ove Office of Evaluation and Oversight

Project Evaluation

OVE´s Review of PCRs and XSRs: The 2021 Validation Cycle

IDB | IDB Invest

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

RE-565
December 2021

Project Evaluation

OVE's Review of Project Completion Reports and Expanded Supervision Reports: The 2021 Validation Cycle

Office of Evaluation and Oversight

Inter-American Development Bank

IDB Invest

Contents

Acknowledgements

This report was prepared by Chiaki Yamamoto (team leader), Stephany Maqueda, Oliver Peña-Habib, and Melanie Putic, under the direction of Ivory Yong, OVE Director, and Monika Huppi, Principal Advisor.

Acronyms and Abbreviations

CO	Operational Closure
E&S	Environmental and Social
EOM	Early Operating Maturity
IDB	Inter-American Development Bank
M&E	Monitoring and Evaluation
NSG	Non-Sovereign-Guaranteed
OVE	Office of Evaluation and Oversight
PBL	Policy-Based Loan
PBP	Programmatic Policy-based Loan
PCR	Project Completion Report
SG	Sovereign-Guaranteed
XSR	Expanded Supervision Report

Executive Summary

The report summarizes the results of the 2021 validation of Project Completion Reports (PCRs) and Expanded Supervision Reports (XSRs) by the Office of Evaluation and Oversight (OVE). 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 project evaluation 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 on 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 non-core criteria, which are rated but do not count toward a project's overall outcome rating. For simplicity, this report refers to the top half of the rating scale as a positive rating and to the bottom half as a negative rating.

OVE supports accountability and learning through the validation of completed PCRs and XSRs. It contributes to the credibility of the IDB Group's project performance reporting system through the independent validation of PCRs and XSRs. Management assigns performance ratings to the projects in the PCRs and XSRs, based on the relevant guidelines. OVE's validation assesses whether these self-assessments are substantiated by evidence and prepared in accordance with guidelines.

During the 2021 validation process, OVE validated 71 PCRs of sovereign-guaranteed (SG) and 38 XSRs for non-sovereign-guaranteed (NSG) operations. These operations represented US$8.8 billion in SG approvals and US$1.1 billion in NSG approvals. All but one SG operations validated reached operational closure in 2019, and the 71 validated PCRs represent 94% of operations that were expected to be validated in this cycle. Most of the 38 NSG operations validated reached early operating maturity (EOM) or exited the portfolio in 2019. In improving the system for timely identification of operations subject to XSRs, Management identified eleven operations that should have been subjected to XSRs in previous validation cycles and also included them in this cycle. The improvements in the system also led to the belated identification of 12 operations that should have been subject to XSRs this year and will need to be covered next year. Therefore, the XSRs validated this year only present a partial performance picture of the operations that reached EOM or exited the portfolio in 2019 and were thus expected to be validated this year.

About half of the validated SG operations achieved positive overall outcome ratings. Relevance was the highest-rated core criterion (83% rated positive). Effectiveness was the lowest-rated core criterion (31% rated positive); it drove the overall development outcome rating down. Failure to reach targeted results, the poor quality of monitoring and evaluation (M&E), cancellations, project design weaknesses and extension of disbursement periods were correlates of negative effectiveness rating. Among the 71 operations assessed during this cycle, effectiveness ratings of 31 suffered from M&E issues that prevented a clear determination of whether the project objectives were achieved. 61% of SG investment operations achieved positive efficiency ratings, and 59% of all SG operations achieved positive sustainability ratings. A little over half of SG operations with negative sustainability ratings (15 projects) were closed while environmental and social (E&S) issues, including resettlement, were still pending. This practice affects the sustainability of operations and exposes IDB to reputational risks. Yet, PCRs do not consistently document project safeguards performance. Bank performance was rated positive for a little less than 60% operations, and Borrower performance was rated positive for slightly more than 60% of operations. The quality of PCRs was variable with 66% being rated excellent or good and 34% rated fair or poor. Differences between Management's self-ratings and OVE's ratings of project performance persist. OVE downgraded the overall outcome rating of 15 out of 71 operations from positive to negative.

Sixty-one percent of NSG operations achieved positive overall outcome ratings. Relevance continued to be the highest-rated core criterion (79% positive) and effectiveness the lowest-rated (39% positive). As in previous cycles, failure to reach targeted results, poor M&E, and flawed project designs were correlates of negative effectiveness ratings. Data was particularly sparse for objectives that went beyond the narrow performance of the enterprise that borrowed the funds, such as for job creation. About half (53%) of NSG operations achieved positive efficiency ratings, and 55% achieved positive sustainability ratings. Among the non-core criteria, work quality remained the lowest rated, considered satisfactory in only 29% of operations. The most frequently cited reasons for negative work quality ratings were shortcomings during appraisal and screening, such as weaknesses in financial analysis; partner identification; and mitigation of foreign exchange, policy, or construction risks. Differences between Management's overall project outcome ratings and OVE's ratings were relatively small. OVE downgraded the overall outcome rating of 2 out of 38 operations from positive to negative.

Given that effectiveness has consistently been the lowest rated core criteria for both SG and NSG operations over the last few years, this report further looks into what lies behind these ratings. Among SG operations, objectives aimed at better service provision were most likely to achieve positive ratings, and objectives aimed at institutional capacity building and sector-level changes were least likely to do so. Among NSG operations, objectives aimed at increasing renewable energy generation were most likely to achieve positive outcome ratings, and objectives aimed at creating jobs were least likely to do so.

Given the role partial cancellations have played in low effectiveness rating of SG operations, the report reviewed the reasons behind cancellation documented in the cancellation notes IDB prepared for SG operations that were closed before significant execution took place. IDB produced 15 cancellation notes between validation cycles 2018 and 2020, when preparation of cancellation notes was mandatory. These operations represented a combined approved amount of US$971 million and disbursements of US$172 million. OVE also identified another 16 SG operations that were fully cancelled during validation years 2018–20. These operations were approved but never reached the disbursement stage (cancellation notes were not required for these cases). These cancelled operations cost the Bank US$5.3 million in preparation and supervision expenses and used nearly 3,000 weeks of staff time.

All 15 operations with cancellation notes were cancelled because of multiple factors. Most prominent were issues that hindered the capacity of executing units. The second-most frequently cited reason for cancellation were changes in government policies and processes. Five operations experienced excessive delays, three of them in reaching eligibility.

Considering these conclusions, OVE recommends:

A. For IDB Management

1. **Systematically document in PCRs how E&S safeguard issues were addressed and what their final status was at project closure.** PCRs did not consistently report on safeguard issues when safeguards had been triggered. Unresolved E&S issues at project closure were a significant driver of less than satisfactory sustainability ratings. PCRs for all projects that trigger safeguards

need to consistently report on how projects complied with Bank safeguard requirements, indicating whether and how E&S issues were resolved by project closure.

2. **Strengthen reporting on cancellations.** Given the validations' findings that partial cancellations are a significant driver of less than satisfactory project effectiveness, OVE recommends that PCRs consistently discuss whether and how partial cancellations have affected the results chain and achievement of the operation's objectives. OVE also recommends that IDB report to the Board regularly the amount and reasons for cancelled operations for which no PCRs are required.

3. **Strengthen the measurement of project performance.** Weaknesses in project M&E continue to be an important reason for less than satisfactory project effectiveness ratings. Management's initiative to strengthen the Development Effectiveness Matrix by requiring the inclusion of at least one valid outcome indicator to assess achievement of each specific objective is in line with the finding that the poor quality of M&E is one of the key contributors to negative effectiveness. Efforts need to be made to strengthen project performance measurement by ensuring that baseline and follow-up data for indicators are routinely collected.

B. For IDB Invest Management

1. **Continue strengthening the system to identify projects subject to XSR each year.** Management's efforts to strengthen the system to identify projects subject to XSRs has pointed to a significant number of operations for which no XSRs were prepared in the past. To identify projects reaching EOM each year in a timely manner and transparently account for their performance, OVE recommends that IDB Invest:

    i. Build on the progress it has made to date and agree with OVE on the final structure of the database that contains all data necessary to determine when each operation (including those in Special Assets) reaches (or has reached) EOM or has exited the portfolio, as well as each operation's XSR status and reason for XSR waiver where applicable;

    ii. Ensure that the agreed upon database is fully updated and cleaned of any inconsistencies by February 1, 2022, and covers all NSG operations approved since 2015, including operations that are or previously were in Special Assets, and provide OVE permanent access to the database for purposes of verification of projects subject to XSRs each year;

iii. Ensure that the database is updated and cleaned of any inconsistencies annually by February 1, and clearly shows which operations are subject to XSRs in the XSR cycle starting in that year, which operations are not and why, so that OVE can verify the list with a February 1, cut-off date each year.

2. **Strengthen the evidence base for measuring project effectiveness.** Validations have repeatedly found that the poor quality of M&E is one of the key contributors to negative effectiveness ratings. OVE therefore recommends that IDB Invest identify all outcome indicators that have frequently had insufficient means of verification and thus poor M&E, such as job creation, sector level changes, and other objectives that go beyond direct benefits to the client and ensure that sufficient means of verification for these indicators are in place from the project design stage.

01

Introduction

1.1 This report summarizes the results of 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. According to both PCR and XSR Guidelines,[1] 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."[2] In addition to the core criteria, the PCR and XSR Guidelines require the inclusion of non-core criteria, which are rated but do not count toward a project's overall outcome rating. Among these non-core criteria, PCRs include the performance of the Bank and its counterparts (Borrowers); XSRs include the financial and nonfinancial additionality of IDB Invest, the outcome of the investment, and IDB Invest's work quality. Annex I provides a full description of IDB's evaluation framework.

1.3 The IDB Group's Office of Evaluation and Oversight (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 PCRs and XSRs. Management

1 See PCR Principles and Guidelines (OP-1696-5), 2018 XSR Guidelines for IDB Group Private Sector Operations.

2 The overall outcome rating is calculated as a weighted average. For investment loans, relevance, efficiency, and sustainability criteria weigh 20% each and effectiveness 40%. For policy-based loans (PBLs), relevance and sustainability weigh 20% each and effectiveness 60%.

assigns performance ratings to the projects in the PCRs and 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 outcome rating. These final ratings are reported by IDB Group Management in its Development Effectiveness Overview. This is the fifth consecutive year in which OVE has validated these self-evaluations in the context of the IDB Group's objectives-based evaluation framework for public and private sector operations.

B. Scope of this report

1.4 This report presents the results of the 2021 validation process, undertaken by OVE between October 2020 and April 2021. OVE reviewed and validated 71 PCRs of sovereign-guaranteed (SG) operations,[3] all but one of which reached operational closure (CO) in 2019.[4] OVE also validated 38 XSRs for non-sovereign-guaranteed (NSG) operations, most of which reached early operating maturity (EOM) or exited the portfolio in 2019.[5] These operations represented US$8.8 billion in SG approvals and US$1.1 billion in NSG approvals. Annexes II and V provide the lists of SG and NSG operations reviewed, respectively.

1.5 In addition, the report provides the results of its deep dive into the correlates of effectiveness ratings. Effectiveness makes up at least 40% of the overall project outcome rating (60% in the case of a policy-based loans, PBLs). It received the lowest ratings of the four core criteria in the previous four validation cycles, for both SG and NSG operations. Understanding what factors are associated with low effectiveness ratings will therefore help shed light on what affect overall project outcome ratings.

1.6 Finally, the report reviews the knowledge generated by the cancellation notes produced between validation years 2018 and 2020. Last year's validation report found that cancellations were an important driver of low effectiveness ratings. Based on this finding, this year's report takes a closer look at cancellations and reviews cancellation notes that were prepared for operations that did not reach significant execution and therefore the PCR stage.

3 OVE validated 57 free-standing investment loans, seven investment loans supported by related investment grants, four free-standing investment grants, one program consisting of three investment loans, one free-standing PBL, and one programmatic PBL series covering two loans.

4 CO-L1019 closed in 2018.

5 Twenty-seven operations reached EOM or exited the portfolio before reaching EOM in 2019, and 11 had reached EOM or exited the portfolio before reaching EOM between 2016 and 2018.

1.7 The report is organized as follows. Section II presents the results of OVE's validations for SG and NSG operations in 2021. Section III examines possible correlates of negative effectiveness ratings. Section IV analyzes the content of cancellation notes. Section V presents conclusions and recommendations.

02

Results of Operations Validated in 2021

2.1 This section presents the results of the 2021 validation exercise. It includes the highlights of this year's validations, a cross-sectoral comparison of overall project development outcome and core criteria ratings, and a summary of non-core criteria ratings. It ends with a comparison between the ratings assigned by Management and OVE and OVE's assessment of the quality of PCRs and XSRs.

2.2 For simplicity, the report groups ratings into positive and negative categories. A positive rating refers to the top three ratings on the six-point scale used for the overall project development outcome rating (i.e., "partly successful" or higher) and to the top two ratings on the four-point scale used for core and non-core criteria and for rating the quality of a PCR or XSR (i.e., "satisfactory" or "excellent" for core and non-core criteria, "good" or "excellent" for the quality of a PCR or XSR). A negative rating refers to the ratings of "partly unsuccessful" or below for the overall outcome rating of the operations, "partly unsatisfactory" and "unsatisfactory" for the core and non-core 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
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
Non-core 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 and XSR Guidelines (2020 revision).

A. Public sector projects reviewed in 2021

2.3 OVE reviewed PCRs for 71 operations, of which 70 closed in 2019 and 1 closed in 2018. These PCRs covered 57 free-standing investment loans, seven investment loans supported by related investment grants, four free-standing investment grants, one program consisting of three investment loans, one free-standing PBL, and one Programmatic Policy-based Loan

(PBP) series covering two loans. They represented US$8.8 billion in total approved amount. Table 2.2 breaks these operations down by sectors.

Table 2.2. PCRs for SG operations validated in 2021, by department

Department	SG operations validated	Approved amount
Infrastructure and Environment Sector (INE)	27	(US$ million)
Institutions for Development Sector (IFD)	15	2,859
Social Sector (SCL)	10	1,256
Climate Change and Sustainable Development Sector (CSD)	17	912
Integration and Trade Sector (INT)	2	40
Total sovereign-guaranteed operations	**71**	**8,827**

Source: OVE, with data from IDB (2021).
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.

2.4 PCRs for 94% of operations subject to validation in 2021 were submitted to OVE on time and are included in this report. The current PCR guidelines require that a PCR be prepared for free-standing closed operations over US$3 million that had a Development Effectiveness Matrix (DEM) at approval and for a programmatic series of loans (or investment grants) with shared objectives and a common results framework after the last operation of the series closes. Seventy-five operations that closed in 2019 fell under this category. Among them, four were not submitted on time and will be validated in the 2022 validation cycle. Management prepared 10 additional PCRs that were not subject to OVE validation because these operations were approved without a DEM (i.e. before April 2011) and the PCRs were prepared according to the old guidelines.

1 Overall project development outcome ratings

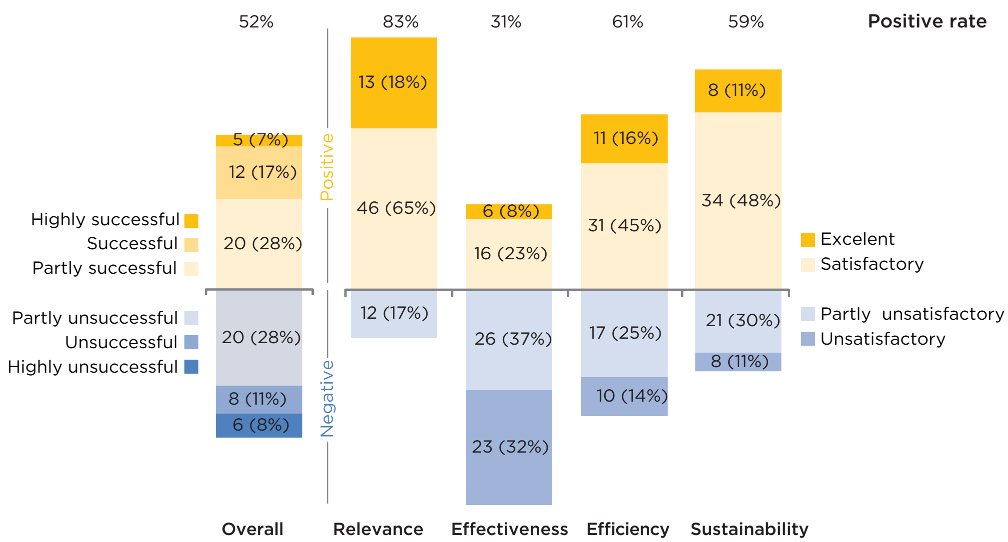
2.5 About half of the validated SG operations achieved positive overall outcome ratings.[6] This share falls short of the target of 70% set in the 2020–23 Corporate Results Framework (CRF). The ratings of 84% of operations fell in the middle four categories of successful, partly successful, partly unsuccessful and unsuccessful, only 7% were rated highly successful, and 8% were rated highly unsuccessful (Figure 2.1).

6 A PBP series consisting of multiple loans that share the same objectives and results frameworks and are evaluated as a single program. For purposes of this report, OVE treats these programs as single operations. Programs including a combination of investment loans and grants or a set of several investment loans that were part of a single program with shared objectives are also evaluated as a whole and counted as a single operation.

Figure 2.1

Overall project outcome and core criteria ratings for sovereign-guaranteed operations validated in 2021

Source: OVE.



2. Core criteria: Relevance, effectiveness, efficiency, and sustainability

2.6 Relevance was the highest-rated core criterion, with 59 (83%) of SG operations achieving positive ratings (Figure 2.1). Twelve operations (17%) received negative ratings for relevance. All of them suffered from a project design that was not aligned with country realities either at design stage, during implementation, or at closing. Misalignment happens, for example, when assessment of the country's institutional structure or local capacity is insufficient, or an operation's vertical logic does not hold when a component is cancelled or country realities change. Project objectives of two operations were also not aligned with the country strategies.[7]

2.7 Effectiveness was the lowest-rated core criterion, with just 22 (31%) of SG projects achieving positive ratings. Of the 49 operations with negative effectiveness ratings, 34 (69%) were affected by poor performance, meaning they failed to achieve targeted results. Thirty-one (63%) operations were affected by the poor quality of M&E, which made it impossible to evaluate whether some of the objectives had been achieved or results were attributable to the operation. Twelve (24%) operations were affected by partial cancellation. Some of the operations with low effectiveness ratings were affected by more than one of these factors. Given the consistently low effectiveness ratings, Chapter III takes a deep dive into effectiveness.

7 The number of operations adds up to more than 12 because categories are not mutually exclusive.

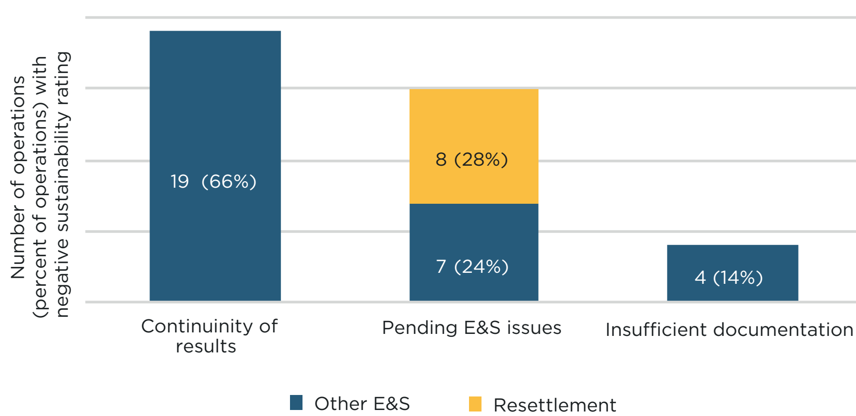
2.8 Sixty-one percent of SG investment operations (42) achieved positive efficiency ratings.[8] Twenty-seven investment operations (39%) obtained negative efficiency ratings. Among them, 13 (48%) were rated negatively for a combination of project delays and lack of credible economic analysis; seven (26%) for lack of credible economic analysis when such analysis was expected;[9] six (22%) for inefficient use of resources as shown by low rates of returns, unfavorable cost comparisons or unused infrastructure, and one for cost overruns.

2.9 Fifty-nine percent of SG operations (42) achieved positive sustainability ratings. Three factors drove negative sustainability ratings: (i) uncertain continuity of results achieved, (ii) environmental and social (E&S) issues that were not fully addressed during project execution and were pending at project closure, and (iii) insufficient documentation of management and implementation arrangements beyond project closure to verify the continuity of results or compliance with E&S safeguards (Figure 2.2). Among the 29 SG operations with negative sustainability ratings, 15 were closed while E&S issue were still pending, including eight that closed without having fully resolved resettlement issues (Box 2.1). Such practice affects the sustainability of operations and exposes IDB to reputational risks.

Figure 2.2

Drivers of negative sustainability ratings

Source: OVE.



Note: Percentages sum to more than 100% because categories are not mutually exclusive.

Box 2.1. Tracking the implementation of the resettlement safeguard (OP-710)

Among 71 SG operations validated in 2021, 11 triggered the resettlement safeguard (OP710), according to information available in Convergence and IDB's Enterprise Data Warehouse. Among these 11 SG operations,

8 PBLs are not rated on efficiency.

9 PCR guidelines stipulate that investment project efficiency is to be determined by cost-benefit or cost effectiveness analysis and that lack of such analysis normally leads to a negative efficiency rating, with some exceptions.

3 PCRs did not mention the implementation of safeguards. It was not clear whether these projects did not entail any resettlement or safeguard implementation was not documented even though it was part of project implementation. Among the eight remaining operations, six closed while resettlement was still pending. Aside from these 11 SG operations, 2 additional operations that did not trigger the resettlement safeguard during preparation included resettlement issues that had not been fully resolved at the time of project closure.

Source: OVE.

2.10 Overall project performance was similar across sectors, but performance on relevance and sustainability varied. To explore potential differences in outcome ratings across sectors, OVE reviewed all operations validated between 2017-21.[10] No statistically significant difference was found for the overall outcome, effectiveness, or efficiency ratings across sectors. In contrast, operations from the Social Sector (SCL) achieved higher relevance ratings than operations in other sectors, and operations in the Institutions for Development (IFD) Sector achieved lower relevance ratings. Sustainability had the widest variation across sectors, with all sectors except SCL revealing statistically significant differences from other sectors. Table 2.3 summarizes the ratings by sector.

Table 2.3. Percent of sovereign-guaranteed projects validated in 2017–21 with positive overall outcome ratings, by sector

Rating	Total	Climate Change and Sustainable Development (CSD)	Institutions for Development (IFD)	Infrastructure and Environment (INE)	Integration and Trade (INT)	Social (SCL)
Overall	56%	47%	56%	59%	33%	63%
Relevance	89%	86%	**81%**	93%	83%	**98%**
Effectiveness	32%	30%	31%	32%	33%	33%
Efficiency	55%	50%	58%	60%	0%	57%
Sustainability	67%	**56%**	**75%**	**56%**	100%	76%
Number of operations[a]	254	43	75	75	6	54

Source: OVE.
Note: A proportions test was performed to test for statistical significance between a sector's share of positive ratings and the average of all other sectors. Numbers in boldface indicate results that are statistically different from the rest. Given the small sample size, INT was excluded from the intersectoral statistical significance test analysis but included in the totals when calculating the ratings of other sectors for the purpose of comparison.
[a] OVE validated 254 SG operations during the 2017–21 validation cycles, including 28 PBL/PBP operations. The efficiency criterion was not assessed for these types of operations, so they were excluded from the efficiency analysis. Validation of operations BO-L1104 and BO-L1107 (a PBP series) was excluded from the sector analysis because one was classified under CSD and the other under INE.

10 Adding more observations expands the sample size and reduces uncertainty in the analysis. PCR Guidelines comparable to the current guidelines were first implemented during the 2017 validation cycle.

3. Project performance: Non-core criteria ratings

2.11 The Bank's performance was rated positive on 59% (42) of SG operations. Multiple and at times overlapping factors affected Bank performance ratings. As in 2020, the most frequently cited reasons for negative Bank performance ratings were issues at the design stage. Among the 29 operations for which Bank performance was less than satisfactory, 17 (about 60%) suffered from overly ambitious project design that underestimated the time and cost and/or overestimated implementation capacity. Seventeen (about 60%) were affected by poor M&E plans that reduced the quality of supervision and evaluability. Nine validations (about 30%) cited failure to identify risks and mitigation strategies, including understanding issues such as the physical condition of the land, long-standing social conflicts, and requirements for environmental management. In four cases (17%), lack of coordination with complementary projects limited the effectiveness of the project. Seven validations (24%) mentioned failure to ensure compliance with safeguard policies during implementation; five (17%) referred to the sub-par quality of supervision, such as lack of proactiveness and high turnover of the team leaders; and three (10%) mentioned failure to adjust the vertical logic[11] after the project scope was adjusted.

2.12 Borrower performance was rated positive on 63% (45) of operations. Among the remaining 26, Borrower performance was rated negatively for 11 operations (42%) because of lack of compliance with fiduciary obligations, such as monitoring and reporting performance indicators, keeping and reporting an adequate record of financial data, and fulfilling audit requirements. Three operations (12%) faced challenges with procurement processes, and another three (12%) failed to adhere to E&S safeguards. Administrative delay was mentioned in 14 cases (54%). As underlying causes for less than satisfactory Borrower performance, 18 validations (69%) pointed to inadequate staffing of executing units, including skill mismatch, high turnover rate, and insufficient time allocated to the task. A few operations also suffered from inadequate coordination mechanisms among government entities and frequent changes in regulations and procedures.

4. Quality of self-assessments

2.13 Differences between Management's and OVE's ratings persist. IDB Management rated overall project outcome as positive in 73% of SG operations, whereas OVE rated only 52% as positive, a slightly larger gap than last year (21 percentage points versus 16 percentage points) (Figure 2.3). OVE downgraded the overall outcome rating

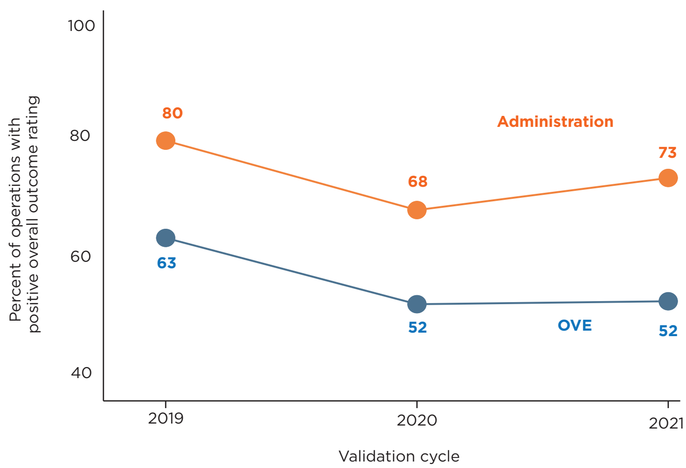
11 Vertical logic defines activities, outputs, and outcomes necessary and sufficient to achieve ultimate impacts an operation intends to achieve.

from positive to negative in 15 of the 71 validated SG operations (21%). The most downgraded criterion was sustainability (14 operations or 20%), followed by effectiveness (13 operations or 18%), efficiency (10 operations or 14%), and relevance (5 operations or 7%). There were no OVE upgrades from negative to positive in the overall outcome rating; there was one in relevance and one in effectiveness.

Figure 2.3

Percent of sovereign-guaranteed operations with positive overall project outcome ratings according to Management and OVE, 2019–21

Source: OVE.



Note: Figure excludes validation cycles and projects for which the rating scales used by Management and OVE were not comparable.

2.14 The quality of 47 of 71 PCRs (66%) was rated excellent or good. OVE rates the quality of the PCRs submitted by Management using a four-point scale (excellent, good, fair, or poor). Of the 71 PCRs, 3 were rated excellent (4%), 44 good (62%), 21 fair (30%), and 3 poor (4%). Among the 24 PCRs rated as fair or poor, the most common shortcomings identified by OVE included insufficient information to properly justify some of the claims (23 PCRs or 96%); the low quality of some of the analyses presented (17 PCRs or 71%); and insufficient transparency or candidness, including on safeguard performance (11 PCRs, or 46%).[12]

B. Private sector projects reviewed in 2021

2.15 Management prepared XSRs for 38 operations, all of which were reviewed by OVE during the 2021 validation process. Of these operations, 27 (71%) had reached EOM or exited the portfolio before

12 Percentages do not add up to 100% because categories are not mutually exclusive. Other reasons for low PCR quality ratings included inconsistencies between the analyses presented and the ratings provided (42%), data quality issues (42%), inconsistencies in reporting (33%), poor observance of PCR Guidelines (25%), and the presentation of lessons learned of little depth (13%).

reaching EOM in 2019.[13] Eleven (29%) had reached EOM or exited the portfolio before reaching EOM between 2016 and 2018[14] but without XSRs having been prepared during the previous validation cycles. The operations evaluated were approved between 2007 to 2017, with a large share (28 or 74%) being approved in or after 2015.[15] These 38 XSRs covered 32 senior loans, three equity investments, three guarantees, and two subordinated loans.[16] Jointly, these operations represent US$1.1 billion in approvals. Table 2.4 breaks these operations down by business segment.

Table 2.4. XSRs for NSG operations validated in 2021, by business segment

Business segment	NSG operations validated	Approved amount (US$ million)
Infrastructure and Energy	15	713
Financial Institutions	12	293
Corporates	11	128
Total non-sovereign-guaranteed operations	**38**	**1,134**

Source: OVE, with data from IDB Invest (2021).

2.16 In the process of improving the system to determine EOM, Management identified operations whose XSRs are still pending. XSR Guidelines stipulate that XSRs are required soon after the projects reach EOM or exit the portfolio by being paid off, sold, or written off. In 2018, OVE's validation report (RE-530-2) recommended that IDB Invest agree with OVE on an independent review process to verify the operations reaching EOM and therefore subject to an XSR every year. In response to this recommendation, IDB Invest developed guidelines to determine EOM and launched a process to implement the improved identification system in 2020. This improved system revealed 10 operations (approved amount of US$367 million jointly) that had reached EOM or exited the active portfolio and two (approved amount of US$8 million jointly) that had closed out of Special Assets in 2019, for which XSRs were not prepared in time for this validation cycle. It also identified two that reached EOM in 2019 while in Special Assets. The XSRs validated this year therefore only present a partial picture of the performance of the operations that reached EOM or exited the portfolio in 2019.

13 EOM signifies that IDB Invest made material disbursement and received at least one set of audited financial statements covering 12–36 months of operation after the disbursement/project completion. The exact definition of EMO varies by project type. Among 27, 25 have reached EOM and two exited the portfolio in 2019.

14 Breakdown of the previous EOM years reached are as follows: 1 (2016), 2 (2017), 7 (2018). One was repaid in 2017.

15 Breakdown of the approval years are as follows: 1 (2007), 1 (2008), 1 (2011), 2 (2012), 2 (2013), 3 (2014), 14 (2015), 10 (2016), 4 (2017).

16 The total exceeds 38 because two operations had both senior loans and guarantees.

In addition, Management identified 37 operations (approved amount of US$713 million jointly) that had reached EOM or exited the active portfolio during 2016–2018 and 11 that closed out of Special Assets for which XSRs had not been prepared during the corresponding validation cycles (2018–2020) (Table 2.5). Management has proposed to prepare the missing XSRs for operations for which IDB Invest continues to have a relationship with the client during the 2022 validation cycle, and cover those for which only limited information can be obtained at this point in a one-off report. OVE will include a review of this one-off report in its 2022 validation report.

Table 2.5. Operations without XSRs that reached EOM or exited the portfolio during 2016–2019

Operations in or out of the active portfolio	#	Approved amount (US$ million)**
Projects that reached EOM or exited the portfolio pre-EOM in 2019	10	367
Projects that reached EOM or exited the portfolio pre-EOM in 2016-2018	37	713
Total	**47**	**1080**
Operations in or out of the Special Assets portfolio	**#**	**Approved amount (US$ million)****
Projects that reached EOM in Special Assets in 2019*	2	9
Projects closed (repaid or write-off) out of Special Assets in 2019	2	8
Projects that reached EOM in Special Assets in 2016-2018*	7	377
Projects closed (repaid or write-off) out of Special Assets in 2016-2018	11	460
Total	**22**	**853**

Source: Data provided to OVE by IDB Invest.
Notes: * Not subject to XSR while in Special Assets.
** Approved amounts may be higher than the actually disbursed amounts.

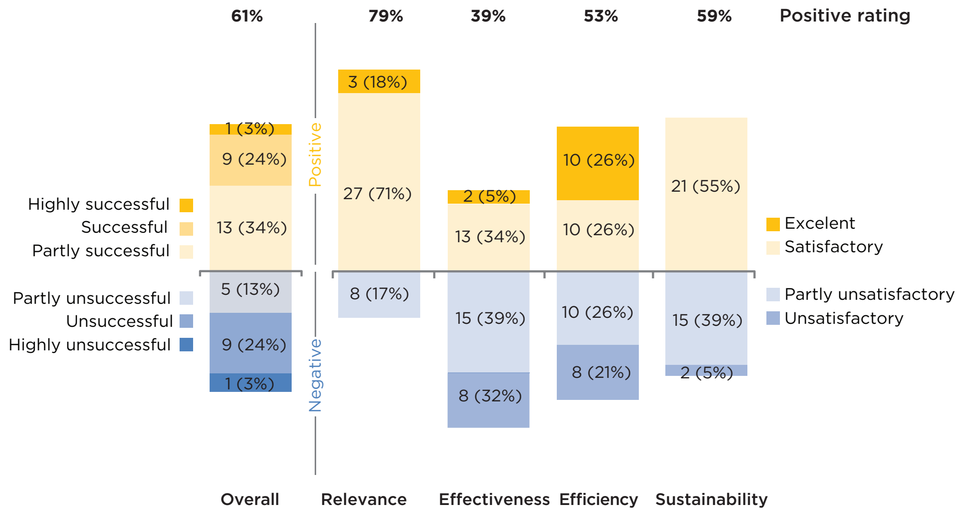
1. Project performance: Overall outcome ratings

2.17 Sixty-one percent of NSG operations (23) achieved positive overall outcome ratings. This figure falls somewhat short of the target of 65% set in the 2020–23 Corporate Results Framework. The ratings of 95% of operations fell in the middle four categories off successful, partly successful, partly unsuccessful or unsuccessful; only one operation each was rated highly unsuccessful and another highly successful (Figure 2.4).

Figure 2.4

Overall project outcome and core criteria ratings for non-sovereign-guaranteed operations validated in 2021

Source: OVE.



2. Project performance: Core criteria ratings

2.18 Relevance continues to be the highest-rated core criterion, with 30 (79%) NSG operations achieving positive ratings. Eight operations (21%) obtained negative relevance ratings. Among them, six (75%) had project designs that did not respond to the market's needs, and five (63%) were not well aligned with the country's development needs.[17]

2.19 Effectiveness was the lowest-rated core criterion, with just 15 (39%) NSG projects achieving positive ratings. As in previous cycles, poor performance, meaning failure to achieve targeted results, and poor M&E were the key reasons for negative effectiveness ratings, with several of the negatively rated operations having been affected by both factors. Of 23 operations with negative effectiveness ratings, 18 (78%) were affected by poor performance and 9 (39%) by poor quality of M&E that made it impossible to evaluate whether the objective was achieved or attributable to the operation. Given consistently low effectiveness ratings, Chapter III further explores factors affecting project effectiveness.

2.20 Fifty-three percent of NSG operations (20) achieved positive efficiency ratings. Efficiency of NSG operations is assessed based on the extent to which the financial and economic benefits of the project exceed project costs. Among the 18 operations (47%) that obtained negative efficiency ratings, 12 (67%) did not achieve a financial return that was higher than the cost of capital. Half of those operations (6) also failed to achieve adequate economic rates of return. Three of these (17%) were in the financial sector. Two achieved positive financial performance, but their economic

17 The number of operations adds up to more than eight because categories are not mutually exclusive.

performance was judged less than satisfactory because the targeted portfolio (e.g., loans to small farmers) did not grow. In the remaining operation, the targeted portfolio was not profitable.

2.21 Fifty-five percent of NSG operations (36) achieved positive sustainability ratings. No operation obtained an excellent rating. Among the 17 operations (45%) with negative sustainability ratings, 10 (59%) received low scores because of the fragility of results continuation and 7 (41%) because of E&S compliance shortfalls, including incomplete actions by the EOM of the project.

2.22 Infrastructure and Energy operations had a higher share of positive overall outcome ratings than other business segments. To assess potential differences in performance ratings across business segments, OVE reviewed all operations validated between 2018–21.[18] Infrastructure and energy operations outperformed others in terms of positive outcome ratings (Table 2.6). Although there was substantial variation in the share of projects with positive effectiveness ratings across the three business segments, the other differences were not statistically significant.

Table 2.6. Percent of non-sovereign-guaranteed projects validated in 2018-2021 with positive overall outcome ratings, by business segment

Rating	Total	Corporates	Financial Institutions	Infrastructure and Energy
Overall outcome	55%	49%	52%	**71%**
Relevance	68%	77%	**57%**	**89%**
Effectiveness	44%	36%	43%	57%
Efficiency	47%	51%	**41%**	61%
Sustainability	55%	51%	51%	71%
Number of projects	151	39	84	28

Source: OVE.
Notes: A proportion's test was performed to test for statistical significance between a business segment's share of positive ratings and the average of all other business segments. Numbers in boldface indicate a statistically significant difference from the average of other sectors.

3. Project performance: Non-core criteria ratings

2.23 About two-thirds of IDB Invest projects (26) achieved positive investment outcomes, and 89% (34) demonstrated additionality. Among the 12 operations validated in 2021 that did not achieve positive investment outcome, eight (67%) were senior loans that were repaid, three (25%) were equity investments for which IDB Invest either received lower than expected returns or was not able

18 Adding more observations expands the sample size and reduces uncertainty in the analysis.

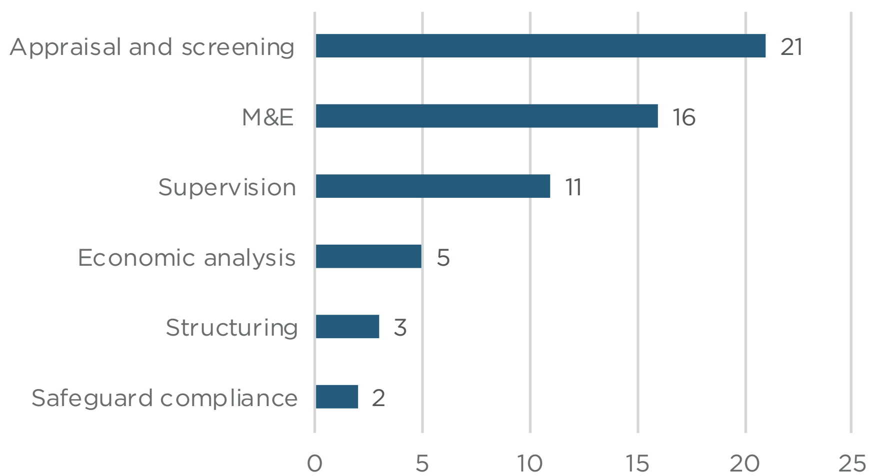
to exit as planned, and one (17%) was a senior loan that defaulted. Three out of four operations with negative additionality ratings did not deliver financial additionality.

2.24 Work quality remained the lowest-rated non-core criterion, considered satisfactory for only 29% (11) of operations. The most frequently cited reasons for negative ratings were shortcomings during appraisal and screening, such as weaknesses in financial analysis; partner identification; and mitigation of foreign exchange, policy, or construction risks (Figure 2.5.) The second-most frequently cited reasons were M&E frameworks with missing or poorly defined indicators and the low quality of supervision, including limited communication with the client, inability to adapt to changing environments, and failure to collect key data. Other reasons included inadequate articulation of development objectives or economic analysis and failure to ensure compliance with E&S safeguards.

Figure 2.5

Causes of negative work quality ratings of non-sovereign-guaranteed operations

Source: OVE.



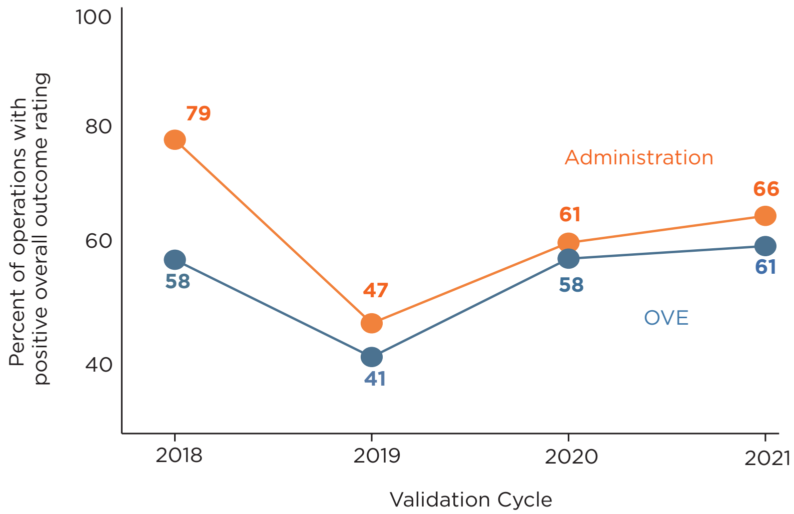
4. Quality of self-assessments

2.25 Differences between Management's and OVE ratings were relatively small. IDB Invest Management rated the overall outcome as positive in 25 of the 38 NSG operations (66%)—close to OVE's share (23 of 38 projects, or 61%) (Figure 2.6). In 2021, OVE downgraded the overall outcome rating from positive to negative for 2 of the 38 validated NSG operations (5%). The core criteria with the most downgrades from positive to negative was effectiveness (6 NSG operations), followed by sustainability (3 operations) and relevance (2 operations). OVE did not upgrade the overall outcome rating of any NSG operation from negative to positive, but there were two upgrades in relevance, two in effectiveness, and two in efficiency.

Figure 2.6

Percent of non-sovereign-guaranteed operations with positive overall project outcome ratings according to Management and OVE, 2018–21

Source: OVE.



Note: Figure excludes validation cycles and projects for which the rating scales used by Management and OVE were not comparable.

2.26 The quality of 34 of 38 XSRs (89%) was rated excellent or good. Five (13%) were rated excellent, 29 (76%) good, and 4 (11%) fair. No XSRs were rated poor. Among the four XSRs rated fair, the most common shortcoming was insufficient information (three XSRs). Other challenges included specifying development objectives that were not exclusively business-oriented (two XSRs), presentation of lessons learned of little depth (two), data quality issues (two), poor observance of XSR Guidelines (two), inconsistencies between the analyses presented and the ratings provided (one), and insufficient transparency or candidness (one).

03

Understanding Efectiveness

3.1 Effectiveness, or achievement of development objectives, has consistently been rated the lowest among the four core criteria across IDB Group operations.[19] Only 31% of SG operations and 39% of NSG operations achieved positive effectiveness ratings in 2021. Low effectiveness ratings play a significant part in the below-target overall project outcome ratings because effectiveness represents 40% of the overall outcome rating in the case of SG and NSG investments and 60% in the case of PBLs.[20] This section reviews the achievement of specific development objectives, as well as whether low effectiveness ratings are correlated with certain characteristics of ineffective operations, such as design shortfalls and delays in execution.[21] Annex I summarizes how effectiveness ratings are assigned according to current PCR and XSR Guidelines.

A. Public sector operations

1. Types of specific objectives and factors affecting their achievement

3.2 OVE reviewed achievement of specific objectives, the building blocks of effectiveness ratings. The effectiveness rating is based on an assessment of multiple specific objectives. The 71 SG operations validated in 2021 together had 177 specific objectives, of which 101 (57%) were rated less than satisfactory. Among these, 66 (66%) were affected by poor performance,[22] 12 (12%) by cancellations[23] and fifty-one (50%) by poor quality M&E.[24,25] The latter hindered credible measurement of achievement and its attribution to the operation (Figure 3.1). The most frequent M&E issue is insufficient

19 Effectiveness was the lowest-rated criteria in all comparable periods (every validation cycle since 2017 for SG operations and since 2018 for NSG operations).

20 The Corporate Results Framework 2020–23 tracks the share of projects with satisfactory development outcome ratings at completion. It sets the target at 70% and above for SG operations and 65% and above for NSG operations.

21 Assessing by how much each of these factors affects the effectiveness ratings would require regression analysis. The limited number of PCRs available at this time prevents such analysis.

22 Poor performance refers to failure to achieve the targeted results. Consistent with the ratings criteria in the PCR guidelines for a less than satisfactory effectiveness rating, poor performance of an objective is defined as an objective that met 79% or less of its targets.

23 OVE's 2020 Validation Report analyzed cancellations and identified 12 operations for which cancellations affected the achievement of results. In this year's report, 12 specific objectives affected by cancellations correspond to 12 operations. In other words, 25% of all SG operations validated last year had a negative effectiveness rating due at least partially to cancellations, whereas that share is 17% (12 out of 71) for operations validated in 2021.

24 Poor quality of M&E refers to shortcomings in the way achievement is measured. They include missing result indicator, inadequate result indicator that do not measure the achievement of results intended, misclassification of output as a development outcome indicator, unclearly defined baseline and target, lack of information (lack of monitoring), and inability to plausibly attribute the results achieved to the IDB operation.

25 The number of objectives does not add up to 101 and the percentages do not sum to 100% because categories are not mutually exclusive.

information to assess achievement of expected results, followed by lack of attribution and missing indicators (Figure 3.2). These findings mirror the discussion in paragraph 2.7, which shows that poor performance, poor M&E, and cancellations were the key factors affecting overall project effectiveness ratings.

Figure 3.1

Objectives of sovereign-guaranteed operations affected by poor performance, poor quality of M&E, or cancellations

Source: OVE.

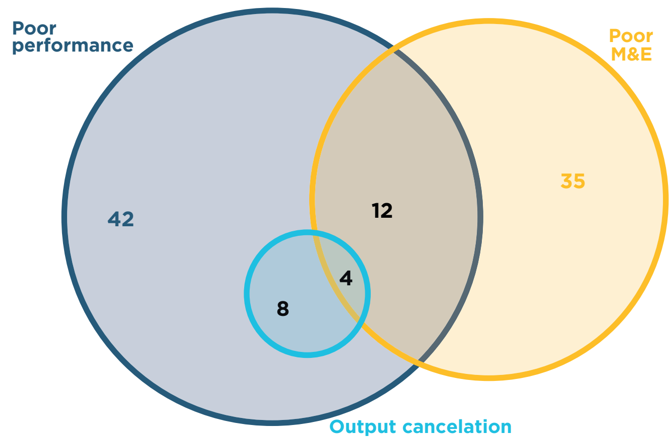
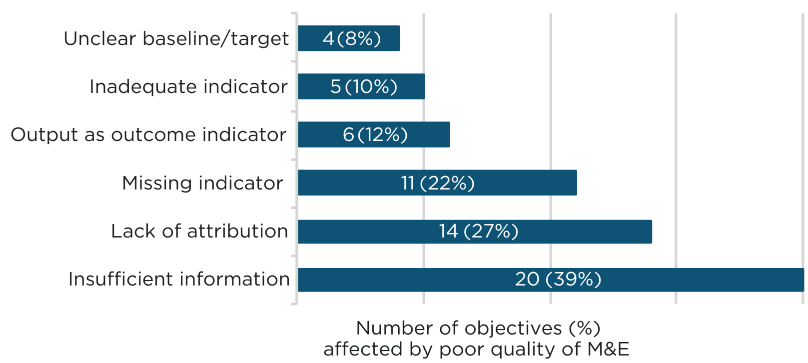


Figure 3.2

Types of M&E issues affecting measurement of objectives of sovereign-guaranteed operations

Source: OVE.



Note: Figure 3.2 shows share of 51 objectives affected by poor quality of M&E. The number of affected objectives does not sum to 51 because categories are not mutually exclusive.

3.3 Specific objectives aimed at improving service provision were most likely to achieve positive ratings; objectives aimed at strengthening institutional capacity and facilitating sector level changes were least likely to do so (Figure 3.3). To gain further insights into project effectiveness ratings, OVE grouped the 177 objectives of the 71 SG operations assessed into five categories: 77 (44%) related to better service provision, 42 (24%) institutional capacity building, 11 (6%) improved environmental sustainability, and 9 (5%) sector-level changes. The remaining 38 (21%) included various other objectives that lacked critical mass. (Table 3.1). Objectives related to improving service provision were found to have the highest share of positive ratings, while those related strengthening institutional capacity and affecting sector level changes had the lowest share, although these results need to be treated with caution, given the limited number of observations in some categories. Performance also varied within each type of objective. Among operations that aimed to improve service provision, objectives to improve physical infrastructure quality were more likely and objectives aimed to reduce the cost of services were less likely to be achieved. Among objectives to build capacity, those that aimed to build government capacity were more likely and those targeting service providers were less likely to be achieved.

Table 3.1. Types of objectives of sovereign-guaranteed operations

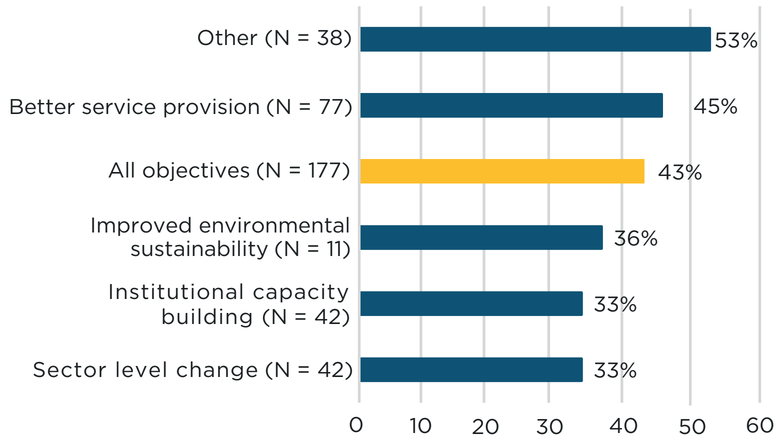
Type of objective	Number	%	Breakdown (Number of objectives)
Better service provision	77	44%	Includes objectives to improve quality (28), expand access to services (18), improve service provision through better physical infrastructure (18), reduce cost (8), and other type of improvements (5).
Institutional capacity building	42	24%	Includes objectives to build the capacity of service providers (20), the government (11), and others (11). *
Improved environmental sustainability	11	6%	Includes objectives to reduce greenhouse gas emissions (4), minimize other types of negative environmental impact (3), promote conservation (2), and control pollution (2).
Sector-level change	9	5%	Includes objectives to develop and diversify tourism sectors (4), develop a legal/policy/strategic framework for infrastructure sectors (4), and develop institutional and executive framework for less fragmented judicial defense of the country (1).
Other	38	21%	Includes various objectives that are not classified as any type of objectives listed above such as increased foreign trade, fiscal sustainability, security improvement, disaster prevention, etc.

Source: OVE.
Note: * When capacity building is provided to public entities to strengthen its public policy functions (e.g., sector coordination, policy-making), 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.

Figure 3.3

Percent of objectives of sovereign-guaranteed operations achieved, by type of objective

Source: OVE



3.4 Objectives related to institutional capacity building were more likely to be affected by shortcomings in M&E than those related to service provision, environmental sustainability and sector level change. Table 3.2 shows how the three correlates of negative effectiveness ratings affected the various types of objectives. Negatively rated objectives targeting better service provision, improved environmental sustainability, and sector-level changes were more likely to be affected by poor performance than other correlates, while institutional capacity building objectives were more likely to be affected by poor underlying M&E than other correlates. The share of objectives affected by cancellations was similar across objectives.

Table 3.2. Factors associated with negative objective ratings of sovereign-guaranteed operations, by type of objective

Reason for negative objective rating	Tipo de objetivo					
	Better service provision (N = 42)	Institutional capacity building (N = 28)	Improving environmental sustainability (N = 7)	Sector-level change (N = 6)	Other (N=18)	Total (N=101)
Poor performance	33	14	5	5	9	66
Poor M&E	15	20	2	3	11	51
Output Cancellations	5	3	1	1	2	12

Source: OVE.
Note: Cells highlighted in brown represent the most frequent reason for negative effectiveness by type of objective.

2. Characteristics of operations associated with negative effectiveness ratings

3.5 SG operations with flaws in project design were less likely to achieve positive effectiveness ratings than operations without such flaws. Possible design flaws were proxied by identifying shortcomings in either the operation's alignment with country realities, its vertical logic, or a negative rating in the Bank's performance at entry. Among 71 SG operations validated in 2021, 63% (45) had some shortcoming at the design stage. Only 13% (6) of these operations achieved positive effectiveness ratings, compared with 62% (16) of operations without design flaws.[26]

3.6 Extensions in the disbursement period are correlated with negative effectiveness ratings. A correlation analysis was performed to assess whether the time spent in reaching legal effectiveness, eligibility, and the extension in the disbursement period is associated with a negative effectiveness rating for SG operations. For the operations validated in 2021, there is a statistically significant, albeit moderate (0.28), positive correlation between the extension of the disbursement period and negative effectiveness ratings. To corroborate this finding, OVE also reviewed all operations validated in the past five years and found a similar correlation (0.21).[27] No statistically significant

26 The differences between the effectiveness ratings for operations with and without design shortcomings are statistically significant.

27 The oldest comparable PCRs were validated in 2017. The cutoff years are determined by the year that guidelines that are comparable to the guidelines used in the 2021 cycle were implemented.

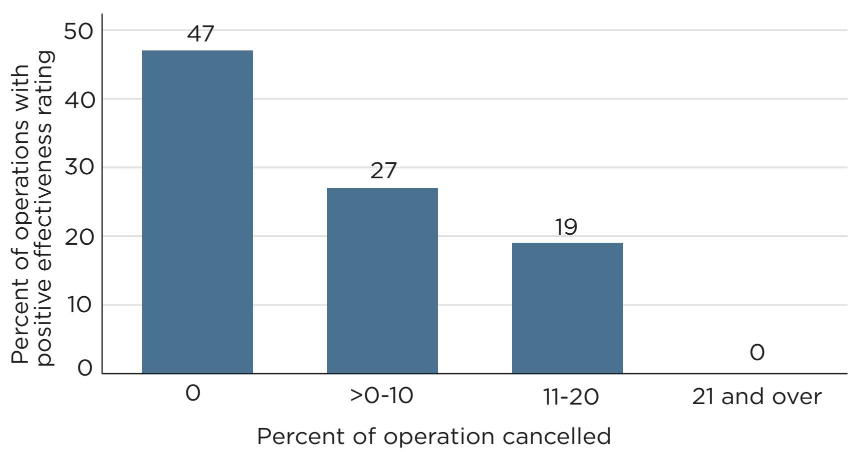
relationship was identified between negative effectiveness ratings and the number of days the project spent reaching legal effectiveness or eligibility.

3.7 Given that cancellations continued to be a notable factor in negative effectiveness ratings in 2021, OVE looked at how partial cancellations affected effectiveness ratings. Nine operations (13%) cancelled more than 20% of their approved amount. None of these operations obtained positive effectiveness ratings. In contrast, 35% of the remaining 62 operations did so.[28] To corroborate the finding, OVE reviewed all operations validated over the past five years and found a strong correlation between cancellations and effectiveness ratings. Sixty percent of operations validated in this period had partial cancellations, though most of the cancellations represented a small share of the loan amount (a median of 2% of the original approved amount). The top decile of operations (25 of 254) cancelled 20% or more.[29] Every one of these operations failed to achieve a positive effectiveness rating, reflecting their inability to meet half or more of their objectives with attribution. In contrast, 65% of operations without large cancellations achieved positive effectiveness ratings.[30] Figure 3.4 groups operations by the size of their cancellations. It shows a clear trend: The share of projects with positive effectiveness ratings decreases with the size of the cancellation.

Figure 3.4

Relationship between effectiveness ratings and partial cancellations in sovereign-guaranteed operations

Source: OVE, with cancellations data from IDB (2021a).



Note: Figure is based on data for 254 SG operations validated by OVE between 2017 and 2021.

28 In their PCRs, all but one operation that had a large cancellation reported that it had affected the achievement of objectives in some way (the one operation that reported otherwise obtained a negative effectiveness rating because of attribution problems and underachievement in other objectives).

29 The 90th percentile of the percentage cancelled is 19.5% of the original approved amount.

30 This difference is highly statistically significant, suggesting a strong association between large cancellations and underachievement of objectives. A simple, two-variable linear probability model of a binary indicator of whether the project had a negative effectiveness rating regressed on a binary indicator of whether the project had a large cancellation yields a mean difference of 35 percentage points, with a p-value < 0.001 using robust standard errors ($N = 254$; $R2 = 0.0511$).

B. Private sector operations

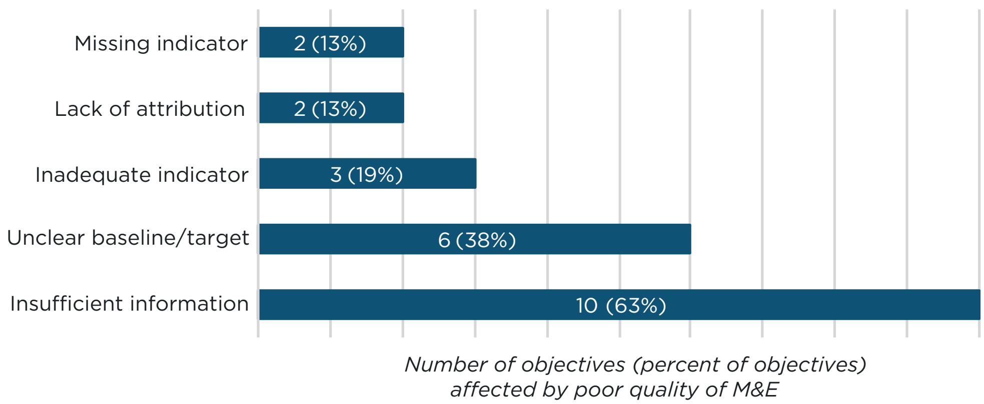
1. Types of objectives and factors affecting their achievement

3.8 To further understand what causes negative effectiveness ratings among NSG operations, OVE also looked at individual objectives. The 38 NSG operations validated in 2021 together had 90 objectives, of which 48 (53%) were rated less than satisfactory. Among these, 32 (67%) were rated negatively because of poor performance (failure to achieve targeted results), 11 (23%) because of the poor quality of underlying M&E framework, and 5 (10%) because of a combination of both factors. The main cause of poor M&E was insufficient information to assess indicator achievement, followed by lack of a baseline or target for some of the indicators (Figure 3.5).

Figure 3.5

Types of M&E issues affecting measurement of objectives of non-sovereign-guaranteed operations

Source: OVE



Note: Figure 3.5 shows reasons for poor M&E for the 16 objectives rated negatively at least in part because of poor-quality M&E. Percentages do not sum to 100% because categories are not mutually exclusive.

3.9 Specific objectives aimed at expanding renewable energy capacity were most likely to achieve positive outcome ratings; objectives aimed at creating jobs were least likely to do so. OVE grouped the 90 objectives of the 38 NSG operations into six broad categories. (Table 3.3). Although these results need to be treated with caution given the limited number of observations, objectives focused on increasing renewable energy generation, reducing GHG emissions, and strengthening export and production in priority sector scored better than objectives aimed at supporting SME growth or creating jobs (Figure 3.6).

Table 3.3. Types of objectives of non-sovereign-guaranteed operations

Type of objective	Number	%	Breakdown (Number of objectives)
Support for SME development	21	23%	Includes support for SME development through improved access to financing (15) and other support (6).

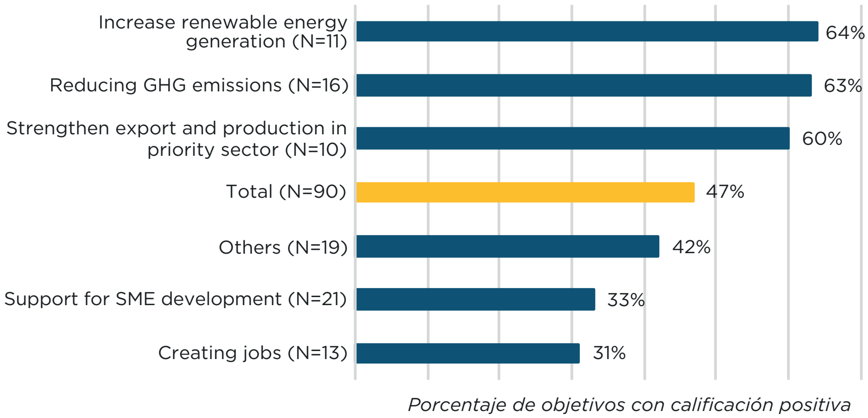
Type of objective	Number	%	Breakdown (Number of objectives)
Reducing GHG emissions	16	18%	Includes reduction of GHG emissions from expanded use of renewable energy (12), improved energy efficiency (3), reduction of emission through import substitution (1).
Creating jobs	13	14%	All 13 objectives aimed to create jobs.
Increase renewable energy generation	11	12%	All 11 objectives aimed to increase renewable energy generation through added capacity.
Strengthen export and production in priority sector	10	11%	Includes increased exports (5) and increased production/ customer reach in priority sectors (5).
Others	19	21%	Includes increasing productivity through better and lower cost infrastructure (6) better financing to non-SME clients (4), sector-level change such as developing capital markets, promoting the new industry, or increasing the county's oil reserves (4), expanded access to non-infrastructure services such as health care and housing (3), supporting expansion of enterprise with positive social impact through investment (2).

Source: OVE.

Figure 3.6

Percent of objectives of non-sovereign-guaranteed operations achieved, by type of objective

Source: OVE



3.10 All major types of objectives were more likely to be rated negatively due to poor performance than because of poor M&E. Table 3.4 shows the share of objectives that were negatively affected by the two main reasons for negative effectiveness ratings. The increase of renewable energy generation, reduction in GHG emissions, strengthening in export and production in priority sectors, and support for SMEs development were rated negatively predominantly because of poor performance. Other objectives rated negatively were at least as likely to have received a negative raging because of poor M&E.

Table 3.4. Factors associated with negative objective ratings on non-sovereign guaranteed operation, by type

Reason for negative objective rating	Tipos de objetivos						
	Support SME development (N=14)	Reduce GHG emissions (N=6)	Create job (N=9)	Increase renewable energy generation (N=4)	Strengthen export and production in priority sector (N=4)	Others (N=11)	Total (N=48)
Poor performance	13	5	6	4	3	6	37
Poor M&E	2	1	5	0	2	6	16

Source: OVE.
Note: Cells highlighted in blue represent the most frequent reason for negative effectiveness by type of objective.

3.11 NSG operations with flawed project designs were statistically less likely to be effective than other operations. Possible design flaws were proxied by identifying shortcomings in either the operation's alignment with country realities or negative ratings in work quality at the screening, appraisal, or structuring stage. Among the 38 NSG operations validated in 2021, 28 (74%) had some shortcoming at the design stage. Thirteen of them (46%) achieved positive effectiveness ratings, compared with 70% of operations without design flaws.[31] It would be useful to explore the correlation between effectiveness ratings and performance flags raised during supervision. However, the current IDB Invest supervision system has been fully operational only since 2018, and insufficient data were available for the NSG operations reviewed by OVE in 2021 to conduct such analysis. Similarly, difficulty obtaining data on partial cancellation makes it infeasible to conduct correlation analysis between effectiveness rating and partial cancellation.

31 The differences between the effectiveness rate of operations with and without design shortcomings are statistically significant.

04

Lessons from Cancellation Notes of Sovereign-Guaranteed Operations

4.1 Between validation years 2018 and 2020, IDB prepared cancellation notes for SG operations that were closed before significant execution took place.[32] Given the role partial cancellations have played in low effectiveness rating, it is worth reviewing the reasons behind cancellation documented in these notes.[33] The cancellation notes were introduced by the 2018 PCR Guidelines (OP-1696-3) as a shortened and adapted version of the PCR to discuss results achieved, factors that determined the cancellation of the project, and findings and recommendations. With the update of PCR Guidelines in 2020, operations cancelled in validation cycle 2021 or later are no longer required to include cancellation notes.[34] During the three years during which they were mandated, IDB produced 15 cancellation notes. These operations represented a combined approved amount of US$971 million and disbursements of US$172 million. Table 4.1 shows the breakdown of projects with cancellation notes by validation year.

Table 4.1. Cancellations of sovereign-guaranteed operations reviewed, 2016–18 (US$ million, except where indicated otherwise)

Cancellation year[a]	Number of operations	Approved amount	Cancelled amount	Disbursed amount	Total preparation and supervision expenses	Total staff weeks in preparation and supervision
Operations canceled before significant execution						
2017	8	605.6	454.4	151.3	1.3	1,038.5
2017	2	140.0	135.1	4.9	0.6	281.3
2018	5	225.7[b]	210.5	15.2	1.6	816.6
Subtotal	15	971.3	800.0	171.4	3.5	2,136.4
Operations canceled before any disbursement						
2016	5	198.0	198.0	0.0	0.2	197.4
2017	9	665.9	665.9	0.0	0.7	460.7
2018	2	280.8	280.8	0.0	0.9	130.7
Subtotal	*16*	*1,144.7*	*1,144.7*	*0.0*	*1.8*	*788.8*
Total	**31**	**2,116.0**	**1,944.7**	**171.4**	**5.3**	**2,925.2**

Source: OVE.
Note: [a]For the 15 operations cancelled before significant execution, the cancellation year used is the operation's CO year. For the 16 operations that were fully cancelled before any disbursement, the year when the cancellation was reflected in the operation's financial balances is used, as no CO year is reported for these cases. b. Figure includes US$36.3 million in cancellations for operation EC-L1107, although these funds remained committed, as they were reallocated to operation EC- L1216 because of a natural disaster emergency.

32 For investment projects, significant execution is defined as a total disbursed amount exceeding 30% of the amount initially committed by the Bank. For PBLs, significant execution is defined as declaration of legal effectiveness. PCRs are not required for operations that do not reach these stages.

33 Operations that are cancelled after significant execution took place are required to have PCRs. However, these PCRs do not systematically elaborate why partial cancellation occurred.

34 Though some documents are captured in Convergence, including a results matrix, project cost, and cancellation requests from the client, Guidelines do not require documentation of IDB's own analysis on reasons for and lessons from cancellation.

4.2 OVE also identified another 16 SG operations that were fully cancelled during validation years 2018–20. These operations were approved but never reached the disbursement stage (cancellation notes were not required for them).[35] They represent a combined approved amount of US$1.1 billion. The most frequent reason for these cancellations (12 of the 16 cases) was that the government did not sign the contract by the due date, often after having requested several extensions;[36] nine of these operations were with Brazilian subnational governments. Other reasons included lack of legislative approval (two cases) and the government's inability to secure enough of its own resources to honor previously agreed commitments (one case). In the remaining case, no documentation was available.

4.3 These cancelled operations cost the Bank US$5.3 million in preparation and supervision expenses and nearly 3,000 weeks of staff time. Operations cancelled before significant execution cost the Bank US$3.5 million in preparation and supervision expenses using 2,136 weeks of staff time (see Table 4.1). The Bank also spent US$1.8 million and dedicated 789 weeks of staff time preparing and supervising the 16 operations that were fully cancelled. All 31 cancelled operations cost the Bank US$5.3 million in expenses and 2,925 weeks of staff time.

4.4 All 15 operations with cancellation notes were cancelled because of multiple factors. Most prominent among them were issues that hindered the capacity of executing units. Many cancellation notes highlighted the need to train the executing agency in procurement, safeguards, Bank methodology, and tools to monitor and evaluate project progress and strengthen their analytical capacity. They also frequently mentioned the need to simplify coordination mechanisms, ensure that the executing unit has an appropriate level of autonomy and authority, improve communication and coordination among executing agencies through workshops, and clarify the roles and responsibilities of participating agencies.

4.5 The second-most frequently cited reason for cancellation was changes in government policies and processes. In order to mitigate changes in sector priorities, several cancellation notes recommended sector-level dialogue with relevant stakeholders

35 OVE identified these total cancellations by reviewing the operation's monthly financial balances during 2016–19 (IDB 2021a). Included here are only operations that would have needed to produce a PCR per the PCR Guidelines had they not been cancelled.

36 In eight of the 18 operations, Management produced an internal cancellation memo (different from the cancellation notes submitted to OVE) explaining why the operation had been cancelled. In nine cases, no formal cancellation memo was available but other relevant documentation was uploaded in the systems (mainly communications between the government and the Bank) that made the reasons for the cancellation clear. Only in one case (DR-L1070) was no related documentation available in the Convergence system.

to help build momentum. Some emphasized the importance of starting the dialogue early during preparation stage; others highlighted the importance of including the private sector in such dialogue. Recommendations for mitigating the changes in government policies and processes included realistically assessing the level of counterpart contributions.

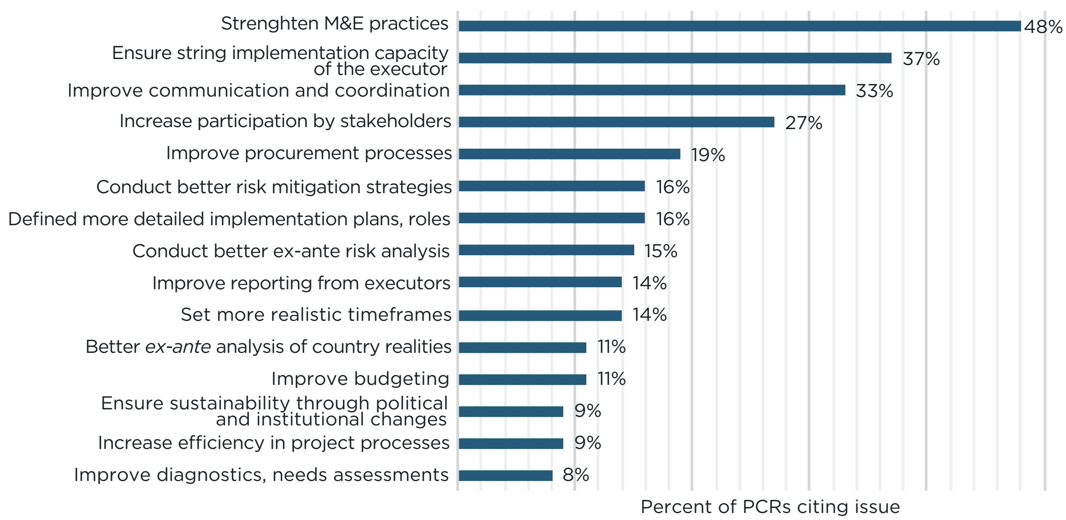
4.6 Five operations experienced excessive delays, three of them in reaching eligibility. Reasons for delays included lack of agreement on the need to comply with IDB's safeguard policies, shortages of construction materials in the local market, underestimation of project cost, IDB's financial charge, and a corruption scandal that paralyzed the executing unit.

4.7 Although these lessons were drawn from a relatively small number of cancellation notes, they are largely consistent with lessons from the 2017–20 validation exercises, presented in last year's report. Ensuring strong implementation capacity of the executing unit, improving communication and coordination among agents involved in the project, increasing participation from stakeholders, and improving procurement processes were four of the five most-cited lessons from recently closed operations (Figure 4.1).

Figure 4.1

Lessons learned from PCRs conducted in 2017–20

Source: OVE



Source: OVE, 2020. OVE's categorization; some lessons learned text data was partly extracted from the PCR Lessons Learned Dashboard and completed by other text extracted by OVE from the full PCR reports.
Note: Figure is based on 177 PCRs reviewed. Percentages do not sum to 100% because categories are not mutually exclusive. Only the 15 most frequently mentioned categories (out of 27 identified) are included.

05

Conclusions and Recommendations

5.1 OVE validated 71 PCRs and 38 XSRs in 2021. Four PCRs with CO 2019 were not completed in time for validation and will be reviewed in 2022. On the NSG side, this report only provides a partial picture of NSG operations that reached EOM or exited the portfolio in 2019. The crop of XSRs validated this year excludes 10 operations that reached EOM or exited the portfolio in 2019, and two that had closed out of Special Assets in 2019. On the other hand, as a result of Management's improvements in the system to determine EOM, XSRs for an additional 11 operations that had reached EOM in prior years were also included in this year's crop.

5.2 Fifty-two percent of SG operations achieved positive overall project outcome ratings. This figure falls short of the target of 70% set in the 2020–23 Corporate Results Framework. As in previous validation cycles, effectiveness was the lowest-rated criterion (31% of operations rated positive), driving down SG operations' overall development outcome ratings. Among specific objectives, those aimed at improving service provision were most likely to achieve positive ratings, and those aimed at strengthening institutional capacity were least likely to do so. Aside from poor performance, correlates of low effectiveness ratings include poor-quality M&E, cancellations, design shortcomings and extensions in disbursement periods. Analysis of cancellations across all operations validated since 2017 shows that the likelihood of not achieving project objectives increases with the size of cancellations. However, PCRs do not consistently discuss the effects of such cancellations on project performance.

5.3 Sustainability of project results was uncertain for over 40% of SG operations validated in 2021. A little over half the operations with negative sustainability ratings closed while E&S safeguards issues, including resettlement, were still pending. Such practice affects the sustainability of operations and exposes IDB to reputational risks. Yet, PCRs do not consistently discuss how environmental and social safeguards issues were addressed when safeguards were triggered.

5.4 Among SG operations validated in 2021, Bank performance and Borrower performance were rated positive for 59% and 63% respectively. Issues at the design stage—namely, overly ambitious designs, inadequate vertical logic, and failure to identify risks and mitigation strategies—as well as inadequate attention to E&S safeguards during design and implementation drove low Bank performance. Lack of compliance with fiduciary obligations and administrative delays were the most-cited reasons for negative borrower ratings. In many cases, inadequate staffing of executing units was highlighted as an underlying cause.

5.5 Thirty-one SG operations amounting to a total approved amount of US$2.1 billion were cancelled in part or in full between 2016 and 2018. For the 15 operations cancelled after significant disbursements, IDB prepared cancellation notes, in line with 2018 PCR Guidelines. Under the updated PCR Guidelines, preparation of cancellation notes will be discontinued, raising the question of how reporting on and learning from cancellations will be ensured.

5.6 Sixty-one percent of NSG operations (23) achieved positive overall project outcome ratings. This figure falls somewhat short of the target of 65% set in the 2020–23 Corporate Results Framework. As in the case of SG operations, effectiveness ratings were the lowest-rated criteria (39% rated positive), driving overall development outcomes down. Negative effectiveness ratings were explained largely by poor-quality M&E and poor performance. Poor M&E particularly affected objectives beyond the narrow performance of the Borrower, such as job creation. Shortcomings in the design phase were correlated with negative effectiveness ratings. Among specific objectives, those aimed at reducing greenhouse gas emissions were most likely to achieve positive ratings, and those aimed at improving access to and conditions of finance were least likely to do so.

5.7 Among non-core criteria, 68% of IDB Invest projects achieved positive investment outcomes and 89% demonstrated additionality. Among the operations that failed to have a positive investment outcome, eight loans were repaid, three equities produced lower than expected returns, and one loan defaulted. Three out of four operations with negative additionality ratings did not deliver financial additionality.

5.8 Work quality was less than satisfactory for 71% of validated NSG operations. Shortcomings during appraisal and screening—such as weaknesses in financial analysis; partner identification; and mitigation of foreign exchange, policy, or construction risks—were the most frequent reasons for low work quality ratings.

5.9 Considering these conclusions, OVE recommends:

A. For IDB Management

1. Systematically document in PCRs how E&S safeguard issues were addressed and what their final status was at project closure. PCRs did not consistently report on safeguard issues when safeguards had been triggered. Unresolved E&S issues at project closure were a significant driver of less than satisfactory sustainability ratings. PCRs for all projects that trigger safeguards

need to consistently report on how projects complied with Bank safeguard requirements, indicating whether and how E&S issues were resolved by project closure.

2. Strengthen reporting on cancellations. Given the validations' findings that partial cancellations are a significant driver of less than satisfactory project effectiveness, OVE recommends that PCRs consistently discuss whether and how partial cancellations have affected the results chain and achievement of the operation's objectives. OVE also recommends that IDB report to the Board regularly the amount and reasons for cancelled operations for which no PCRs are required.

3. Strengthen the measurement of project performance. Weaknesses in project M&E continue to be an important reason for less than satisfactory project effectiveness ratings. Management's initiative to strengthen the Development Effectiveness Matrix by requiring the inclusion of at least one valid outcome indicator to assess achievement of each specific objective is in line with the finding that the poor quality of M&E is one of the key contributors to negative effectiveness. Efforts need to be made to strengthen project performance measurement by ensuring that baseline and follow-up data for indicators are routinely collected.

B. For IDB Invest Management

1. Continue strengthening the system to identify projects subject to XSR each year. Management's efforts to strengthen the system to identify projects subject to XSRs has pointed to a significant number of operations for which no XSRs were prepared in the past. To identify projects reaching EOM each year in a timely manner and transparently account for their performance, OVE recommends that IDB Invest:

    i. Build on the progress it has made to date and agree with OVE on the final structure of the database that contains all data necessary to determine when each operation (including those in Special Assets) reaches (or has reached) EOM or has exited the portfolio, as well as each operation's XSR status and reason for XSR waiver where applicable;

    ii. Ensure that the agreed upon database is fully updated and cleaned of any inconsistencies by February 1, 2022, and covers all NSG operations approved since 2015, including operations that are or previously were in Special Assets, and provide OVE permanent access to the database for purposes of verification of projects subject to XSRs each year;

iii. Ensure that the database is updated and cleaned of any inconsistencies annually by February 1, and clearly shows which operations are subject to XSRs in the XSR cycle starting in that year, which operations are not and why, so that OVE can verify the list with a February 1, cut-off date each year.

2. Strengthen the evidence base for measuring project effectiveness. Validations have repeatedly found that the poor quality of M&E is one of the key contributors to negative effectiveness ratings. OVE therefore recommends that IDB Invest identify all outcome indicators that have frequently had insufficient means of verification and thus poor M&E, such as job creation, sector level changes, and other objectives that go beyond direct benefits to the client and ensure that sufficient means of verification for these indicators are in place from the project design stage.

References

IDB (Inter-American Development Bank). (2020). Proposed Adjustments to the Development Effectiveness Framework Tools for Sovereign Guaranteed Operations. (OP-1696-5) Washington, DC: Inter-American Development Bank.

OVE (Office of Evaluation and Oversight). (2018). 2018 XSR Guidelines for IDB Group Private Sector Operations. Washington, DC: Inter-American Development Bank.

———. (2020). OVE's review of PCRs and XSRs: The 2020 validation cycle. Washington, DC: Inter-American Development Bank. Available at https://publications.iadb.org/publications/english/document/OVEs-Review-of-Project-Completion-Reports-PCRs-and-Expanded-Supervision-Reports-XSRs-The-2020-Validation-Cycle.pdf.

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

linkedin.com/showcase/idb-ove

@BID_evaluacion