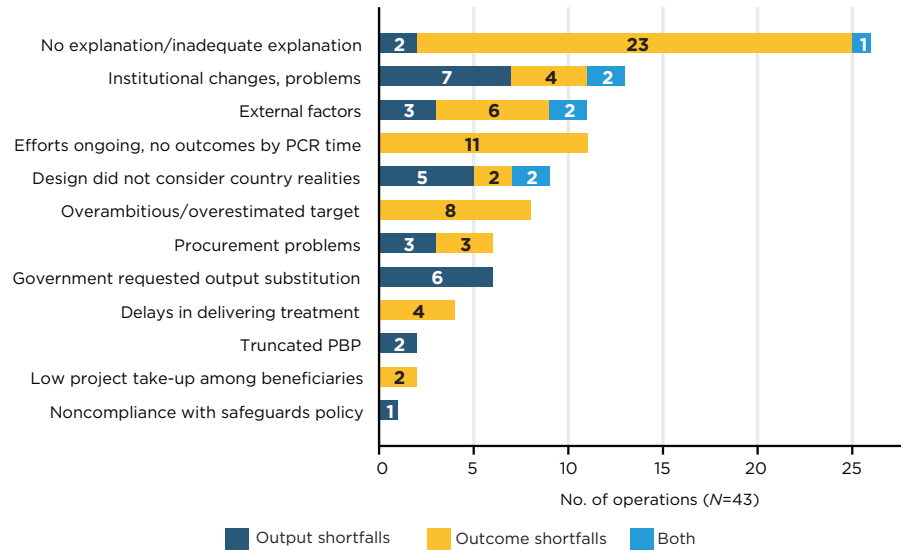
2.8 Multiple factors contributed to underachievement, though many PCRs did not sufficiently explain why outcomes were not achieved. OVE conducted a content analysis of the explanations provided by Management in the PCRs for why the operations faced underachievement. Figure 2.3 presents a summary of these explanations. The three most commonly cited reasons for underachievement were institutional changes or problems (e.g., a shift in government priorities or changes in country laws and regulations; or administrative, technical, and budget problems

6 Among the 45 negatively rated operations, those that had large partial cancellations ( $\geq 20\%$  of the original approved amount) were associated with a 40-percentage-point increase in the probability of having an output shortfall that led to underachievement (linear probability model using robust standard errors, p-value < 0.05, R2 = 0.089).

that affected project delivery); external factors, such as Covid-19, macroeconomic and market shocks, increased violence in project areas, and others; and that efforts were still ongoing, but they had not materialized into outcomes by the time the PCR was prepared.<sup>7</sup> Importantly, however, over half of PCRs for underperforming operations did not provide an adequate explanation of why the operations faced underachievement in at least one of their outcome indicators,<sup>8</sup> which suggests a missed opportunity for institutional learning.

**Figure 2.3**  
Reasons behind underachievement according to PCRs, SG operations

Source: OVE's categorization, based on PCR contents.



Note: Numbers do not add up to 43 because reasons are not mutually exclusive.

2.9 Poor M&E encompassed various types of issues. The most common types of M&E issues were unreliable data (17 out of 45 negatively rated operations) and attribution problems or failure to establish the project's plausible contribution to the observed results (15) (Figure 2.4). Other issues included missing data (13), use of indicators that were inadequate for accurately measuring achievement in the specific objectives that they were expected to measure (13), use of indicators that reflected outputs but not outcomes (5), and insufficient indicators for measuring all relevant dimensions of the specific objective in question (4).

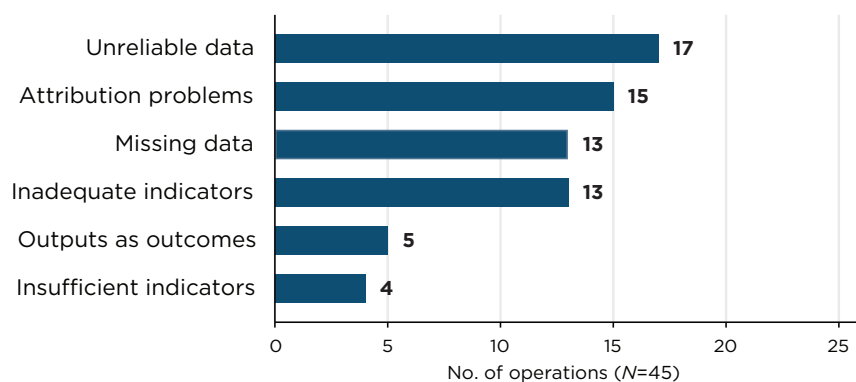
7 While the unit of account for effectiveness is the project, a similar analysis was also conducted to see what factors account for the lack of achievement of individual objectives. The results are similar to those at the project level.

8 Twenty-four of 37 (65%) operations with outcome shortfalls and 3 of 17 (18%) operations with output shortfalls did not explain why they faced underachievement in at least one of their outcome indicators, or provided only explanations that OVE found inadequate given that they were not supported by the evidence presented in the PCR.

**Figure 2.4**

**Types of M&E issues, SG operations**

Source: OVE.



Note: Numbers do not add up to 45 because the M&E issues are not mutually exclusive.

2.10 Among operations validated in 2021 and 2022, there was no statistically significant variation in the likelihood of achieving various types of objectives. OVE analyzed the specific objectives of the 133 SG operations validated in 2021 and 2022.<sup>9</sup> A total of 340 specific objectives were categorized into five types of objectives: 134 (39%) aimed to provide better services, 106 (31%) to build institutional capacity, 22 (6%) to facilitate sector-level changes, 18 (5%) to improve environmental sustainability, and 14 (4%) to increase financing. The remaining 46 objectives (14%) were classified as “other”<sup>10</sup> (Table 2.3). Objectives related to improving environmental sustainability had the highest share of positive ratings (50%), while those affecting sector-level changes had the lowest share (27%), with the caveat that differences are not statistically significant<sup>11</sup> and that the number of observations for some types of objectives is very small.

**Table 2.3. Types of objectives, SG operations validated in 2021 and 2022**

Type of objective	Number of objectives <sup>a</sup>	Percentage positive	Breakdown (number of objectives)
Improved environmental sustainability	18	50	Includes objectives to minimize diverse types of negative environmental impacts (7), reduce greenhouse gas emissions (5), promote conservation (4), and control pollution (2).
Better service provision	134	44	Includes objectives related to improving quality of services (45), <sup>b</sup> improving service provision through better physical infrastructure (41), expanding access to services (28), reducing their cost (15), and other types of improvements (5).

9 While this report’s main focus is the performance of operations validated in 2022, the analysis of specific objectives incorporates data from operations validated in 2021 to expand the number of observations.

10 “Other” objectives are those with fewer than 10 observations.

11 While some types of objectives were achieved at higher rates than others, the differences between the types are not statistically significant. The differences between each type of objectives and the rest of all objectives combined are also not significant. These tests are based on linear probability models using robust standard errors.



Type of objective	Number of objectives <sup>a</sup>	Percentage positive	Breakdown (number of objectives)
Increased financing	14	43	Includes objectives related to increased access to credit (8), expanded financing (4), backing up credit guarantees for micro, small, and medium enterprises (1), and increased availability of financing (1).
Institutional capacity building	106	42	Includes objectives related to building capacity of the government (53), of other stakeholders (30), and of service providers (23). <sup>c</sup>
Sector-level changes	22	27	Includes sector-level changes oriented toward developing a legal, policy, or strategic framework for tourism diversification (7), infrastructure sectors (4), digital transformation (3), business environment (2), education (1), improvement of financial intermediation (1), inclusion (1), social security (1), developing an institutional and executive framework for a less fragmented judicial defense in the country (1), and competitiveness (1).
Other <sup>d</sup>	46	48	Includes other types of objectives related to behavioral changes, citizen security improvement, foreign trade, fiscal sustainability, increased private investment, resettlement, disaster prevention, and sustainable development.

Source: OVE.

Notes: <sup>a</sup>Breakdown of number of objectives by 2021 and 2022, respectively, is as follows: 7 and 11 for improved environmental sustainability, 57 and 77 for better service provision, 9 and 5 for increased financing, 64 and 42 for institutional capacity building, 13 and 9 for sector-level changes, and 13 and 33 for “other.” <sup>b</sup>Four of these objectives also aim to expand coverage. <sup>c</sup>When capacity building is provided to public entities to strengthen their public policy functions (e.g., sector coordination, policymaking), objectives are classified as capacity building for the government. When it is provided to public entities to strengthen their ability to provide services directly to end beneficiaries (e.g., public utilities), objectives are classified as capacity building for service providers. <sup>d</sup>Fewer than 10 objectives under a single category.

2.11 Of 56 investment operations, 33 (59%) achieved a positive efficiency rating,<sup>12</sup> and the remaining 23 obtained a negative efficiency rating mostly due to lack of credibility of the ex post economic analyses. PCRs for investment operations are expected to demonstrate how efficiently a project has used its resources through either a cost-benefit analysis (CBA) or a cost-effectiveness analysis (CEA). In exceptional circumstances, where such an analysis is not possible, the PCR needs to justify why. When a CBA or CEA is not presented and there is no credible justification for the lack of the analysis, a project cannot achieve a positive efficiency rating, and cost and time overrun information is used to determine the efficiency rating. If OVE finds that the CBA or CEA presented lacks credibility and thus cannot be relied upon to rate efficiency, it also reverts to cost and time overrun information.<sup>13</sup> Furthermore, the PCR Guidelines state that for a positive efficiency rating, aside from the CBA or the CEA pointing to satisfactory efficiency, the evaluator also needs to ascertain

12 Efficiency is not rated for policy-based loans per PCR Guidelines.

13 If there is no credible justification for a missing CBA or CEA, and cost and time overrun information from Progress Monitoring Reports do not point to any alert or problem status, the project’s efficiency rating is “partly unsatisfactory”; if the project was in alert or problem status, the efficiency rating is “unsatisfactory.” The same rules apply when a CBA or CEA is rejected for lack of credibility.

that there are no other factors that may have affected efficiency negatively. For 14 of the 23 projects with negative efficiency ratings, OVE determined that the CBA or CEA undertaken in the PCR lacked credibility,<sup>14</sup> and for 4 projects the PCR presented no CBA or CEA when such an analysis could have been expected. Many of these operations were also affected by other factors that reduced efficiency, such as implementation delays. For 3 other operations, the economic analysis presented pointed to inefficient resource use, and in 1 case where such an analysis was deemed not possible, cost and time overrun information led to a negative rating. Finally, 1 operation was rated negatively on efficiency because other factors pointed to suboptimal efficiency of the investments.

2.12 Out of 62 operations, 39 (63%) achieved positive sustainability ratings. Risks to continuation of results of these operations were minor or nonexistent. Three factors led to negative sustainability ratings of the other 23 operations: uncertain continuity of results, unresolved issues around environmental and social (E&S) safeguards, and insufficient documentation to ascertain that risks to continuity of results were adequately mitigated. Uncertainty in the continuity of achieved results affected 19 of these 23 operations, due to a combination of not mutually exclusive factors such as inadequate resources for maintenance (13); the macroeconomic and political environment (8), including the effects of Covid-19 and subsequent reductions in the government's fiscal space (4); low institutional capacity (7); and low ownership and commitment (6). For 3 additional projects, the PCR did not provide sufficient information to ascertain that risks to continuity of results were adequately mitigated. Operations affected by safeguards performance issues totaled 13, including 5 for which E&S safeguards issues were significant or remained unresolved at project closure, and another 8 for which the PCR failed to report whether E&S issues that had arisen during implementation had been adequately resolved at project closure.

### **3. Project performance: Noncore criteria ratings**

2.13 Bank performance was rated positive for 34 operations (55%), while the remaining 28 were affected by a combination of factors, most notably an overly ambitious project design coupled with poor M&E.<sup>15</sup> The Bank was successful in both ensuring quality at entry and providing quality supervision for the 34 operations

---

14 Credibility issues included: inclusion of benefits that were not achieved or not attributable to the project, assumptions on benefits or the continuation of benefits not supported by evidence, failure to account for a significant share of project costs (including relevant operating costs), and lack of a detailed enough presentation of the assumptions and calculations of the analysis to verify the results.

15 Out of 28 operations with negative Bank performance, 10 suffered from this combination.

rated positive. Among those with negative ratings, shortcomings were observed during the entry phase of 26 operations and during the supervision of 18 operations. At entry, overambitious project design that was misaligned with the technical and institutional capacities of the executing units contributed to a negative Bank performance rating for 16 operations (57%). Negative Bank performance ratings were also related to weak M&E design (14),<sup>16</sup> inadequate up-front analysis and project design (14),<sup>17</sup> and shortcomings in risk identification and mitigation (13). During the supervision phase, poor M&E<sup>18</sup> affected Bank performance in 15 operations (54%). Other supervision shortcomings included lack of proactiveness to address challenges in a timely manner during implementation (9), high turnover of the team leaders (4), and failure to ensure safeguards compliance (7).<sup>19</sup>

2.14 Borrower performance was rated positive for 36 operations (58%). Among the remaining 26 operations where borrower performance was rated negative, 21 operations were affected by lack of commitment and readiness of the executing units as well as inadequate staffing and high turnover. Challenges complying with contractual agreements were mainly related to M&E weaknesses (12) and failure to fulfill fiduciary responsibilities (8). For 12 operations (46%), borrowers failed to substantially comply with E&S safeguards. Administrative challenges, such as excessive time required for approval and complex administrative procedures, affected 8 operations (31%), while inadequate involvement of other stakeholders affected 7 operations (27%).

## **B. Private sector projects reviewed in 2022**

2.15 Management prepared and OVE validated XSRs for 47 operations in 2022. Of these, 31 reached EOM<sup>20</sup> in 2020 or exited the portfolio in 2020 before reaching EOM. In turn, for 12 operations, Management submitted catch-up XSRs that had not been completed in previous years; 6 of these operations reached EOM in 2019, 3 in 2018, and 3 in 2016.<sup>21</sup> The remaining 4 reached EOM

16 Inadequate definition of indicators and targets was the most common deficiency.

17 This includes flawed diagnosis of country challenges, unclear vertical logic, and an inadequate instrument selection.

18 M&E weaknesses included deficiencies in the registration and reporting of information as well as failure to conduct relevant analyses and evaluations.

19 Factors are not mutually exclusive.

20 EOM means that IDB Invest made material disbursement and received at least one set of audited financial statements covering 12–36 months of the operation after disbursement/project completion. The exact definition of EOM varies by project type (see XSR Guidelines).

21 The 12 operations that reached EOM during 2016–2019 (and thus had past-due XSRs) were only recently identified by Management, as part of its ongoing efforts to strengthen the EOM calculation system.

in 2021.<sup>22</sup> The operations were approved between 2009 and 2018, and most of them (29) were approved in 2016 or after, reflecting post-Merge-Out, IDB Invest operations.<sup>23</sup> The 47 operations covered 40 senior loans,<sup>24</sup> 4 equity operations, and 3 guarantees. Jointly, these operations represent US\$2.6 billion in approvals. Table 2.4 breaks this amount down by business segment.

**Table 2.4. NSG operations by business segment, 2022**

Business segment	NSG operations validated	Approved amount (US\$ million)
Infrastructure and Energy	17	1,055
Corporates	12	883
Financial Institutions	18	674
<b>Total NSG operations</b>	<b>47</b>	<b>2,612</b>

Source: OVE, with data from IDB Invest (2016, 2022b).

Note: Approvals data up to December 31, 2021.

2.16 OVE also included in this exercise 4 NSG operations that exited the portfolio from Special Assets in 2020 and were lacking an XSR. Following OVE’s 2021 Validation Report recommendations, Management shared the list of operations that exited the portfolio from Special Assets during 2020. Starting with this validation exercise, OVE will report on such operations for accountability purposes. In the absence of an XSR, OVE assigned a negative overall outcome rating to all 4 operations. These 4 operations reached EOM in 2015 or later and were approved for amounts of over US\$1 million.<sup>25</sup> Two of the operations were in the infrastructure and energy (I&E) segment with a combined approved amount of US\$82 million, and the other two in the corporate segment with a combined approved amount of US\$9 million.

### 1. Overall project development outcome ratings

2.17 Of 47 NSG operations validated this cycle, 29 achieved positive overall outcome ratings. The ratings of 89% of operations fell in the middle four categories of “successful,” “partly successful,” “partly unsuccessful,” or “unsuccessful”; only 1 operation was

22 Three of the four operations were groups of two projects for which IDB Invest had initially identified only the first project, which had an earlier EOM year, prompting IDB Invest to prepare the XSR. However, during XSR preparation, the second project with EOM 2021 was identified, and IDB Invest decided to continue preparing the XSR covering both projects. In the fourth case, IDB Invest perceived a risk that information could not be obtained from the client at a later point, so it conducted the XSR early. EOM data provided to OVE by IDB Invest (2022d). In cases of XSRs covering multiple jointly assessed projects, OVE uses the EOM year of the project with the latest EOM.

23 Breakdown of the approval years are as follows: 1 (2009), 1 (2012), 3 (2013), 5 (2014), 8 (2015), 8 (2016), 17 (2017), 4 (2018). In cases of several operations assessed as a program, the earliest approval date is considered.

24 Of the 40, 34 were stand-alone senior loans, while 2 were combined with subordinated loans and 2 with guarantees, and 2 were Trade and Supply Chain Finance loans.

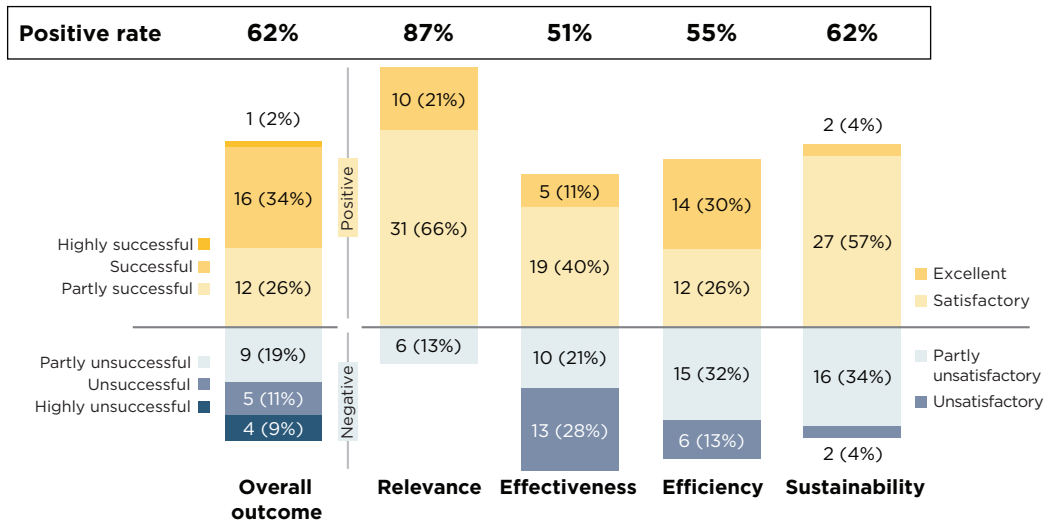
25 XSRs are not prepared for operations approved for less than US\$1 million.

rated “highly successful” and 4 were rated “highly unsuccessful” (Figure 2.5). Aside from 18 validated operations with negative overall outcome ratings, 4 operations with EOM in 2015 or later and approved amounts of over US\$1 million exited the portfolio in 2020 from Special Assets and thus were assigned a negative overall outcome rating by OVE. Therefore, the share of operations with positive overall outcome ratings is 57%, short of the target of 65% set in the 2020–2023 Corporate Results Framework.

**Figure 2.5**

**NSG operations’ ratings, 2022 (validated operations only; excludes operations that exited the portfolio from Special Assets)**

Source: OVE.



Note: N = 47 validated NSG operations. See Annex VII.

## 2. Project performance: Core criteria ratings

2.18 Relevance continues to be the highest-rated core criterion, with 41 operations (87%) achieving positive ratings. Objectives of all operations validated this year were aligned with country needs and the relevant IDB Group strategies. Project designs of 41 operations with positive ratings were also aligned with country priorities and IDB Group strategies. Among the 6 operations (13%) that obtained negative relevance ratings, 5 had shortcomings in project design such as currency mismatch (3), inadequate selection of partner (2), and inadequate selection of instrument (1). One operation had shortcomings in its vertical logic.<sup>26</sup>

2.19 About half of the operations validated this year achieved positive effectiveness ratings. These 24 operations achieved more than half of their objectives, attribution of outcomes to the project was plausible, and none of their objectives were rated “unsatisfactory.” Of 23 operations with negative effectiveness ratings, 21 (91%) failed to achieve the targeted results, and 6 (26%) were affected by poor M&E that made it impossible to evaluate whether the objectives were achieved or attributable to the operation.

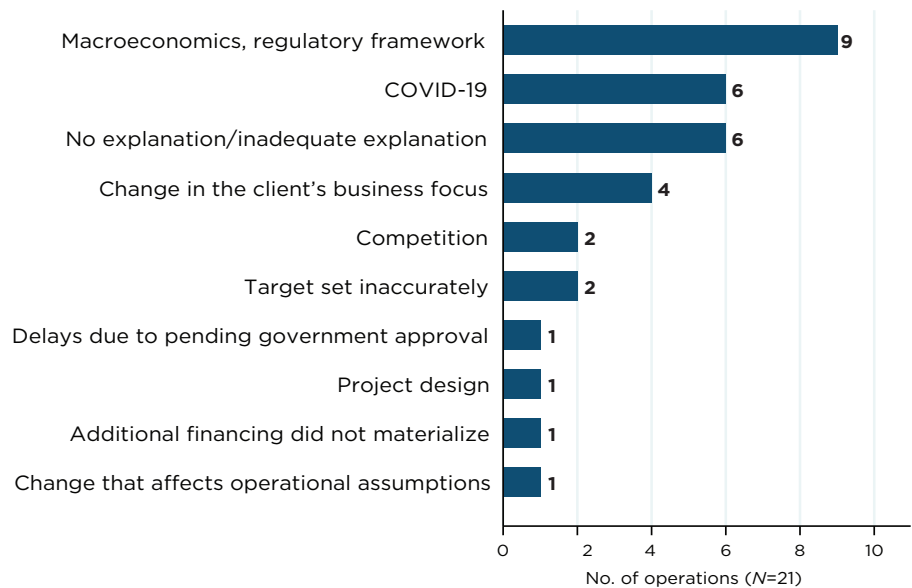
<sup>26</sup> Factors are not mutually exclusive.

2.20 XSRs cited macroeconomic and regulatory issues most frequently as the reasons for underachievement, though the depth of analysis varied across XSRs. Among the 21 operations for which underachievement contributed to negative effectiveness ratings, XSRs cited macroeconomic shock and regulatory framework most frequently (9 XSRs) as contributing factors. The negative impact of Covid-19 (6), change in clients' business focus (4), strong competition in the market (2), and an inaccurately set target (2) were also mentioned, among others (see Figure 2.6). The depth of analysis of the underlying cause of underachievement varied among XSRs. Six XSRs provided either no or inadequate explanations for underachievement.

**Figure 2.6**

**Reasons behind underachievement according to XSRs, NSG operations**

Source: OVE's categorization, based on XSR contents.



Note: Numbers do not add up to 21 because reasons are not mutually exclusive.

2.21 Among operations validated in 2021 and 2022, objectives on increasing renewable energy generation and reducing greenhouse gas (GHG) emissions were the most likely to be achieved, while objectives on improving financing for non-MSME clients were the least likely to do so. OVE analyzed the specific objectives of the 85 NSG operations validated in 2021 and 2022.<sup>27</sup> A total of 190 specific objectives were categorized into five types: 47 (25%) were objectives aimed at supporting the development of micro, small, and medium enterprises (MSMEs); 36 (19%) at reducing GHG emissions; 26 (14%) at increasing generation of renewable energy; 19 (10%) at improving financing for non-MSME clients; and 17 (9%) at creating jobs. The remaining 45 objectives (24%) were classified as “other,” as they fell under types of objectives with fewer than 10 observations. Objectives with the highest share of positive ratings were those

<sup>27</sup> While this report's main focus is on the performance of operations validated in 2022, the analysis of specific objectives incorporates the data from operations validated in 2021 to expand the number of observations.



related to generating renewable energy (73%) and reducing GHG emissions (72%). MSME support objectives were 38% positive, while those on job creation and improving financing for non-MSME clients were the least positive, with 24% and 16%, respectively (Table 2.5). Statistical significance tests largely confirmed these findings.<sup>28</sup>

**Table 2.5. Types of objectives, NSG operations**

Type of objective	Number of objectives	Percentage positive	Breakdown (number of operations)
Increase generation of renewable energy	26	73	All objectives aimed at increasing renewable energy generation through added capacity (26).
Reduce GHG emissions	36	72	Includes reduction of GHG emissions from expanded use of renewable energy (27), improved energy efficiency (8), and reduction of emissions through import substitution (1).
Support MSME development	47	38	Includes support for MSME development through improved access to financing (38), support for MSMEs through the value chain (1), support for MSMEs with E&S goals (1), knowledge transfer to MSMEs (1), and other support (6).
Create jobs	17	24	All objectives were aimed at creating employment (17), mostly during the construction phase of operations.
Improve financing for non-MSME clients	19	16	Includes financing to support housing (7), green portfolios (4), low-income populations (3), women (2), housing improvements (2), and mortgages (1).
Other	45	53	Increased productivity through better and lower cost of infrastructure (9), increased production in priority sectors (9), increased exports (7), sector-level changes (4), reduced cost of energy (3), increased access to services (3), improved physical infrastructure (3), corporate social responsibility (2), and others (5).

Source: OVE.

Note: Breakdown of number of objectives by 2021 and 2022, respectively, is as follows: 11 and 15 for increasing generation of renewable energy, 16 and 20 for reducing GHG emissions, 21 and 26 for supporting MSME development, 13 and 4 for creating jobs, 4 and 15 for improving financing for non-MSME clients, and 25 and 20 for other.

2.22 Of the 47 NSG operations, 26 (55%) had positive efficiency ratings. Financial and economic benefits of these 26 operations exceeded project costs, taking into account the time-value of money. Of the

28 On average, when compared against the rest of all objectives combined, GHG emissions objectives and renewable energy objectives were 28 and 27 percentage points more likely to be achieved, respectively, while non-MSME objectives were the least likely to be achieved (37 points less likely). When comparing each type of objective against the other types, GHG emissions and energy objectives were also more likely to be achieved than any other type—between 34 and 57 percentage points more likely, depending on the comparator. Importantly, MSME objectives were 23 points more likely to be achieved than non-MSME objectives. These findings are restricted to operations validated in 2021 and 2022, and interpretation should consider the small number of observations for some types of objectives. Results are based on linear probability models using robust standard errors; all point estimates mentioned were statistically significant at least at the 95% level.

21 operations rated negative, 14 were with financial institutions (FIs), 4 were in I&E, and 3 were corporate operations. These 21 operations were affected by lower-than-expected growth in their target portfolios or sub-borrowers (12), followed by insufficient information that prevented a full efficiency assessment (9), financial or economic returns lower than the cost of capital—or threshold return for equity projects—(8), and higher-than-expected nonperforming loan (NPL) rates (7).<sup>29</sup> These factors affected both financial and economic efficiency in most cases (17 of 21). Of the 21 operations, 12 (57%) were affected by a single factor, with low financial or economic returns being the most common (7). The remaining 9 (43%) were affected by a combination of two or three of these reasons. In addition, 7 of the 21 operations (33%) were reportedly affected by Covid-19, namely through NPL rates (4 operations), given the difficulties the pandemic imposed on sub-borrowers to repay; the financial or economic returns of the clients (2); and the lower-than-expected growth of the target portfolio or sub-borrowers (1).

2.23 Of the 47 NSG operations, 29 (62%) achieved positive sustainability ratings. For these 29 operations, IDB Invest clients are in material compliance with E&S requirements, and the project results already achieved are likely to continue. Among the 18 operations with negative sustainability ratings, 16 received low scores because of the fragility of results continuation. Of these, 13 operations (72%) were affected by project-specific issues such as a shift in the client's focus away from the target portfolio (e.g., lending to small and medium enterprises) and uncertainty about continued financing sources. Macroeconomic and political issues risked compromising the sustainability of 9 operations' results. In no project was the Covid-19 pandemic considered the primary cause of sustainability concerns. E&S compliance shortfalls affected the sustainability of 7 operations (38%), including 2 in which the inability to ensure E&S compliance of the clients' investees or suppliers affected the rating.

### **3. Project performance: Noncore criteria ratings**

2.24 Of the 47 NSG operations, 40 (85%) achieved positive investment outcomes and 40 (85%) demonstrated additionality. Gross profit contribution of these 40 operations exceeded the rating benchmarks set at approval for minimally expected performance. Among the 7 operations validated in 2022 that did not achieve a positive investment outcome, 4 (57%) were equity investments for which IDB Invest received either negative or lower-than-expected investment returns and was not able to exit as planned, 2 (29%) were senior

---

<sup>29</sup> The breakdown by business segment is as follows. Of the 12 operations with lower-than-expected growth in target portfolios or sub-borrowers, 10 were FI operations and 2 were corporate operations (both of them companies offering financial services). Of the 10 operations with higher-than-expected NPLs, 5 were FIs and 2 were corporate operations (the same two companies offering financial services). Of the 9 operations with insufficient information, 7 were FIs and 2 were corporate operations. Of the 8 with low returns, 4 were in I&E, 3 were FIs, and 1 was a corporate operation.

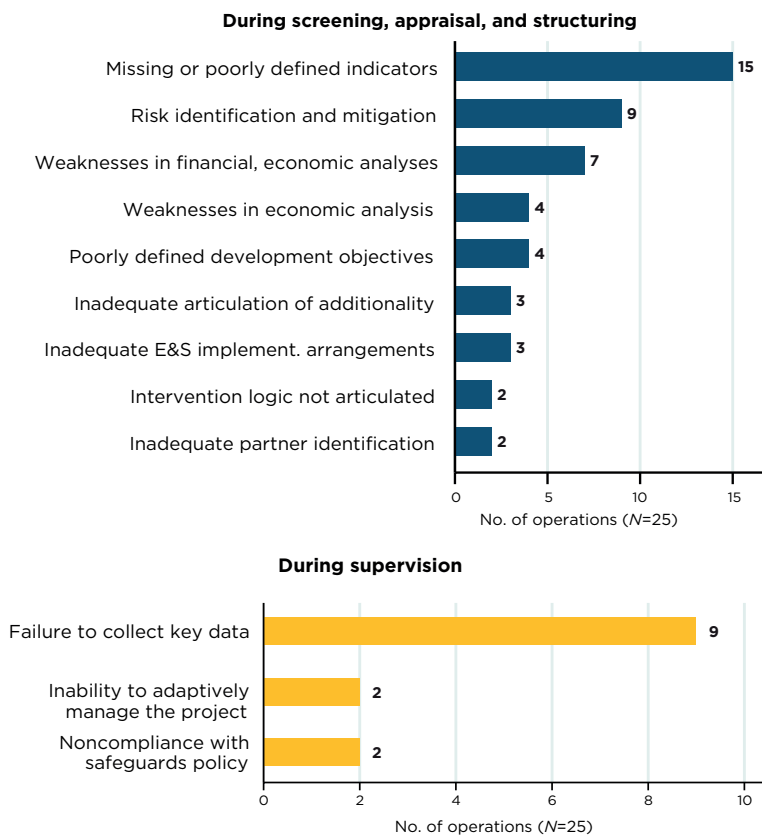


loans that were repaid, and 1 (14%) was a senior loan that is at risk of default. Among the 40 operations with positive additionality, 39 demonstrated financial additionality and 29 nonfinancial additionality. Conversely, among the 7 with negative additionality ratings, 5 did not deliver financial additionality and 6 did not deliver environmental, social, and governance additionality.

2.25 Work quality remained the lowest-rated noncore criterion, considered satisfactory for a little less than half of operations. IDB Invest’s screening, appraisal, and structuring, as well as monitoring and supervision following commitment, were of a high standard, and all procedures and good practice standards were met for the 22 operations rated positive. The 25 operations that did not achieve positive work quality ratings suffered from a combination of shortcomings during screening, appraisal, and structuring, including missing or poorly defined indicators (15), shortcomings in identification and mitigation of risks (9), and weaknesses in financial and economic analyses (7). There were 10 operations that also suffered from shortcomings during supervision. Failure to collect key data most frequently affected the quality of supervision (9), followed by inability to adapt to changing environments (2) and noncompliance with E&S safeguards (2) (Figure 2.7).

**Figure 2.7**  
Factors behind negative work quality ratings, NSG operations

Source: OVE.



Note: Numbers do not add up to 25 in either graph because shortcomings are not mutually exclusive.

## C. Review of Management report on prepaid and repaid NSG projects between 2016 and 2019

2.26 In July 2022, IDB Invest Management submitted a one-off report on 24 operations that repaid or prepaid between 2016 and 2019 and for which XSRs had never been prepared. As discussed in OVE's 2021 Validation Report,<sup>30</sup> Management had made significant efforts to improve the system to identify operations requiring an XSR in a given year, which led to the identification of repaid or prepaid operations that should have had but did not have any XSRs during the previous validation cycles. To ensure that all operations are accounted for, Management prepared XSRs during the 2022 validation cycle for those operations for which IDB Invest continues to have a relationship with the relevant client. For 24 operations for which IDB Invest no longer has any engagement with the client, Management and OVE agreed on a simplified methodology for project reviews with a significantly reduced scope, the "Mini-XSR." Management prepared a one-off report summarizing the findings from the 24 Mini-XSRs, which OVE reviewed as part of this year's validation exercise (Annex VI provides the list of operations with Mini-XSRs). The results of Mini-XSRs were not reported to the Board in Management's DEO report<sup>31</sup> earlier this year because the Mini-XSR rating methodology was still under discussion between Management and OVE at the time and the validated results became available only after the DEO publication.

2.27 During 2022, Management identified 35 additional repaid or prepaid operations with missing XSRs, for which Mini-XSRs will be prepared next year. Following OVE's recommendation in the 2021 Validation Report,<sup>32</sup> in 2022 Management systematically reviewed the full list of operations that should have had XSRs in previous cycles, which was then reviewed by OVE. This led to the identification of an additional 35 operations that had been repaid or prepaid between 2015 and 2019, for which IDB Invest no longer has any engagement with the relevant clients.<sup>33</sup> Management has proposed, and OVE agreed, that an additional one-off report be prepared in 2023 to cover these 35 operations. The majority (20) of the 35 operations reached EOM in 2015,<sup>34</sup> while 2 operations each reached EOM in 2016, 2018, and 2019, and 5 in 2017.

---

30 Document [RE-565-1](#), para 2.16.

31 Document [GN-3111-1](#).

32 Document [RE-565-1](#), para 5.9.

33 For operations that were repaid or prepaid during the same period that IDB Invest continues to have engagement with the clients, Management plans to prepare full XSRs in the 2023 validation cycle.

34 Given that the current evaluation framework has been applied to operations that reached EOM in 2015 or later, operations that reached EOM in 2014 or earlier that never produced XSRs are exempt from catching up.

2.28 The Mini-XSR differs from the regular XSR in several critical ways and is expected to draw lessons specifically from repaid and prepaid operations. Management stated that collecting data necessary for the full analysis required for XSRs is difficult when significant time has passed since the repayment or prepayment and IDB Invest's disengagement with client. Accordingly, as an exceptional measure, OVE and Management agreed on a specific and simplified methodology for the Mini-XSR that differs from the XSR methodology (see Annex VIII for details on the Mini-XSR methodology). Specifically, Mini-XSRs differ from regular XSRs in three critical aspects. First, while XSRs rate the overall outcome on a six-point scale and underlying criteria on a four-point scale, Mini-XSRs use a two-point scale of "satisfactory" or "unsatisfactory." In view of the exceptional circumstances when a relationship with a client was severed several years before preparation of the Mini-XSR and no further data can be collected, operations may also be rated as "lack of data" when justified. Second, the discussion does not cover in-depth analysis of each criterion. Third, the Mini-XSR does not assess the additionality of the operation.<sup>35</sup> The Mini-XSR and the summary report are expected to elaborate lessons learned specifically from repaid and prepaid operations. Given the differences in methodology, the ratings of these operations cannot be combined with those evaluated through full XSRs.

2.29 The majority of the 24 operations repaid or prepaid without previous XSRs were in the corporate and FI segments. Of all 24 operations, 12 operations were with FIs, 11 were corporate operations, and 1 was in I&E. Except for 1 guarantee operation, all were senior loans. Combined, they represented \$316 million in approvals. Prepaid operations totaled 15, while 9 were repaid. On average, the operations were prepaid and repaid 2.5 years and 3.1 years, respectively, after their first disbursement.

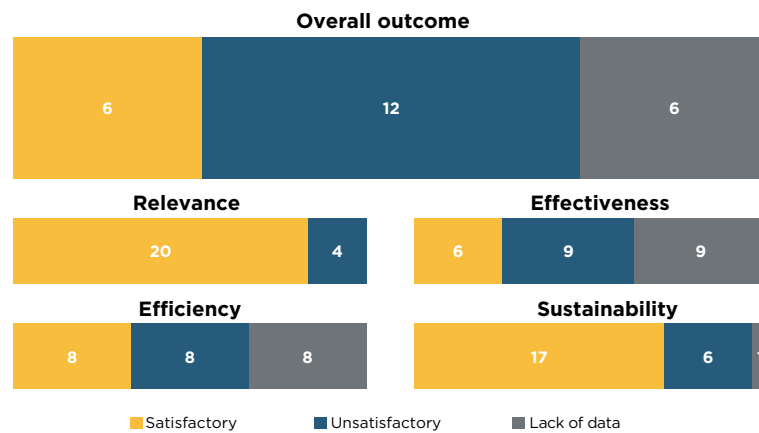
2.30 Of the 24 operations, 6 (25%) had "satisfactory" overall outcome ratings. There were 12 operations (50%) rated "unsatisfactory," and there was insufficient data to assign a rating for 6 operations (25%). Figure 2.8 shows the number of operations rated "satisfactory," "unsatisfactory," and "lack of data" for overall outcome and for each of the core criteria.

---

<sup>35</sup> Mini-XSRs are meant to focus on assessing outcome and core criteria. However, among three noncore criteria in the XSR Guidelines, investment outcome is assessed to understand the impact of repayment and prepayment on profitability. Work quality is also assessed for its potential link with repayment/prepayment.

**Figure 2.8**  
**Ratings of NSG**  
**operations in repaid/**  
**prepaid report, 2022**  
**(number of operations)**

Source: OVE.



Note: N = 24 operations.

2.31 Of the operations in the one-off report, 20 (83%) achieved “satisfactory” relevance ratings while 6 (25%) had “satisfactory” effectiveness ratings. The four operations with “unsatisfactory” relevance ratings had design shortcomings, including inadequate partner selection (2), inadequate instrument selection (1), and overly complex structure (1). One operation also was not aligned with the country strategy at approval. All 9 operations with “unsatisfactory” effectiveness ratings suffered from underachievement of the expected development outcomes. In addition, 4 operations lacked sufficient data to assess one of the key outcome indicators, including 1 operation that did not track indicators related to its main development outcome. Among 9 operations rated as “lack of data” for effectiveness, the most common reason was that the main expected development outcomes were not adequately addressed at approval (e.g., they lacked baselines or were not included in reporting requirements), nor were they tracked during supervision, and given that the relationship with the client had been severed, data could not be collected at the time of Mini-XSR preparation. More operations were rated as “lack of data” for effectiveness than for any other core criterion, indicating that data collection on development outcomes in these operations was particularly poor.

2.32 Of the 24 operations, 8 (33%) had “satisfactory” efficiency ratings and 17 (71%) had “satisfactory” sustainability ratings. Drivers for “unsatisfactory” efficiency ratings were a combination of reasons including poor financial performance of the client (6), NPL rates higher than expectations (4), and lower-than-expected target portfolio growth (3). Among the six operations with “unsatisfactory” sustainability ratings, four were affected by deteriorating performance of the client that made sustainability or results unlikely and two fell short of complying with safeguards. Sustainability of one operation was compromised by external shocks.

- 2.33 Half of the operations achieved an “unsatisfactory” investment outcome. All 12 operations with “unsatisfactory” investment outcome ratings received less than 60% of the expected interest income. In all but one case, where the Mini-XSR clearly stated that the prepayment penalty was waived, it was unclear whether IDB Invest contractually required such a penalty, whether it was paid, and whether it was included in the investment outcome calculation. Assessment of investment outcome was of particular interest for prepaid and repaid operations, and the Mini-XSRs should have elaborated on this point more clearly to draw lessons for future structuring.
- 2.34 Work quality was rated “unsatisfactory” in 20 out of 24 operations (83%). Inadequate project design (8), failure to include relevant indicators to track development outcomes (8), and unclear intervention logic (3) were shortcomings during appraisal and structuring. During supervision, key development indicators were not tracked in 15 cases. In 1 case, the supervision report was missing.
- 2.35 Lessons learned in the areas of project appraisal and supervision are in line with past evaluations and good practices. Lessons learned with regard to prepaid and repaid operations as presented in Management’s report are summarized in Box 2.1. Lessons on assessing client readiness (a), selecting committed and specialized partners (b), and assessing foreign exchange risks are in line with previous lessons from OVE evaluations.<sup>36</sup> Documenting the reasons for prepayment (d), flagging them early to ensure collection of development results data (f), structuring projects to ensure evaluability (g), and executing stronger supervision (i) are clearly good practice. Better disaggregation of data by client group (h) can also be valuable, when relevant for the target population and useful to assess development results.

### *Box 2.1. Lessons learned with regard to repaid/prepaid NSG operations*

Management’s report contains the following recommendations with the clarification that some of the lessons are already being addressed.

- a) During project origination and structuring, it is important to pay attention to client readiness, to ensure that each client has the technical and financial capacity to deliver on the project’s expected development impact.

<sup>36</sup> OVE’s evaluation on the IDB Group’s work through financial intermediaries refers to the need to align goals between the FIs and IDB Invest (document [RE-486-2](#)); OVE’s 2020 Validation Report raised the poor investment outcomes for prepaid projects (document [RE-552](#)), and OVE’s 2019 Validation Report mentioned that lack of knowledge about development achievements constrained evaluability (document [RE-544](#)).

- b) Partnering with specialized and committed clients often results in successful projects, particularly in the case of projects with FIs.
- c) Look out for prepayments after external shocks, and conduct supervision from a broad sector and country perspective.
- d) Foreign exchange risk should be carefully considered at entry or further discussed in approval documents, considering the client's operations.
- e) The documentation of the reasons for prepayment needs to be consistently available to better understand project performance, particularly IDB Invest's financial additionality.
- f) Operations near repayment or prepayment should be flagged automatically to ensure that development information is collected before clients leave the portfolio.
- g) During structuring, it is important to ensure that projects are designed to be evaluable.
- h) While outcome indicators included in the results matrix need to be consistent with the development objectives established for an operation, further data disaggregation should be supported, when possible, as it could provide valuable information to understand project performance and to track how benefits are distributed across different target populations.
- i) Strong supervision processes are needed to ensure that high-quality information is available to report on development impact and to generate lessons learned during execution.

2.36 The report should have delved more deeply into lessons from prepaid operations. Given that Management's report reviewed the performance of repaid and prepaid operations, lessons on prepayments should have included further reflections on structuring and the possibility of prepayment penalties. While IDB Invest has a protocol for prepayment penalty fees, the report pointed out that the reasons for prepayment were not always clearly documented. Similarly, given the report's findings that only 20% of the repaid operations achieved satisfactory ratings,<sup>37</sup> that prepaid operations had a tendency to take longer to reach the first disbursement,<sup>38</sup> and that prepayments follow changes in market competitiveness and the clients' financial conditions, an overarching lesson may be drawn that where a sufficiently long time elapses between approval, commitment, and disbursement, any material adverse development in the intervening period should cause the project to be reappraised, and the baselines for the relevant development results indicators should be updated.

---

37 This figure is 13% by OVE's calculation, as OVE considered one of the three prepaid operations rated "satisfactory" by Management to be lacking sufficient data to assign a rating.

38 On average, 13 months from approval to first disbursement, nine months longer than repaid counterparts, according to Management's report.

2.37 The conclusion that a simplified self-assessment approach can be applied to a wider set of operations than those covered in the one-off report is not evidence-based and, if adopted, would compromise the comprehensiveness, credibility, and intertemporal comparability of future self-assessments. Management's report argues that conducting one thematic rather than multiple individual evaluations for similar transactions would not compromise the accountability or quality of lessons learned, and that neither would streamlining XSR content or introducing different formats to evaluate operations with low execution, those with prepayments, and transactions with repeated clients that have the same use of proceeds or objectives. These lessons are not evident from the review of Mini-XSRs, whose analytical depth does not compare to that of regular XSRs. While OVE welcomes the general efforts to make XSRs concise, its position is that each operation that reaches EOM or leaves the portfolio early due to prepayment requires an XSR. OVE further cautions against introducing varying types of self-assessments that make comparison and aggregation of the results difficult. There needs to be a full accounting of the performance of the entire portfolio, and this can be done only when the same evaluation method is applied to each operation.



# 03

Performance  
Over Time



3.1 Comparability of performance data over time is essential for decision making. It helps institutions understand whether efforts undertaken to strengthen their operations' development effectiveness have been successful or whether course corrections are needed. At the IDB Group, comparable project performance data are available for NSG operations over the past five years, given that the same evaluation methodology has been used since 2018.<sup>39</sup> This allows for an analysis of trends in performance over time. In contrast, for SG operations, repeated changes to PCR Guidelines have prevented the creation of time series data that are comparable over time. As a result, performance data for SG operations is strictly comparable for only the last two validation cycles.<sup>40</sup>

## **A. Performance of public sector operations, 2021-2022**

3.2 Overall outcome ratings were similar in 2021 and 2022, as gains in sustainability were offset by losses in the remaining criteria. In 2021 and 2022, OVE validated 133 SG operations. The share of operations with a positive overall outcome rating was 52% in 2021 and 53% in 2022 (Figure 3.1). In 2022, sustainability increased by 4 percentage points, but all other criteria decreased by a few points (most notably relevance, by 7 points), leaving the share of projects with overall positive outcome ratings virtually unchanged. The proportion of “highly successful” and “successful” operations decreased, and a much higher share fell in the “partly successful” rating, the minimum to obtain a positive rating.<sup>41</sup>

---

39 While there is data available since the 2017 validation cycle, the ratings for the 2017 cycle were based on [2015 XSR Guidelines](#). Since the 2018 cycle, [2018 XSR Guidelines](#) have been in use without updates to the methodology, allowing for full comparability between 2018 and 2022.

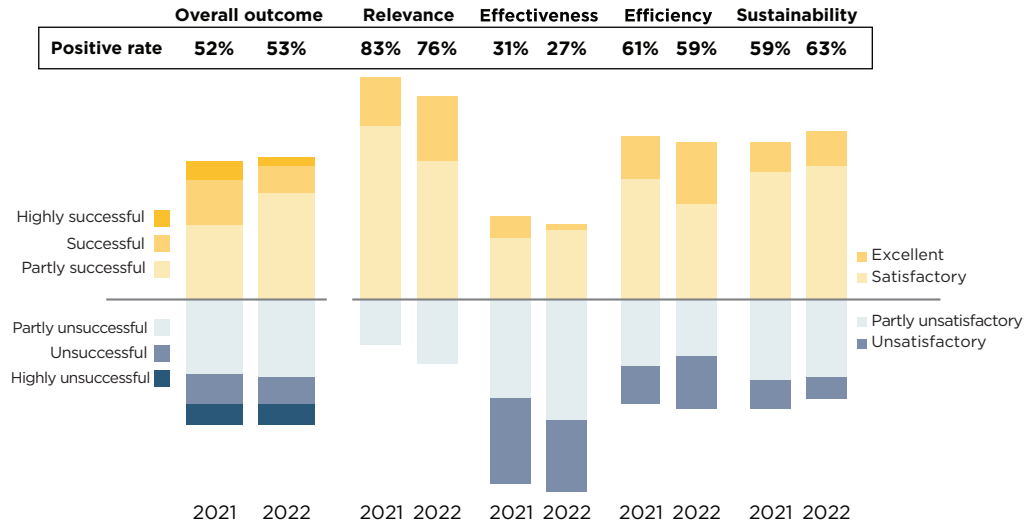
40 The 2017 validation cycle (document [RE-520](#)) was the first during which an objectives-based methodology was used to assess performance of all SG and NSG operations. For SG operations, this cycle used 2016 PCR Guidelines (document [OP-1242-5](#)). The PCR Guidelines were updated in 2018 (document [OP-1696-1](#)) and served as the basis for the following cycles of 2018 (document [RE-530-2](#)), 2019 (document [RE-544](#)), and 2020 (document [RE-552](#)). In 2020, the PCR Guidelines were updated again (document [OP-1696-6](#)), and this update was used for both the 2021 cycle (document [RE-565](#)) and the present 2022 cycle. Each update has introduced methodological changes that prevent full comparability across ratings.

41 The proportion of operations that fell under the three negative ratings remained almost constant between 2021 and 2022: from 28% to 29% under “partly unsuccessful,” from 11% to 10% under “unsuccessful,” and a constant 8% under “highly unsuccessful.” The biggest change was in the distribution of positive ratings— “partly successful” operations (the minimum for a positive rating) increased from 28% to 40%, while “successful” ones decreased from 17% to 10% and “highly successful” ones from 7% to 3%. These changes almost offset each other, leading to only a small increase in the overall positive rating, from 52% to 53%. However, a higher proportion of positively rated projects were much closer to the negative cutoff than last year.

**Figure 3.1**

**SG operations' ratings, six- and four-point scales, 2021-2022**

Source: OVE.



Note: The number of operations was 71 in 2021 and 62 in 2022 for all criteria except for efficiency, which is not rated for policy-based loans. For efficiency, the number of investment operations was 69 in 2021 and 56 in 2022.

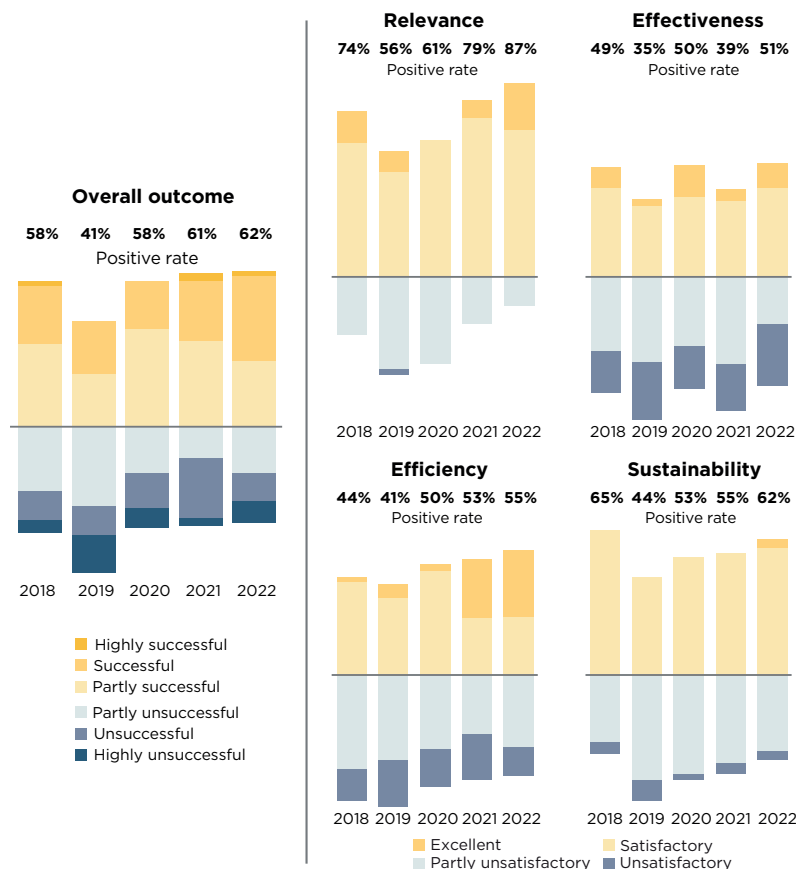
## B. Performance of private sector operations, 2018-2022

3.3 Overall outcome ratings have improved only modestly in recent years, due to an unsteady trend in effectiveness despite clear improvements in all other core criteria. These trends exclude projects that exited the portfolio from Special Assets. Between 2018 and 2022, OVE validated 198 NSG operations. There was a drop in scores on all core criteria between 2018 and 2019, which led to a plummet of 17 percentage points in the share of operations rated positive in overall outcome (Figure 3.2). However, since 2019, there has been a clear improving trend in all core criteria but effectiveness. Despite the improvements in the other criteria, this unsteady trend in effectiveness has led to only modest improvements in overall outcome ratings over the last three years. Over time, ratings have somewhat tended to concentrate toward the extremes—while positively rated operations have achieved higher levels of success, those negatively rated saw a slight increase in the number of “highly unsuccessful” operations.<sup>42</sup> These trends, however, should be interpreted with caution given that they do not include data for projects that exited the portfolio from Special Assets and therefore underperformed (such projects were not reported on in the past, as explained before).

42 A comparison of 2018 against 2022 shows that, among positively rated operations, the share of “successful” projects (the second-highest rating) grew from 23% to 34% and the share of “partly successful” decreased from 33% to 26%, while the share of “highly successful” projects remained constant at 2%. These trends show that there is an increase in the level of success among operations positively rated. However, a similar trend is also observed among the negatively rated operations.

**Figure 3.2**  
**NSG operations’ ratings, six- and four-point scales, 2018–2022**

Source: OVE.



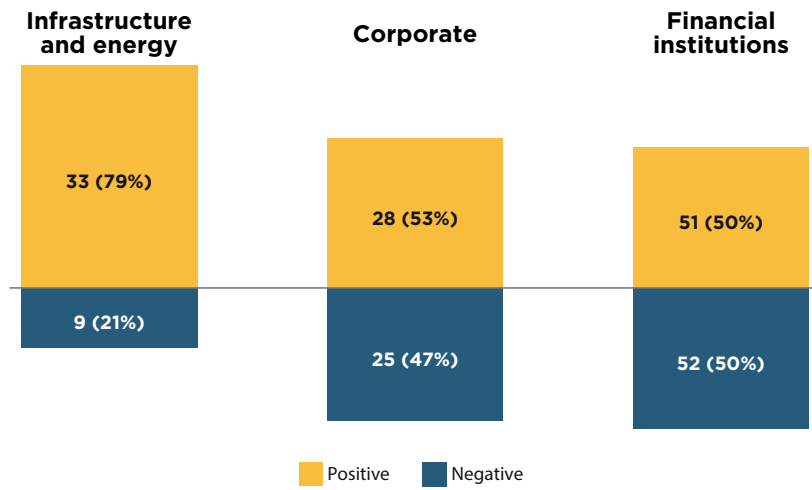
Note: Excludes operations that exited the portfolio from Special Assets.

3.4 By business segment, I&E operations outperformed the other segments. Considering all 198 NSG operations validated over the period 2018–2022 together, I&E achieved the highest (79%) percentage of positive ratings in overall outcome (Figure 3.3). In contrast, the corporate and FI segments had positive ratings of 53% and 50%, respectively. I&E is also the segment with the smallest number of operations, at 42, compared with 103 in FIs and 53 in the corporate segment.<sup>43</sup>

43 Per 2018–2022 validations data, I&E operations were 28 percentage points more likely to obtain a positive overall outcome rating than the rest of validated operations combined, while FI operations were 15 points less likely; these differences were statistically significant. The difference between corporate operations and the rest of operations was not statistically significant. When comparing between segments, I&E operations were 29 and 26 points more likely to obtain a positive rating than FI and corporate operations, respectively. Differences between the corporate and FI segments were not statistically significant. Results are based on linear probability models using robust standard errors; all point estimates mentioned were statistically significant at least at the 95% level.

**Figure 3.3**  
**NSG operations' overall outcome ratings by business segment, 2018–2022**

Source: OVE, with business segment data from IDB Invest (2022a, 2022d).







# 04

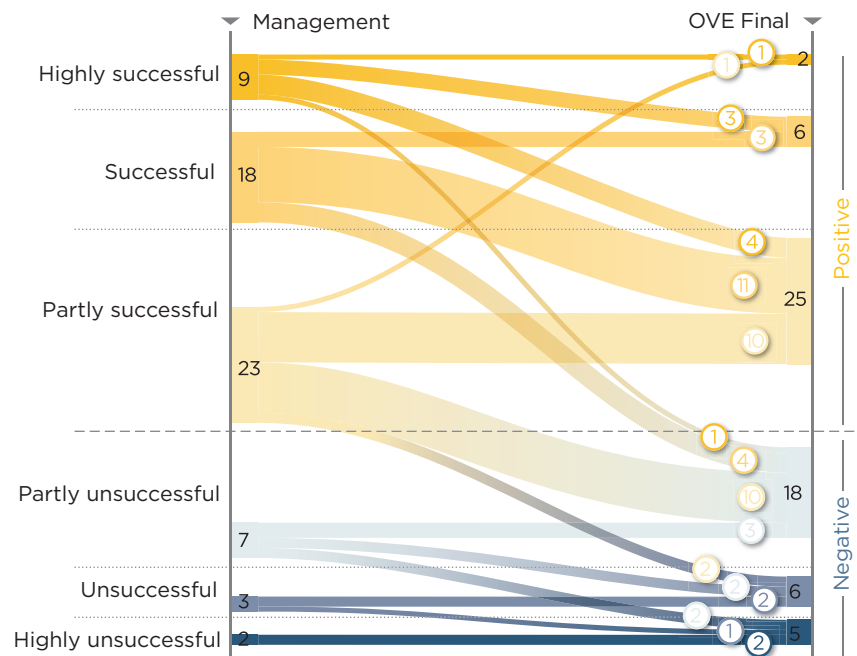
## Quality of Self- Assessments

## A. Quality of PCRs validated in 2022

4.1 During the 2022 validation exercise, OVE’s and Management’s overall outcome ratings differed for 41 out of the 62 validated operations. When the six-point rating scale is considered, the outcome rating for 40 SG operations was downgraded, while it was upgraded in 1 case (Figure 4.1). In 14 cases, the difference between Management’s and OVE’s ratings was by more than one point on the scale.<sup>44</sup> Among the 40 overall outcome ratings that were downgraded, efficiency was the most common downgrade, with 27 cases, closely followed by relevance and effectiveness, with 26 cases each. Sustainability was downgraded in 22 cases. Four of the operations were downgraded in all four core criteria, 9 in three core criteria, 15 in two core criteria, and 12 in one criterion. The 1 overall upgrade was due to a two-point upgrade in effectiveness (see Annex III).

**Figure 4.1**  
SG operations’ overall outcome ratings by Management and OVE, six-point scale, 2022

Source: OVE.



4.2 OVE downgraded overall outcome ratings from positive to negative for 17 operations. There was no upgrade in overall outcome ratings from negative to positive. All but 4 downgrades from positive to negative in the projects’ overall outcome rating were affected by downgrades in more than one core criterion (Table 4.1). Relevance, effectiveness, and efficiency each were downgraded in 10 cases and sustainability in 9. The most frequent reasons behind downgrades in relevance were

<sup>44</sup> Of these, 13 cases were downgraded by more than one point on the scale, and 1 case was upgraded by more than one point.

weaknesses in addressing institutional capacities, followed by output modifications and cancellations that affected the vertical logic of the projects, which had not been adequately taken into account in the PCR's previous rating. Reasons for downgrading effectiveness included the poor definition of indicators to measure specific objectives, lack of evidence to attribute achievement of objectives to the operation, and insufficient data to validate achievement. Downgrades in efficiency were mostly due to lack of credibility of the economic analysis, which prevented OVE from using it to determine the efficiency rating. Sustainability was downgraded mainly due to inadequate safeguards compliance and insufficient evidence to demonstrate appropriate implementation of safeguards and continuity of results.

**Table 4.1. Differences in Management's and OVE's ratings for SG operations with downgrades from positive to negative in overall outcome rating**

Differences in core criteria ratings	No. of operations (N = 15)
One core criterion	4
Two core criteria	6
Three core criteria	5
Four core criteria	2
<i>Total</i>	17

Source: OVE.

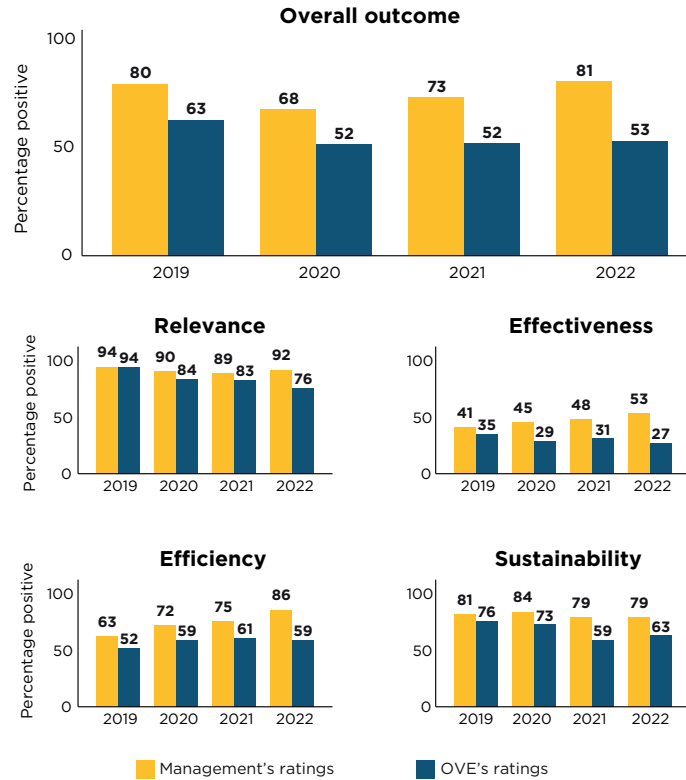
4.3 Differences between OVE's and Management's ratings for SG operations have continued to widen. While OVE rated the overall outcome positively in 53% of validated SG operations, Management did so in 81% of them. In the past four validation cycles, 28 percentage points is the largest discrepancy observed (Figure 4.2, see note under figure). Among the four core criteria, the discrepancies are largest in the efficiency (27 points) and effectiveness ratings (26 points). While the gap between OVE's and Management's sustainability ratings narrowed during this cycle compared with last year, the four-year trend shows that it is still wider than in the two prior validation years.



**Figure 4.2**

**SG operations' ratings by Management and OVE, positive rates, 2019–2022 (comparable rating scales only)<sup>a</sup>**

Source: OVE



Note: <sup>a</sup>While ratings are available starting with the 2017 cycle, the rating scale used by Management has not always been comparable to the one used by OVE. For older operations, Management prepared PCRs using 2014 guidelines, which rated on a 0–1 numeric scale, while OVE has used the current ordinal scales in all cycles since 2017. This difference makes ratings between OVE and Management not comparable for all 21 operations in the 2017 cycle and for 15 out of 35 in the 2018 cycle. Therefore, these two years are excluded from the graph. By the 2019 cycle, only 10 out of 64 operations used the numeric scale, so the graph shows the percentage positive for the remaining 54; by 2020, only 1 out of 63 did so, so the graph shows the positive shares for the remaining 62; and 2021 and 2022 show the complete set of validated projects (71 and 62, respectively), as all scales are comparable in these two cycles.

4.4 Fewer than half of the PCRs were of satisfactory quality. While OVE validated 62 operations, 3 were assessed jointly by Management in a single PCR; therefore, OVE reviewed the quality of a total of 60 PCRs based on the criteria described in Annex V. Of these, 29 (48%) were rated positive. These PCRs cited sufficient data content in a clear manner; captured all important project results in a balanced way; assigned ratings that were consistent with the evidence presented; and showed only minor shortcomings in clarity, consistency, candor, and adherence to PCR Guidelines. OVE rated the quality of the remaining 31 PCRs as “fair” (25) or “poor” (6). While Management has undertaken actions to improve the quality of PCRs, such as the development of an online nano-course, the majority of PCRs reviewed presented weaknesses as shown in Table 4.2.

**Table 4.2. Factors behind negative PCR quality ratings**

Section of the PCR and main weaknesses	No. of PCRs with this weakness (out of 31)
<b>Efficiency</b> Includes: questionable assumptions in economic analysis and inadequate cost determination, unreliable data (sources, errors), lack of sensitivity analysis, questionable comparators or control group, unclear justification for why a CBA or CEA was unfeasible	20
<b>Effectiveness</b> Includes: miscalculations or unclear calculation methods, insufficient explanations of indicators used/proposed, weak attribution analysis, inadequate data (including targets, baselines, and units of measurement used or proposed by the PCR), weak analysis of how changes during implementation affected effectiveness, misidentification of outputs as outcomes	19
<b>Relevance</b> Includes: poor analysis of how changes during implementation affected relevance, weak analysis of alignment with country realities, weak analysis of the project's vertical logic	14
<b>Sustainability</b> Includes: key information missing on safeguards and continuation of results.	12
<b>Lessons learned</b> Includes: lessons lacking detail or depth, lack of lessons for future operations, lessons not relevant.	10
<b>Bank and borrower performance</b> Includes: insufficient reporting on key performance issues	2

Source: OVE.

Note: Numbers do not add up to 31 because weaknesses are not mutually exclusive.

4.5 Over half of the PCRs were missing key information necessary to assess performance, which required Management to submit additional information and OVE to revisit evidence and reassess relevant ratings. Of 60 PCRs, 32 lacked key information, which Management provided only in response to OVE's draft validation note. In 18 cases, this additional information led OVE to change the rating of at least one core criterion for the final version of the validation notes (Box 4.1). This points to an important lack of completeness in PCR reporting. Lack of completeness leads to inefficiencies in the review process, as Management must respond to draft validations that were based on incomplete information, and OVE must assess additional information and revise the validations, both of which take considerable time. Some PCRs also lacked candor, some by failing to discuss large cancellations faced by the operation or changes in project scope.

**Box 4.1. Examples of additional information provided by Management that resulted in changes to core criteria ratings**

OVE validates SG operations based on the information documented in PCRs. Management has an opportunity to review the draft validation and provide comments, if any, before the validation note is finalized. There were multiple cases where Management provided substantial additional information that should have been included in the PCRs, which resulted in a change in core

criteria ratings. Out of the total 60 validated PCRs, Management provided additional information on 32, and this information led to a change in ratings for 18 operations. In the majority of these 18 cases, the additional information Management provided was available at the time of PCR preparation and should have been included in the PCR. In 15 cases out of these 18, OVE rated the PCR quality as negative due, in part, to this important lack of completeness in reporting (the remaining 3 scored well in the other criteria considered—see Annex V—resulting in a positive rating). Examples in this box come from this subset of 15 operations.

Operation “PROCIDADES Cascavel” (BR-L1344) was initially assigned a negative relevance rating. While the operation had a strong vertical logic at approval, it faced cancellations of several outputs during implementation, which cast doubt on whether the vertical logic had remained strong during implementation. The PCR did not thoroughly discuss how these cancellations affected the vertical logic. In response to OVE’s draft note, Management stated that “the PCR does not argue why the vertical logic is maintained [...] we would like to add more elements that should have been included in the document.” Management provided additional information showing that, despite the cancellations, other outputs that were implemented played a key role in maintaining the alignment of the project’s design with its specific objectives. As a result of this additional information, OVE upgraded the rating.

Operation “Technology Transfer to Small Farmers” (HA-L1059 and HA-G1025) was initially assigned a negative efficiency rating. The PCR rated efficiency as “excellent” based on the results of a cost-benefit analysis that Management did not attach to the PCR and OVE could not verify. In response to OVE’s draft note, Management provided the analysis and OVE assigned a positive rating upon review.

Operation “Program for Strengthening the Digital Agenda” (AR--L1304) was initially assigned a negative sustainability rating. This was a truncated PBP, and OVE questioned that results had continued in light of the truncation and the little evidence provided by the PCR on this issue. In response to OVE’s draft note, Management recognized that it had not provided sufficient information on recent progress that had been made by the government in key areas that were supported by the first (and only) operation despite the truncation. Management collected new data and provided it to OVE, resulting in an upgrade to a positive rating.

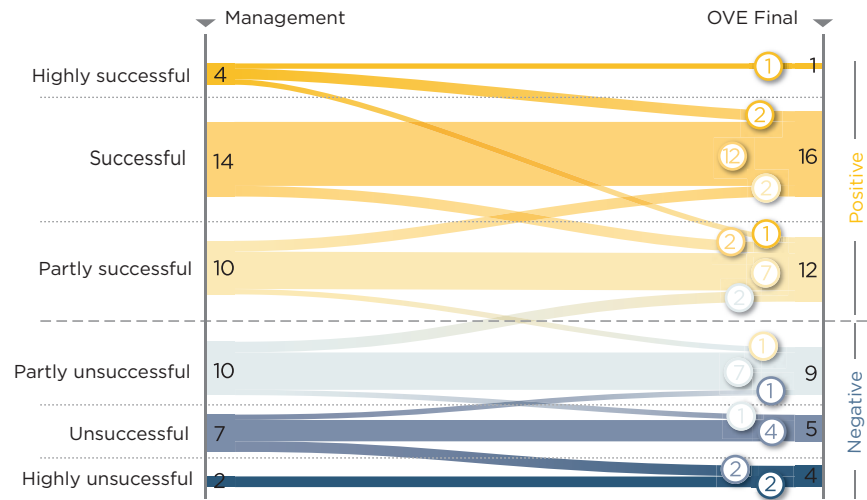
## B. Quality of XSRs validated in 2022

4.6 During the 2022 validation cycle, OVE’s and Management’s overall outcome ratings differed for 14 of the 47 validated operations. When considering the six-point scale, outcome ratings for 9 operations were downgraded and those of 5 operations were upgraded (Figure 4.3). For only 1 operation did OVE’s and Management’s outcome ratings differ by more than one point on the scale. Discrepancies among the four core criteria were relatively small,

varying from none for efficiency to 1 upgrade and 4 downgrades for relevance. In 5 operations, Management’s and OVE’s ratings differed on more than one of the four core criteria.

**Figure 4.3**  
**NSG operations’ overall outcome ratings by Management and OVE, six-point scale, 2022**

Source: OVE

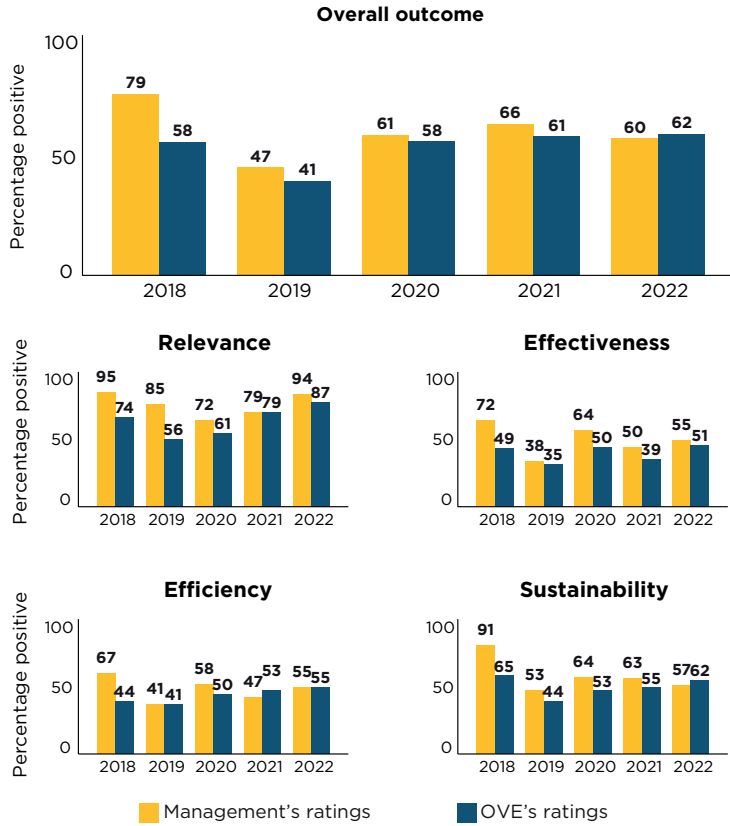


4.7 OVE downgraded the overall outcome rating of one operation from positive to negative and upgraded two operations from negative to positive. One operation downgraded from positive to negative was affected by downgrades from “partly unsatisfactory” to “unsatisfactory” on two criteria. Two operations upgraded from negative to positive were both affected by an upgrade on one criterion, in both cases from “unsatisfactory” to “partly unsatisfactory.”

4.8 OVE’s and Management’s overall outcome and core criteria ratings have been relatively close in the past few years, and closest in the 2022 validation cycle. While OVE rated 62% of the NSG operations’ overall outcomes positively, Management rated 60% positively. Two percentage points is the smallest discrepancy observed in the past five validation cycles (Figure 4.4). Discrepancies among the four core criteria have also been small, ranging from none in efficiency ratings to seven points in relevance ratings. The five-year trend points to a convergence of OVE’s and Management’s ratings.

**Figure 4.4**  
**NSG operations’ ratings by Management and OVE, positive rates, 2018–2022**

Source: OVE



4.9 The overall quality of XSRs was high, with over 80% achieving a positive quality rating. OVE rates the quality of XSRs based on the criteria described in Annex V. Of the XSRs, 4 (9%) were rated “excellent,” 34 (72%) “good,” and 9 (19%) “fair.” No XSRs were rated “poor.” The majority of the 9 XSRs rated “fair” had combinations of issues, including lack of or inaccuracies in key analyses and insufficient or inaccurate information. Other challenges included missed opportunities to draw relevant lessons, internal inconsistencies, and inconsistencies between the analysis presented and the ratings provided. In six cases, Management provided additional information during the feedback process that resulted in changes in core criteria ratings. Efficiency ratings changed in 4 operations, and sustainability ratings did so in 3. In an additional 2 cases, Management provided additional information on efficiency of the operation that did not result in a change in OVE’s ratings.



# 05

## Conclusions and Recommendations

- 5.1 OVE validated PCRs for 62 SG operations and XSRs for 47 NSG operations in 2022. In addition, OVE reviewed the one-off report that Management prepared to catch up on 24 additional NSG operations through Mini-XSRs. Those operations were repaid or prepaid between 2016 and 2019, and IDB Invest no longer has engagement with the clients. Considering the difficulty in collecting information necessary for regular XSRs when significant time has passed between the repayment and the time that the operation was flagged for XSR preparation, Management and OVE agreed that their performance would be accounted for through Mini-XSRs on an exceptional basis. Such a time lag is expected to be eliminated going forward, with an improved EOM database that should systematically identify operations subject to an XSR each year. Thus, all operations repaid or prepaid in 2020 or later will be subject to regular XSRs.
- 5.2 Of SG operations, 53% achieved positive overall project outcome ratings. This figure falls short of the target of 70% set in the 2020–2023 Corporate Results Framework. Effectiveness continued to be the lowest-rated criterion (with 27% of operations rated positive), driving down SG operations’ overall outcome ratings.
- 5.3 Most operations with negative effectiveness ratings were affected by a combination of factors, most commonly outcome shortfalls and poor M&E. Multiple reasons contributed to underachievement, including institutional changes, external factors, and design shortcomings. Importantly, however, many PCRs did not consistently discuss why the operations faced underachievement. This is a significant shortcoming that compromises the learning value of PCRs.
- 5.4 Nearly all SG operations with negative efficiency ratings were affected by deficient economic analysis. In most of these cases the analysis lacked credibility, and in a few cases economic analysis was not presented at all when such an analysis would have been feasible.
- 5.5 Out of 47 NSG operations validated in 2022, 29 achieved positive overall outcome ratings. Aside from 18 validated operations with negative overall outcome ratings, 4 operations with EOM in 2015 or later and approved amounts of over US\$1 million exited the portfolio in 2020 from Special Assets and were also assigned a negative overall outcome rating by OVE. Therefore, the share of operations with positive overall outcome ratings is 57%, short of the target of 65% set in the 2020–2023 Corporate Results Framework. Among core criteria, relevance continues to be the highest-rated and effectiveness continues to be the lowest-rated criterion. The majority of operations with negative effectiveness ratings failed to achieve targeted results, and several suffered from

M&E shortcomings. XSRs cited macroeconomic and regulatory issues most frequently as the reasons for underachievement, though the depth of analysis varied across XSRs.

- 5.6 Work quality was satisfactory for a little less than half of NSG operations. A combination of shortcomings drove the negative ratings. Shortcomings during the screening, appraisal, and structuring stage—such as unclear intervention logics, failure to identify and address risks, and failure to set up necessary arrangements for the fulfillment of fiduciary requirements—were cited more frequently. However, issues during supervision, such as failure to collect key data, noncompliance with E&S safeguards, and inability to adapt to changing environments, also affected 60% of cases where work quality was rated negative.
- 5.7 OVE found that 6 of the 24 operations included in the one-off report achieved a “satisfactory” overall outcome, 12 were rated negatively, and 6 lacked sufficient data to assess outcomes. Of the 24 operations, 9 (38%) lacked sufficient data to assess effectiveness, indicating that data collection on development outcomes in these operations was particularly poor. Lessons learned included in the one-off report that related to project appraisal and supervision were in line with past evaluations and best practices, though the report should have delved more deeply into lessons on prepayment.
- 5.8 The report presents five years of comparable data for NSG operations’ performance, while only two years of comparable data are available for SG operations due to frequent changes in PCR Guidelines. Overall outcome ratings of SG operations were similar in 2021 and 2022, as gains in sustainability were offset by losses in the remaining criteria. Overall outcome ratings of NSG operations have improved modestly in recent years, due to an unsteady trend in effectiveness despite clear improvements in all other core criteria.
- 5.9 Differences between OVE’s and Management’s ratings for SG operations continued to widen, pointing to quality shortcomings in PCRs, while the difference was negligible for NSG operations. The quality of over half of PCRs was rated negative, including 6 that were rated “poor.” The analytical inputs in efficiency and effectiveness assessments were particularly deficient. More than half of all PCRs validated this year were originally missing key information that was provided only later by Management. The overall quality of XSRs, on the other hand, was strong, with over 80% of XSRs rated positive, and no XSR rated “poor.” The majority of the 9 XSRs rated “fair” had combinations of issues, including lack of or inaccuracies in key analyses, or insufficient or inaccurate information.



Considering these findings, OVE recommends the following.

**A. For IDB Management:**

1. Strengthen PCR quality. Over half the PCRs reviewed by OVE this year had considerable shortcomings, and many of them lacked the information necessary to make an adequate assessment of project performance. Incompleteness in key information results in significant inefficiencies in the review process. It requires Management to respond to draft validations that were based on incomplete information. OVE then must assess additional information submitted by Management subsequently and revise the validations, both of which take considerable time. In addition, while it is essential for IDB to systematically analyze why its operations may not achieve the targeted objectives, OVE found that in about half of PCRs for operations with negative effectiveness ratings, no or insufficient explanation was provided for underperformance on at least one outcome indicator. Quality of economic analysis also continues to be a key factor that drives down the quality of PCRs, even after Management developed sector-specific guidance notes on the preparation of CBAs and CEAs, following OVE's 2018 recommendation. The following actions are necessary for strengthening PCR quality.
  - i. Strengthen Management's internal quality assurance to ensure the PCRs submitted to OVE include all key information needed to substantiate all ratings and are prepared in accordance with the PCR Guidelines.
  - ii. Strengthen institutional learning by assessing and documenting in PCRs the reasons why operations underachieve.
  - iii. Adequately implement the CBA and CEA guidance and include the underlying analysis in an Excel file for validation.
  - iv. Conduct PCR training jointly with OVE.

**B. For IDB Invest Management:**

1. Complete another one-off report to fully catch up on operations repaid or prepaid between 2015 and 2019 without XSRs, focusing on lessons learned, and ensure that full XSRs are prepared for operations repaid or prepaid in 2020 and onward. Management made significant progress in strengthening the system to identify projects subject to XSRs each year, including updating the historical list of operations. OVE supports Management's proposal to prepare another one-off report in 2023 to complete catch-up reporting for

the operations that reached EOM between 2015 and 2019. Given that the report will exclusively review the performance of repaid and prepaid operations, lessons on prepayments should include further reflections on structuring and the possibility of prepayment penalties.

2. Establish a system to ensure that all information necessary to prepare an XSR is collected at the time that a client indicates intention to prepay or repay a loan. Given the difficulty in accounting for development results when substantial time has elapsed after repayment or prepayment and IDB Invest's disengagement with the client, relevant data should be collected systematically when the client indicates intention to repay or prepay.
3. Set a time point after which reappraisal is mandatory when a long time elapses between approval, commitment, and disbursement; focus reappraisal on any material adverse development in the intervening period, as well as an update to the baselines for the relevant development results indicators. Considering the findings of the one-off report that prepaid operations had a tendency to take longer to first disbursement, and that prepayments followed changes in market competitiveness and the clients' financial conditions, it is important to establish a process to ensure continued alignment of project design with market conditions and client focus, thus improving the likelihood of operational success and smoother data collection.



# References

Inter-American Development Bank (IDB). 2022. Enterprise Data Warehouse. *Tables ODS.OPER\_ODS\_OPER, EDW.SL\_FINBAL\_PROD\_RPT [Data set]*.

IDB Invest (2016). IDB-IIC *Combined Historical Approvals up to 2015 [Data set]*. Data obtained by information request from OVE to IDB Invest.

---- (2022a). *Azure SQL Database: iic.Maestro Transaction [Data set]*.

---- (2022b). *Azure SQL Database: metrics.Operational Approvals RPT [Data set]*.

---- (2022c). *Azure SQL Database: metrics.Operational Use of Funds RPT [Data set]*.

---- (2022d). *DVF's XSR Hub: Full Report OVE 07/18/2022 [Data set]*.



## Office of Evaluation and Oversight - OVE

Established in 1999 as an independent evaluation office, OVE evaluates the performance and development effectiveness of the activities of the Inter-American Development Bank Group (IDB Group). These evaluations seek to strengthen the IDB Group through learning, accountability and transparency.

OVE evaluations are disclosed to the public in accordance with IDB Group policies to share lessons learned with the region and the development community at large.



[iadb.org/evaluation](http://iadb.org/evaluation)



[linkedin.com/showcase/idb-ove](https://www.linkedin.com/showcase/idb-ove)



[@BID\\_evaluacion](https://twitter.com/BID_evaluacion)