Office of Evaluation and Oversight

Project Evaluation

IDB Group Project Performance: The 2023 Validation Cycle

IDB

IDB | Invest

Inter-American Development Bank, 2023
Office of Evaluation and Oversight
1350 New York Avenue, N.W.
Washington, D.C. 20577
www.iadb.org/evaluation

RE-591
November 2023

Project Evaluation

IDB Group Project Performance: The 2023 Validation Cycle

Office of Evaluation and Oversight

IDB IDB | Invest

Contents

This report has a new title, but it continues the series of annual reports titled *OVE's Review of Project Completion Reports and Expanded Supervision Reports: The [year] Validation Cycle*. You can find the most recent editions of this report below: 2017 Report, 2018 Report, 2019 Report, 2020 Report, 2021 Report, and 2022 Report.

Acknowledgements

This report was prepared by Oliver Peña-Habib (team leader), Jorge Gallego, Luis Fernando Corrales, Jorge González, Julie King, and Melisa Wong, under the supervision of Ana María Linares, Cluster Leader and Senior Advisor, and the direction of Ivory Yong, OVE Director. Chiaki Yamamoto led the 2023 validation cycle.

This report is the product of the collective efforts of numerous individuals and teams across the Inter-American Development Bank (IDB) Group. The Office of Evaluation and Oversight (OVE) is grateful to all parties involved, including the multiple specialists across the institution who prepared the project self-evaluations. OVE would also like to thank the IDB's Office of Strategic Planning and Development Effectiveness (SPD) and IDB Invest's Development Effectiveness Division (DVF) for overseeing the production of Project Completion Reports (PCRs) and Expanded Supervision Reports (XSRs), respectively. Both SPD and DVF collaborated closely with OVE throughout the 2023 validation cycle.

The following OVE members participated in the 2023 validation cycle, which was led by Chiaki Yamamoto, Lead Specialist, under the supervision of Ana María Linares, Cluster Leader and Senior Advisor, and the direction of Ivory Yong, OVE Director:

Josette Arévalo, Andre Averbug, Fernando Carlos Barbosa, César Bouillon, Michelle Bruno, Eglis Chacón, Luis Fernando Corrales, Ernesto Cuestas, Stefania De Santis, Claudia Figueroa, Pablo Fleiss, Jorge Gallego, Gunnar Fabian Gotz, Monika Huppi, Dana Michael King, Julie King, Regina Legarreta, Julián Loayza, Odette Maciel, Stephany Maqueda, Roland Michelitsch, Rasec Niembro, Hans J. Paris, Lina Pedraza, Oliver Peña-Habib, José Claudio Pires, Juan José Pompilio Sartori, Luisa Riveros, María Fernanda Rodrigo, José Ignacio Sembler, Alejandro D. Soriano, Roni Szwedzki, and Judy Twigg.

Acronyms and Abbreviations

ASR	Annual Supervision Report
AUG	Office of the Executive Auditor
CAN	Country Department Andean Group
CBA	Cost-benefit analysis
CCB	Country Department Caribbean

CEA	Cost-effectiveness analysis
CID	Country Department Central America, Haiti, Mexico, Panama, and Dominican Republic
CO	Corporates
CRF	Corporate Results Framework
CS	Country Strategy
CSC	Country Department Southern Cone
CSD	Climate Change and Sustainable Development Sector
CTOA	Cost and time overrun analysis
DEO	Development Effectiveness Overview
DVF	Development Effectiveness Division
E&S	Environmental and social
EOM	Early operating maturity
FI	Financial institution
IDB	Inter-American Development Bank
IE	Infrastructure and Energy (IDB Invest)
IFD	Institutions for Development Sector
IIC	Inter-American Investment Corporation
INE	Infrastructure and Energy Sector (IDB)
INT	Integration and Trade Sector
INV	Investment operation
M&E	Monitoring and evaluation
MSME	Micro, small, and medium-sized enterprises
NPL	Nonperforming loan
OVE	Office of Evaluation and Oversight
PBL	Policy-based loan
PBP	Programmatic policy-based loan
PCR	Project Completion Report
PMR	Progress Monitoring Report
PROFISCO	Program to Support Fiscal Management and Integration
SAT	Special Assets Division
SCL	Social Sector
SME	Small and medium-sized enterprises
XSR	Expanded Supervision Report

Executive Summary

This report summarizes the 2023 results of the Office of Evaluation and Oversight's (OVE's) annual review of project performance at the Inter-American Development Bank (IDB) Group. Every year, IDB Group Management prepares self-evaluations on the final results of its operations. These are called Project Completion Reports (PCRs) at the IDB and Expanded Supervision Reports (XSRs) at IDB Invest. PCRs and XSRs are then validated by OVE to ensure that they are substantiated by evidence and are prepared in accordance with their respective guidelines. Self-evaluations are intended to serve as accountability and learning tools.

The IDB Group's project evaluation methodology is objectives-based. Project performance is rated based on four core criteria: (i) *relevance* of the project's goals and design, (ii) *effectiveness* or the degree to which the project achieved the specific objectives for which it was approved, (iii) *efficiency* with which the objectives were achieved, and (iv) *sustainability* of the results achieved. Each project then receives an overall outcome rating, calculated as a weighted average of the core criteria ratings, where effectiveness has the highest weight. In addition, two noncore criteria are rated for IDB operations (Bank performance and borrower performance), while three are rated for IDB Invest operations (additionality, investment profitability, and work quality). OVE also rates the quality of the self-evaluation reports by Management. The overall outcome is rated on a six-point scale, while core criteria, noncore criteria, and PCR and XSR quality are rated on a four-point scale.[1] For simplicity, this report groups ratings into "positive" and "negative," where positive corresponds to the top half ratings in the scale, and negative to the bottom half.

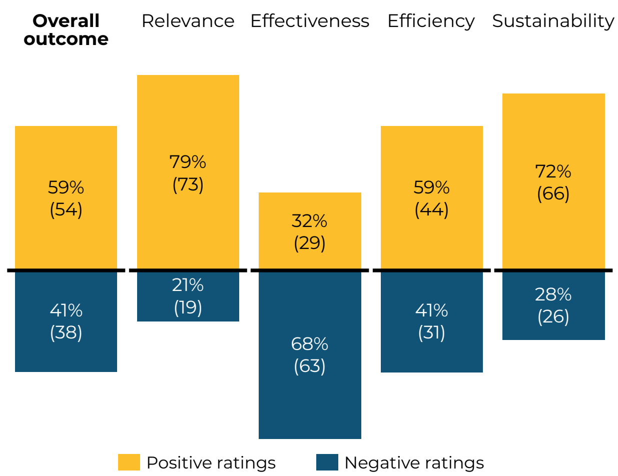
A. IDB performance in 2023

In 2023, OVE validated the PCRs of 92 IDB operations. Of these, 59% achieved a positive overall outcome rating. While this positive rate falls short of the 70% target in the 2020–2023 Corporate Results Framework, it is higher than the rates observed in the previous three years.

1 The overall outcome scale ranges from "highly unsuccessful" to "highly successful." Core and noncore criteria range from "unsatisfactory" to "excellent." PCR and XSR quality ranges from "poor" to "excellent."

Relevance was the highest-rated core criterion, with 79% of operations rated positive; those rated negative were mostly affected by a weak vertical logic. In addition, relevance was affected by changes to projects under implementation that were arguably substantial and fundamental enough to have warranted Board approval, in accordance with the Bank's current regulations.

IDB operations ratings in 2023



Source: OVE.

Effectiveness was the lowest-rated core criterion, rated positive for 32% of operations due to a combination of underachievement and monitoring and evaluation (M&E) issues. Underachievement was far more frequent, mostly due to underdelivery of the outputs that the operations were expected to produce, which then led to not achieving most project objectives. According to the PCRs, the most common reason for underachievement was institutional changes or problems on the borrower's side, such as shifts in government priorities, low capacity among executing units, and organizational and regulatory changes that affected project performance. However, many PCRs did not sufficiently explain why the operations underachieved. In turn, the most common M&E issue was attribution problems, as there was insufficient evidence that the observed outcomes could be plausibly attributed to the operations.

Efficiency was rated positive for 59% of operations, with most negatively rated operations affected by the poor quality of the economic analyses presented to assess the use of project resources.

Sustainability was rated positive for 72% of operations. Operations rated negative were affected by an uncertain continuity of results and/or an unsatisfactory performance in environmental and social (E&S) safeguards.

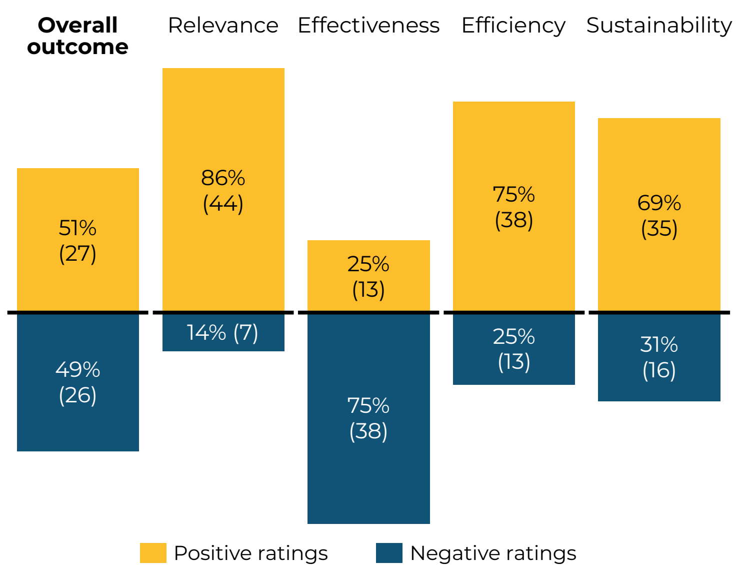
All in all, of the 92 operations, 54 achieved a positive overall outcome rating, of which 19 had positive ratings in all core criteria. This corresponds to 10 operations from the Infrastructure and Energy Sector, 6 from the Social Sector, and 3 from the Institutions for Development Sector. They were implemented in 13 of the 26 borrowing member countries. These operations have the potential to generate learning for future operations.

Bank performance was rated positive for 59% of operations; most shortcomings were identified during project preparation. Borrower performance was rated positive for 72% of operations; most negatively rated operations were affected by the capacity of the executing unit.

B. IDB Invest performance in 2023

In 2023, OVE validated the XSRs of 51 IDB Invest operations. Including two additional Special Assets (SAT) operations for which no XSRs were prepared (and were thus rated negative), 51% of operations achieved a positive overall outcome rating. This falls short of the 65% target in the 2020–2023 Corporate Results Framework.

IDB Invest operations ratings in 2023



Source: OVE.

Relevance was the highest-rated core criterion, with 86% of operations rated positive. Those rated negative suffered from a weak vertical logic at project approval, in combination with other issues.

Effectiveness ratings were positive for 25% of operations. These results are explained by a combination of underachievement and M&E issues, though underachievement was more common. According to the XSRs, the most frequent reason behind underachievement was external shocks (such as changes in market conditions, macroeconomic shocks, sociopolitical unrest, and climate events), followed by the COVID-19 fallout and changes in client priorities. In one-third of the operations with financial institutions (FIs) reviewed in this cycle (8), changes in client priorities led to a shift away from the target segments. In most cases, IDB Invest was able to detect this shift during supervision, but OVE found no evidence of follow-up measures. Moreover, 5 FI operations with negative effectiveness ratings were with clients that had prior operations also rated negative at the time. IDB Invest has recently prepared an action plan to improve FI performance that does not include specific measures to deal with cases where shifts in targets segments are identified. Effectiveness ratings were also affected by M&E issues like inaccuracies in outcome indicator targets and missing data, but IDB Invest's efforts helped reduce the magnitude of the problem.

Efficiency was rated positive for 75% of operations. Most negatively rated operations were affected by low financial or economic returns and subpar portfolio growth among FI operations.

Sustainability was rated positive for 69% of operations. Operations rated negative were affected by an uncertain continuity of results and/or an unsatisfactory E&S performance.

All in all, of the 51 operations, 27 achieved a positive overall outcome rating, of which 12 had positive ratings in all core criteria. This corresponds to 7 operations from the Corporates business segment, 4 from Infrastructure and Energy, and 1 from Financial Institutions. They were implemented in 9 of the 26 borrowing member countries. These operations have the potential to generate learning for future operations.

Additionality was rated positive for 90% of operations. Investment profitability was rated positive for 84% of operations; negatively rated operations did not meet gross profit expectations due to prepayments or cancellations. Work quality was the lowest-rated noncore criterion, positive for 41% of operations, with most shortcomings identified during project preparation.

C. IDB Invest operations with Mini-XSRs

The report also presents the results of an additional 52 IDB Invest operations for which no XSRs were prepared when due. Identifying when operations are ready for XSR preparation entails several challenges. To address this issue, IDB Invest created an automated system and found that 52 repaid or prepaid operations had not prepared XSRs when due. As an exceptional measure, IDB Invest agreed with OVE on a simplified methodology to account for these operations—the "Mini-XSR." Mini-XSRs were prepared in two batches during 2022–2023. This report presents the results of all 52 operations. No additional Mini-XSRs are expected in the future. The rating scale in Mini-XSRs is "satisfactory," "unsatisfactory," or "lack of data."

Of the 52 operations, overall outcome was "satisfactory" for 50%. Among core criteria, relevance was "satisfactory" for 90%, effectiveness for 50%, efficiency for 48%, and sustainability for 65%. Effectiveness, efficiency, and sustainability had similar shares of operations rated "unsatisfactory" (ranging 25%–29%), while "lack of data" ratings were highest in effectiveness (25%) and efficiency (23%). The factors affecting operations with "unsatisfactory" ratings were largely similar to those found for negatively rated operations with full XSRs.

Work quality was the lowest-rated criterion with 42% "satisfactory," while investment profitability was 71% "satisfactory."

D. Performance over time

Based on the 2018–2023 period, trends over time for the IDB show that the positive rate in overall outcome ratings rose to 59% in 2023, after oscillating around 52% over the past three years. Relevance slightly reversed its downward trend, while effectiveness remained the lowest-performing criterion. Efficiency has stagnated after an upward trend, while sustainability increased in the past two years.

At IDB, performance is heterogeneous across sectors and regions, though not so across lending instruments. Regression analyses were conducted to test for statistically significant differences in performance, finding that the Social Sector has performed strongest while the Climate Change and Sustainable Development Sector weakest. Southern Cone outperformed the other regions, while there is no significant difference between the performance of investment and policy-based operations.

Trends over time for IDB Invest show that the positive rate in overall outcome ratings dropped to 53% in 2023 (excluding SAT operations without XSRs),[2] after oscillating around 60% over the past three years. Relevance remained nearly unchanged in 2023 after an upward trend since 2019. Effectiveness remained lowest, with a zigzagging trend that fell to its lowest value yet. By contrast, efficiency and sustainability have continued their upward trend since 2019.

2 For the trends over time analysis, data excludes SAT operations without XSRs (which since 2022 are given a negative overall outcome rating) to avoid introducing comparability issues with previous years.

At IDB Invest, performance is heterogeneous across business segments and regions, and somewhat heterogeneous between pre– and post–Merge Out operations.[3] Regression analyses found that the Infrastructure and Energy segment scored highest, while Financial Institutions lowest. The Caribbean was outperformed by the other regions, although this is based on a very small number of operations. Operations approved after the Merge Out scored higher than those approved before, but only in some of the tests performed.

E. Quality of self-evaluations

About half of the PCRs were of satisfactory quality. The share of PCRs rated positive was 49%, similar to last year's 48%. As in the 2022 validation cycle, most shortcomings were related to the quality and completeness of the analyses in the effectiveness, efficiency, and relevance sections of the PCRs. OVE's 2022 report recommended to strengthen the quality of PCRs, and Management prepared an action plan that included four main actions deemed partially relevant by OVE because of shortcomings in the definition of the specific activities. OVE underlines the importance of addressing these shortcomings and promptly completing and implementing the actions. As for differences between Management's self-assigned ratings and OVE's ratings, while these remain large, they narrowed compared to recent years for all core criteria but relevance.

For the first time, all XSRs were of satisfactory quality. XSRs in this cycle closely followed the guidelines, cited sufficient data, and were balanced and complete, among other aspects. Quality of XSRs has remained over 80% positive since 2019, and even though a downward trend was observed in 2021–2022, the 100% positive in 2023 showed an important recovery.

Considering these findings, OVE recommends the following.

For IDB Management:

1. **Ensure: (i) the overall quality of PCRs, including the quality of their economic analyses, by strengthening quality assurance instruments and their application; (ii) that PCRs systematically report on the reasons for underachievement; and (iii) that operational staff receive adequate training, working in collaboration with OVE.** As in 2022, OVE's report for the 2023 validation cycle found issues in the quality of PCRs, including the lack of sufficient explanation of why objectives were not achieved in the effectiveness section and the low quality of economic analyses in the efficiency section. Management should address

3 The Merge Out was a process that consolidated the private sector lending activities of the IDB and the Inter-American Investment Corporation (IIC) into IDB Invest. It took effect on January 1, 2016.

these issues as part of the action plan prepared in response to OVE's 2022 recommendation on the matter, incorporating OVE's feedback on the plan's relevance.

For IDB Invest Management:

1. **Define measures in the action plan for FI operations for when project supervision finds that the client has shifted away from the target segment and ensure their systematic implementation.** Based on the finding that, in most cases, IDB Invest detected when clients shifted away from the target segment during supervision, but there is no evidence that measures were taken to address such shifts, Management should define measures for dealing with these cases in its action plan for FIs and ensure that they are implemented systematically.

01

Introduction

1.1 This report summarizes the results of the Office of Evaluation and Oversight's (OVE's) annual review of project performance at the Inter-American Development Bank (IDB) Group. Every year, IDB Group Management prepares self-evaluations on the final results of the projects it finances, which are then reviewed and validated by OVE. The self-evaluations are called Project Completion Reports (PCRs) at the IDB and Expanded Supervision Reports (XSRs) at IDB Invest.[1] Self-evaluations are intended to serve as accountability and learning tools. The accountability goal addresses the need for the IDB Group to ensure that project resources were used to achieve the objectives for which they were approved, with due attention to effectiveness, efficiency, and sustainability. The learning goal aims to replicate successes and avoid mistakes in the future, by providing lessons to inform the design of new projects and the implementation of both ongoing and new projects.[2]

A. Evaluation methodology and the role of OVE

1.2 The IDB Group's project evaluation methodology is objectives-based. Project performance is measured against four core criteria:

(i) *relevance* of the project's goals and design (vertical logic);

(ii) *effectiveness* or the degree to which the project achieved the specific objectives for which it was approved;

(iii) *efficiency* with which the specific objectives were achieved;

(iv) *sustainability* of the results achieved through the mitigation of risks and compliance with environmental and social safeguards.

1.3 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, which is calculated with a weighted average of the core criteria ratings.[3] The overall outcome rating uses a six-point scale ranging from "highly unsuccessful" to "highly successful." For all scales, the bottom half ratings are considered negative ratings, while the top half are considered positive ratings. In addition, self-evaluations include a set of noncore criteria that are rated on the same four-point scale as the core criteria but that do

1 OVE does not review self-evaluations of IDB Lab operations.

2 See 2020 PCR Guidelines (document OP-1696-6, Annex 1) and 2018 XSR Guidelines.

3 Relevance, efficiency, and sustainability weigh 20% each, while effectiveness weighs 40%. For IDB's policy-based loans (PBLs), efficiency is not rated, and effectiveness weighs 60% instead.

not count toward the project's overall outcome rating. For PCRs, the noncore criteria are Bank performance and borrower performance. Noncore criteria for XSRs include additionality, investment profitability, and the work quality of IDB Invest. OVE also rates the quality of the self-evaluation reports by Management. Annex I summarizes the IDB Group's project evaluation framework, and the full methodology is described in the PCR Guidelines and XSR Guidelines.

1.4 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 IDB and IDB Invest. At the IDB, PCRs are prepared after operations have finished disbursing and have closed (operational closure). At IDB Invest, XSRs are prepared after operations have reached "early operating maturity" (EOM), indicating that the operation is sufficiently mature, has had enough time to achieve results, and is ready for evaluation.[4] Management prepares PCRs and XSRs and assigns ratings to each core and noncore criterion, based on the PCR and XSR guidelines. OVE then reviews Management's self-evaluations to ensure that they are substantiated by evidence and are prepared in accordance with the guidelines. OVE assigns its own ratings, which are considered final and are used for corporate reporting in the *Development Effectiveness Overview* (DEO).[5]

B. Scope of this report

1.5 This report presents the results of OVE's validation of the PCRs and XSRs for 143 operations in the 2023 validation cycle. OVE validated the PCRs of 92 IDB operations and the XSRs of 51 IDB Invest operations. Annex II lists these operations and their associated ratings. Annex III provides OVE's validation notes for IDB operations only, as IDB Invest validation notes are confidential. Most PCRs and XSRs assessed a single project, although some were assessed as a program covering a group of related projects with shared objectives.[6] Throughout this report, the more general term *operation* is used to refer to both validated projects and programs.[7]

4 Factors determining EOM vary by project type (see details in XSR Guidelines, Annex II).

5 The DEO is an annual report by the IDB Group that shares the results of its work in the region.

6 See paragraph 1.12 of the PCR Guidelines for a definition of when a series of operations are to be evaluated in a single PCR.

7 The 92 IDB operations covered 63 projects and 29 programs (109 projects in total). The 51 IDB Invest operations covered 47 projects and 4 groups of projects (56 projects in total).

1.6 This report also presents the results of the validation of 52 IDB Invest operations for which self-evaluations had not been prepared when due. In 2020–2021, IDB Invest implemented a system to calculate EOM and flag operations ready for XSR preparation. In the process, it identified 52 operations that had reached EOM between 2015 and 2019 but for which no XSRs had been prepared. These operations had been prepaid or repaid. IDB Invest and OVE agreed on a simplified methodology to evaluate these operations, called the "Mini-XSR". In 2022, IDB Invest presented Mini-XSRs for 24 of these operations and OVE validated them.[8] In 2023, IDB Invest presented Mini-XSRs for the remaining 28. This report presents the results of all 52 operations for which Mini-XSRs were prepared. Annex II lists these operations and their ratings.

1.7 This report is structured as follows. Section II presents the 2023 project performance results for the IDB Group, analyzing the factors that influenced performance, among other aspects. Section III compares the 2023 results with previous years to identify trends over time. Section IV addresses the quality of PCRs and XSRs, followed by conclusions and recommendations.

8 The validated results of the 24 Mini-XSRs can be found in OVE's 2022 report (document RE-575).

02

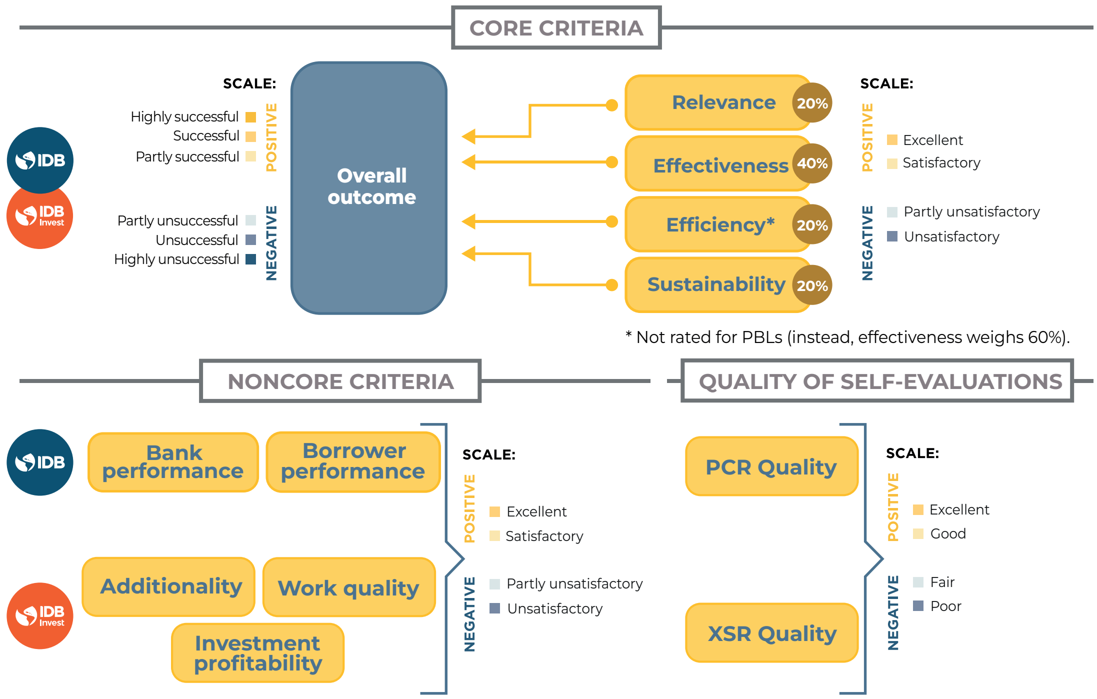
Performance in 2023

2.1 This section presents the results of the 2023 validation cycle for IDB and IDB Invest operations. It presents ratings for overall outcome, core criteria, and noncore criteria, and includes analyses on the main factors explaining project performance.[9] For simplicity, the report groups ratings into positive and negative, where positive ratings are those in the top half of the respective rating scale, while negative ratings are those in the bottom half. Figure 2.1 provides an overview of the evaluation criteria discussed below and their respective rating scales.

Figure 2.1

Evaluation criteria and rating scales

Source: OVE, based on PCR and XSR Guidelines.



A. IDB performance in 2023

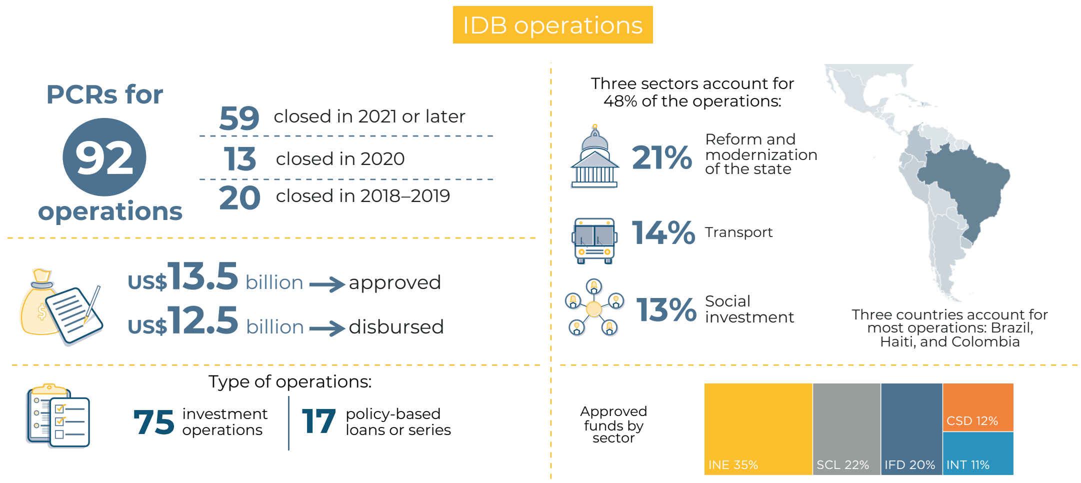
2.2 OVE validated the PCRs of 92 operations. Figure 2.2 shows key characteristics of the 92 IDB operations for which Management submitted PCRs for validation by OVE in the 2023 cycle. This has been the largest number of operations validated in any year to date.

9 Such factors were derived using content analysis with computer-assisted qualitative data analysis software (CAQDAS). This method was applied to determine the presence of certain topics within OVE's validation notes, the PCRs, and the XSRs, where the frequency of such topics is considered indicative of the main drivers of performance.

Figure 2.2

Characteristics of IDB operations with PCRs validated in 2023

Source: OVE, with data from IDB (2023a).



Note: Brazil had 24 operations validated in 2023, of which 14 were investment loans within the Program to Support Fiscal Management and Integration (PROFISCO), a credit line for state governments (BR-X1005).

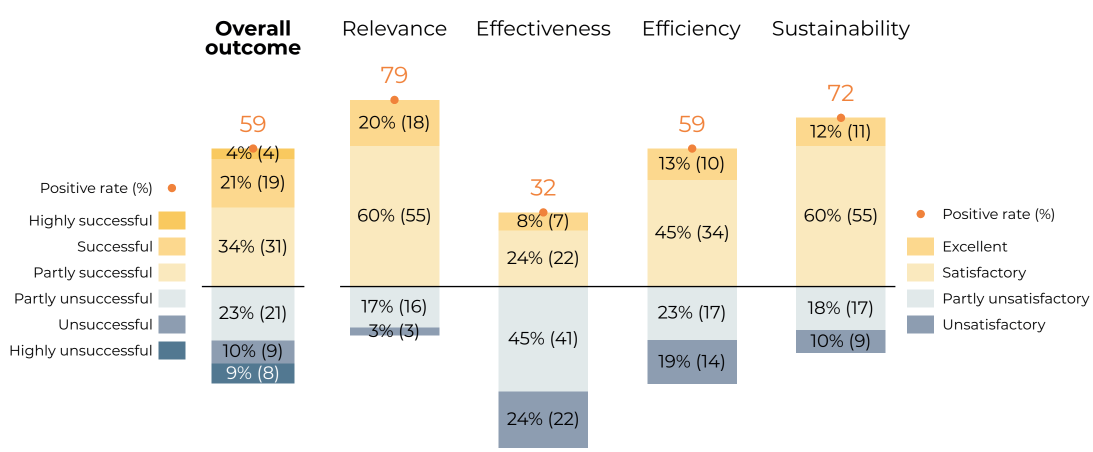
1. Overall outcome ratings

2.3 Of 92 operations, 54 (59%) achieved a positive overall outcome rating. Even though this positive rate falls short of the 70% target set in the 2020–2023 Corporate Results Framework (CRF), it is higher than the rates observed in the previous three years (see Section III). Most operations (80 operations or 87%) fell in the middle four ratings ("unsuccessful," "partly unsuccessful," "partly successful," and "successful"), while 4 (4%) fell in the highest possible rating of "highly successful" and 8 (9%) in the lowest possible rating of "highly unsuccessful" (Figure 2.3).

Figure 2.3

IDB operations ratings, 2023

Source: OVE.



Note: N = 92 operations rated for all criteria except for efficiency (N = 75), given that policy-based loans (PBLs) are not rated for efficiency. Percentages may not add up to 100 due to rounding.

2. Core criteria ratings

2.4 Relevance was the highest-rated core criterion, with 73 of 92 operations (79%) rated positive. These 73 operations had objectives that were relevant to address country development needs, were aligned with government plans and priorities and with IDB Group Country Strategies (CSs), and had adequate project designs that considered country realities. The remaining 19 operations were rated negative in relevance due to a combination of three factors (which are not mutually exclusive):[10]

- In 15 operations (79% of those rated negative), project design had a weak vertical logic.[11] Among them, four operations presented vertical logical issues from project approval. Another seven experienced changes during implementation that weakened the original vertical logic, such as cancellations, modifications, and the truncation of programmatic policy-based loan (PBP) series. Some of these changes were of such nature that they could be considered substantial and fundamental (Box 2.1). The remaining four had a combination of vertical logic issues both at approval and during implementation.
- In 14 operations (74%), project design did not fully consider country realities. This included implementation arrangements not aligned with the executing unit's capacity and insufficient analysis of contextual factors affecting project implementation.
- In 8 operations (42%), project objectives were not aligned with the CSs in place at the time of project approval, during implementation and/or at project closure. Of these, 2 were not aligned with the CS that was valid at the time of project approval, while 3 were not aligned with the CS in place during project implementation or at project closure, and 3 were not aligned with either.

Box 2.1. Projects that experienced major changes during implementation

According to IDB's OA-420 and OA-430, approved operations undergoing changes require Board approval if the proposed change substantially and fundamentally modifies the structure, cost, objectives, or beneficiaries of the project or other implementation agreements, including major exceptions to the Bank's procurement policies; or if in Management's judgement, the proposed change is substantial and fundamental under the circumstances presented in a particular case.

In 2022, the Office of the Executive Auditor (AUG) reviewed the process for reformulating approved operations at the IDB and examined 39 operations

10 Of the 19 operations, 8 had a combination of two factors, 6 only had one factor, and 5 had all three.

11 Vertical logic refers to how a project's development objectives are to be achieved through a results chain that links activities to outputs and outputs to outcomes.

that had been reformulated or modified with Board approval between 2018 and 2021 (AUG, 2022). The scope of AUG's review did not include examining non-reformulated operations that underwent major changes.

In the 2023 validation cycle, out of the 92 validated operations, OVE identified 4 with major changes that could be considered substantial and fundamental but were not submitted for Board approval:

- **HA-L1048 (US$30 million).**[a] At approval in 2010, the project's objective was *to improve the quality of life of the low-income families affected by the 2010 earthquake through the provision of housing solutions that satisfy basic needs of shelter and security* (document PR-3539). During implementation, the operation suffered significant changes that included the reallocation of about half of its approved amount to build a hospital, provide budget support to the executing unit, pay for polio and diphtheria vaccination campaigns, and conduct studies. These modifications took place through multiple project extensions spanning over seven years.

- **HA-X1002 (US$3.4 million)[b] and HA-G1023 (US$9 million).**[c] HA-X1002 was approved in 2009 (document AT-1468) as a single, stand-alone operation financed by the Global Environment Facility (GEF) and was never formally reformulated with Board approval. Four years later, in 2013, HA-G1023 was approved also as a stand-alone operation (document PR-4029). While the second operation (HA-G1023) was designed to complement the first (HA-X1002), it had its own separate results matrix, different from the results matrix of the first operation. However, the second operation ended up completely modifying the results matrix of the first one as the Bank sought to execute and monitor both operations as one, significantly changing the first operation.

- **AR-L1084 (US$200 million).**[d] The specific objectives of this Multiple Works Program approved in 2010 included *improving and expanding the water service in the central and western parts of the city of Salta* (document PR-3557). Plans for Salta were abandoned during implementation, and Salta was replaced by the city of Jujuy. Even though the Multiple Works Program modality allows for flexibility in the final selection of investments to be financed (see document PR-202), in this case Salta was explicitly stated in the specific objectives of the project. As a result, replacing Salta for another city was a change in the project's specific objectives.

Note: Project names: [a] Support to the Shelter Sector Response Plan. [b] Sustainable Land Management of the Upper Watersheds of Southwestern Haiti. [c] Sustainable Management Upper Watersheds Southwestern Haiti-Macaya National Park. [d] Water and Sanitation Program for Urban and Suburban Centers.

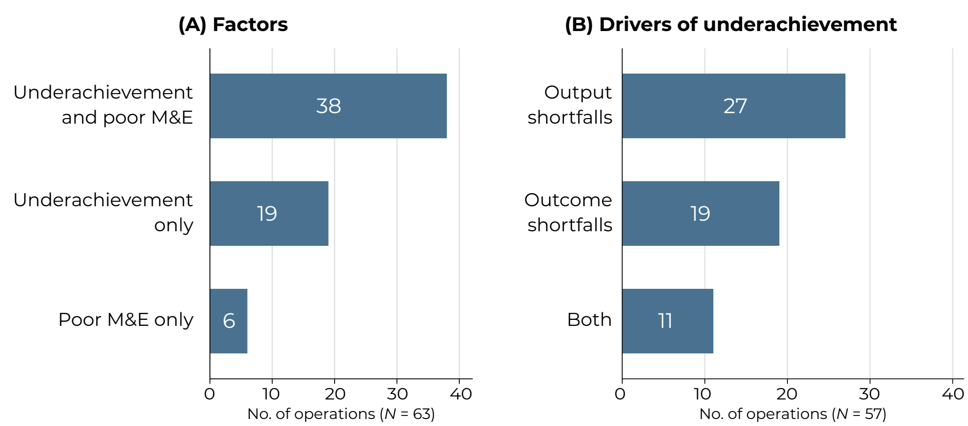
2.5 Effectiveness was rated positive for 29 of 92 operations (32%), making it the lowest-rated core criterion. Negative effectiveness ratings were due to a combination of two factors: underachievement and poor monitoring and evaluation (M&E). Underachievement means there is evidence that the operation did not achieve most of its specific objectives. Underachievement can be further broken down by its key determinants into *output shortfalls* and *outcome shortfalls*. Output shortfalls occur when the project does not deliver the planned outputs (or not to their full extent) and therefore the expected outcomes are not achieved. In turn, outcome shortfalls occur when the project does deliver the planned outputs, but the expected outcomes are still not achieved. While in both cases the expected outcomes are not achieved, knowing whether the determinant was the outputs or the outcomes helps identify where the chain broke along the operation's vertical logic. The second factor, poor M&E, means that there is insufficient evidence to demonstrate achievement of objectives due to measurement issues.

2.6 Most operations were affected by a combination of underachievement and poor M&E, though underachievement was far more frequent. As Figure 2.4 shows, of the 63 operations with negative effectiveness ratings, 38 (60%) suffered from the combination of both, while 19 (30%) were affected by underachievement, and 6 (10%) were affected by M&E problems. A breakdown of the 57 operations facing underachievement shows that most (27 or 47%) experienced exclusively output shortfalls. Outcome shortfalls came in second place (19 or 33%), and a smaller share (11 or 19%) faced both output and outcome shortfalls. Partial cancellations were associated with lower effectiveness, though to a lesser extent than in previous validation cycles.

Figure 2.4

Factors behind negative effectiveness ratings, IDB operations

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



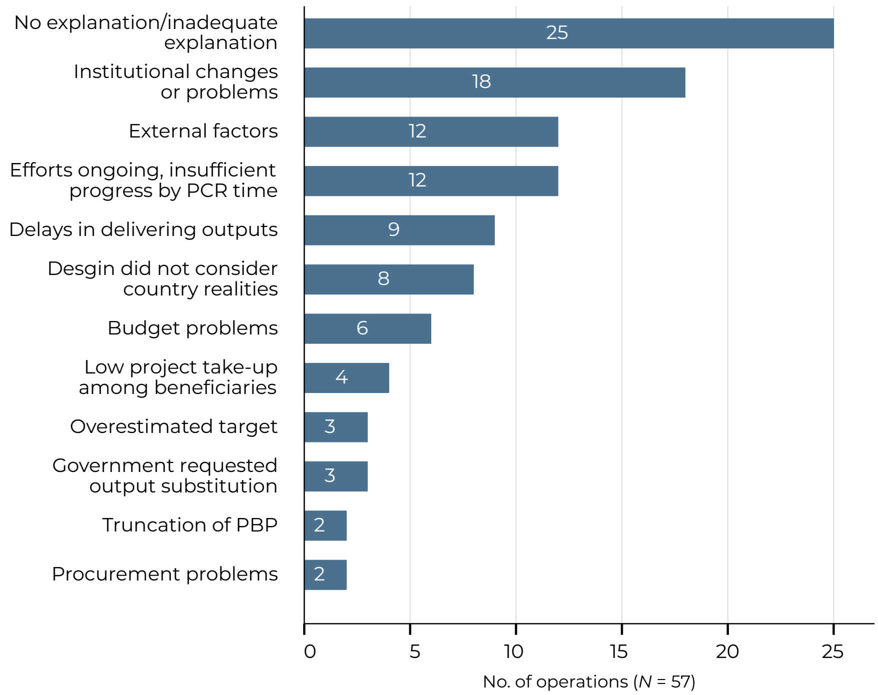
2.7 Multiple factors contributed to underachievement, though many PCRs did not sufficiently explain why objectives were not achieved. Figure 2.5 summarizes the reasons for underachievement according to PCRs; these are largely the same as those found last year. The most frequently cited was institutional changes on the borrower's side, such as: (i) a shift in priorities or scope away from the initial project design, often leading to the cancellation of planned outputs or their delivery at lower-than-expected levels; (ii) low capacity of executing units or other stakeholders involved in project execution; and (iii) organizational changes in the government and regulatory changes that affected project performance.

The second most frequently cited reason was external factors, such as the COVID-19 pandemic, the macroeconomic and financial context, and natural disasters. Several PCRs indicated that, at project closure, it was too soon to observe outcomes, as delivery was still underway—even though these projects spent, on average, 7.4 years to implement. Compared with last year, a higher proportion of projects cited delays in the delivery of outputs and issues with timely availability of counterpart funds. However, nearly half of the PCRs for operations facing underachievement failed to provide sufficient or adequate explanations for their lack of results, similar to last year, missing a learning opportunity.

Figure 2.5

Reasons behind underachievement according to PCRs

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



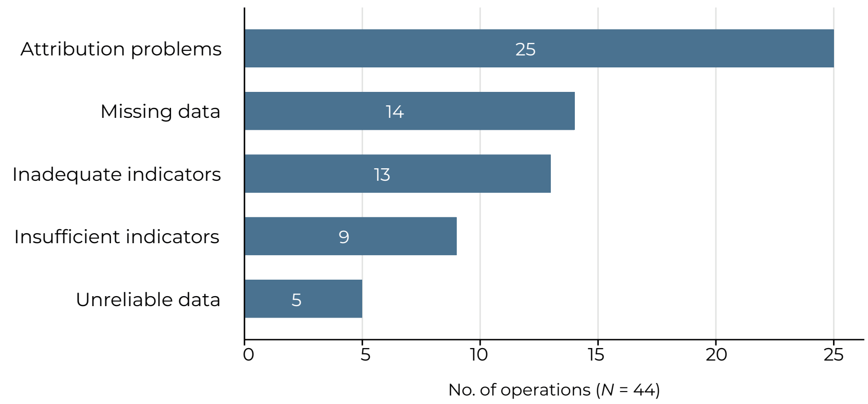
Note: Numbers do not add up to 57 because reasons are not mutually exclusive.

2.8 Effectiveness ratings were also affected by several types of M&E issues. The most common type this year was attribution problems (25 of the 44 operations rated negative because of M&E issues), as there was insufficient evidence that the observed outcomes could be plausibly attributed to the operations (Figure 2.6). Missing data was also common, due to challenges in data collection or outright lack of tracking with no justification provided. Some operations had indicators that were inadequate measures of the objectives they intended to achieve; others did have at least some indicators that were adequate but insufficient as a whole to measure key dimensions of the project objectives. Only a few operations were affected by unreliable data (5 of the 44), a significant improvement compared with last year, when unreliable data was the most common M&E issue.

Figure 2.6

Types of M&E issues, IDB operations

Source: OVE.



Note: Numbers do not add up to 44 because the M&E issues are not mutually exclusive.

2.9 Of 75 investment operations, 44 (59%) achieved a positive efficiency rating.[12] PCRs are expected to demonstrate how efficiently the project has used its resources through an economic analysis—either a cost-benefit analysis (CBA) or a cost-effectiveness analysis (CEA). In exceptional circumstances where such an analysis is not feasible, the PCR needs to justify why. When the analysis is not presented and there is no reasonable justification, a cost and time overrun analysis (CTOA) is used, which is based on the performance classification from the Progress Monitoring Reports (PMRs). In addition, if OVE finds that the CBA or CEA analysis lacks credibility and thus cannot be relied upon to rate efficiency, the rating also reverts to the CTOA. In neither of these two cases can the project achieve a positive efficiency rating.[13] Of the 31 operations with negative efficiency ratings, 19 presented an economic analysis that lacked credibility, so a CTOA had to be used instead (with PMR performance leading to "partly unsatisfactory" in 11 cases and to "unsatisfactory" in 8 cases). In addition, 6 operations did not present an economic analysis when such analysis was feasible, so the rating also reverted to a CTOA (with PMR performance leading to "partly unsatisfactory" in 3 cases and "unsatisfactory" in the other 3). Moreover, 5 operations presented credible analyses, but these found that resources had not been used efficiently. For the remaining operation, an economic analysis was not feasible, thus a CTOA was used (with PMR performance leading to an "unsatisfactory" rating). These results are similar to those in OVE's 2022 report, which also found that the low quality of economic analyses was the main reason behind negative ratings.

12 Efficiency is not rated for PBLs per PCR Guidelines.

13 If there is no reasonable justification for a missing CBA or CEA, and cost and time overrun information from PMRs 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." Same rules apply when a CBA or CEA is rejected for lack of credibility.

2.10 Of 92 operations, 66 (72%) achieved positive sustainability ratings. For these, risks to the continuity of results achieved were minor or nonexistent, and environmental and social (E&S) safeguards performance was satisfactory. Of the 26 operations rated negative in sustainability, 13 were affected by a combination of issues related to continuity of results and unsatisfactory E&S performance, while 12 were only affected by issues concerning continuity of results and 1 only by E&S issues. In terms of continuity of results, operations faced important risks that had not been mitigated and whose impact was likely to be significant, such as weaknesses in country capacities, questionable financial sustainability, uncertain economic or political environment, and low commitment from key stakeholders. In seven operations, there was insufficient information on whether risks to continuity of results had been adequately mitigated. In turn, operations affected by E&S performance had important issues that remained unsolved at project closure, including the mitigation of environmental impacts, social conflict situations, as well as resettlement processes and compensations for affected property. In three operations, there was insufficient information on whether or how E&S issues had been adequately resolved by project closure.

2.11 All in all, among the 92 operations, 54 achieved a positive overall outcome rating, of which 19 had positive ratings in all four core criteria. These 19 operations have a potential to generate learning for future operations. Figure 2.7 shows one bullet point for each of these 19 operations, summarizing the objectives that they achieved (among all project objectives defined at approval) and the countries where they were implemented. Annex IV provides further details on these operations and their achieved objectives.

Figure 2.7

IDB operations with positive ratings in all core criteria

Source: OVE, based on PCR validations (see details in Annex IV).

IDB operations

Infrastructure and Energy Sector

- Improved electricity service (AR-L1079)
- Strengthened water and sanitation policy (BO-L1200)
- Improved housing conditions (BR-L1297)
- Improved road safety and efficiency (BR-L1373)
- Supported the National Logistics Policy (CO-L1109)
- Improved potable water service (DR-L1057)
- Enhanced energy transmission (EC-L1117)
- Restored energy and road infrastructure (HA-L1130)
- Improved electricity service (SU-L1009)
- Increased road infrastructure financing and efficiency (UR-L1107)

Social Sector

- Supported income during COVID-19 (BL-L1034)
- Supported income and employment during COVID-19 (BR-L1554)
- Strengthened labor market sector (CH-L1135)
- Improved social services for people with disabilities (CO-L1252)
- Enhanced social protection services (HO-L1204)
- Improved Ministry of Labor programs (PR-L1066)

Institutions for Development Sector

- Enhanced fiscal management in Mato Grosso (BR-L1234)
- Enhanced fiscal management in the Federal District (BR-L1250)
- Supported financing for COVID-19 expenses (EC-L1272)

Note: AR: Argentina. BL: Belize. BO: Bolivia. BR: Brazil. CH: Chile. CO: Colombia. DR: Dominican Republic. EC: Ecuador. HA: Haiti. HO: Honduras. PR: Paraguay. SU: Suriname. UR: Uruguay.

3. Noncore criteria ratings

2.12 Bank performance was rated positive for 54 of 92 operations (59%); most were affected by shortcomings during project preparation. This noncore criterion rates the quality of IDB's work during project preparation and supervision. In 20 of the 38 operations rated negative (53%), IDB performance had shortcomings only during preparation, while 12 (32%) presented shortcomings during both preparation and supervision, and 6 (16%) only during supervision.

- **Preparation.** Among the 32 operations with preparation shortcomings, the most frequent was that the IDB did not fully consider country realities when designing the project (24 operations) as a result of limitations in needs assessments, diagnostic studies, and/or other ex-ante analyses, including time and cost calculations. In addition, there were weaknesses in the operations' M&E frameworks (16) due to a selection of indicators that were inadequate to measure the expected outcomes or insufficient to fully capture the objectives. Other preparation shortcomings were related to projects with a weak vertical logic (4). Finally, for several operations, risk identification and design of mitigation measures were insufficient (13), including on potential E&S risks (4).

- **Supervision.** Among the 18 operations with supervision shortcomings, the most common were deficiencies in project monitoring (9), such as failure to reflect important project changes in the M&E framework (e.g., key changes in outputs), together with lack of training of executing units on how to comply with the Bank's M&E reporting requirements. An equally frequent shortcoming concerned challenges that arose during implementation that were not addressed in a timely manner, including risks that materialized but were not mitigated (9). In some cases (4), there were changes to projects under implementation that were arguably substantial and fundamental enough to have warranted Board approval. In other cases (3), project changes affected the collection of key data, introduced outputs outside the scope of the operation, or were incorporated even though their potential contribution to the achievement of objectives was not clear. Finally, poor E&S supervision also affected the ratings of a few operations (4).

2.13 Borrower performance was rated positive for 66 of 92 operations (72%). Among the 26 operations rated negative, most were affected by the capacity of the executing unit (21 operations) due to high turnover and lack of appropriate skills. Administrative challenges also affected the operations (19), as manifested by long delays in delivering the outputs and coordination problems with other stakeholders. There were also challenges in complying with contractual agreements (18), especially on M&E responsibilities, fiduciary matters (including integrity), and E&S requirements.

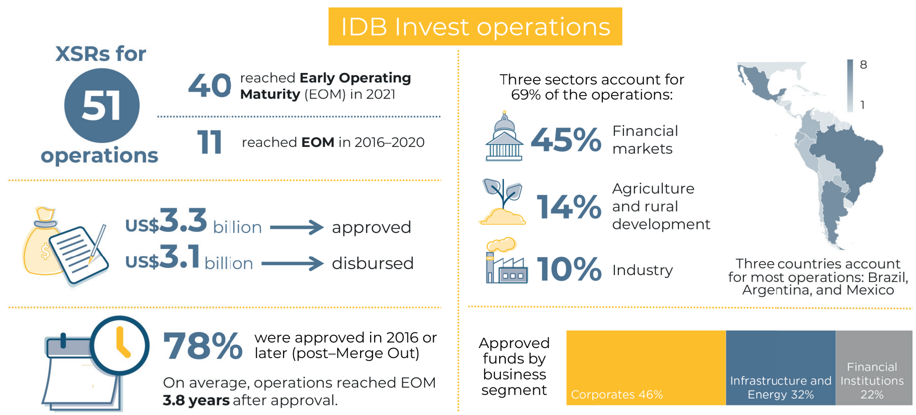
B. IDB Invest performance in 2023

2.14 OVE validated the XSRs of 51 operations. Figure 2.8 shows key characteristics of the 51 IDB Invest operations for which Management submitted XSRs for validation by OVE in the 2023 cycle.

Figure 2.8

Characteristics of IDB Invest operations with XSRs validated in 2023

Source: OVE, with data from IDB Invest (2016, 2023a, 2023b, 2023c, 2023d).



Note: Approval and disbursement data up to December 31, 2022. Operations in this validation cycle were expected to have reached EOM in 2021. For those that did so before, their EOM dates were not detected on time or they were in Special Assets.

2.15 This report also includes four operations that exited the portfolio from Special Assets. The Special Assets Division (SAT) at IDB Invest manages the financially impaired portfolio, seeking to maximize recoveries and preserve capital through recovery strategies. In the past, these operations had not been systematically accounted for, but since last year they have started to be reported on for accountability purposes, following OVE recommendations.[14] Due to their SAT status, these operations may face increased evaluability challenges. Since 2023, Management assesses whether each SAT operation is evaluable and decides if an XSR can be prepared. OVE then reports SAT operations as follows: (i) if an XSR is prepared, it is validated by OVE and the operation is included in the annual cohort of validated operations; (ii) if no XSR is prepared, the operation is assigned a negative overall outcome rating. In this validation cycle, out of four SAT operations that exited the portfolio,[15] two had XSRs prepared and validated, thus

14 OVE's 2022 report (document RE-575) included four operations that exited the portfolio from SAT with no XSRs prepared; as a result, they were assigned a negative overall outcome rating.

15 They exited the portfolio in 2021, in line with the EOM year of most operations in the 2023 cohort.

they have already been included in the cohort of 51 validated operations. The other two operations were considered not evaluable by IDB Invest, so no XSRs or validations were prepared, and OVE assigned them a negative overall outcome rating.[16]

1. Overall outcome ratings

2.16 Of 51 validated operations plus the 2 SAT operations without XSRs, 27 (51%) achieved a positive overall outcome rating. This is the lowest positive rate observed since 2019, a setback after several years of modest improvements (see Section III). Moreover, the 51% positive rate falls short of the target of 65% set in the 2020–2023 CRF. Most operations (47) fell in the middle four ratings ("unsuccessful," "partly unsuccessful," "partly successful," and "successful"), while 3 fell in the highest possible rating of "highly successful" and 1 in the lowest rating of "highly unsuccessful" (Figure 2.9).

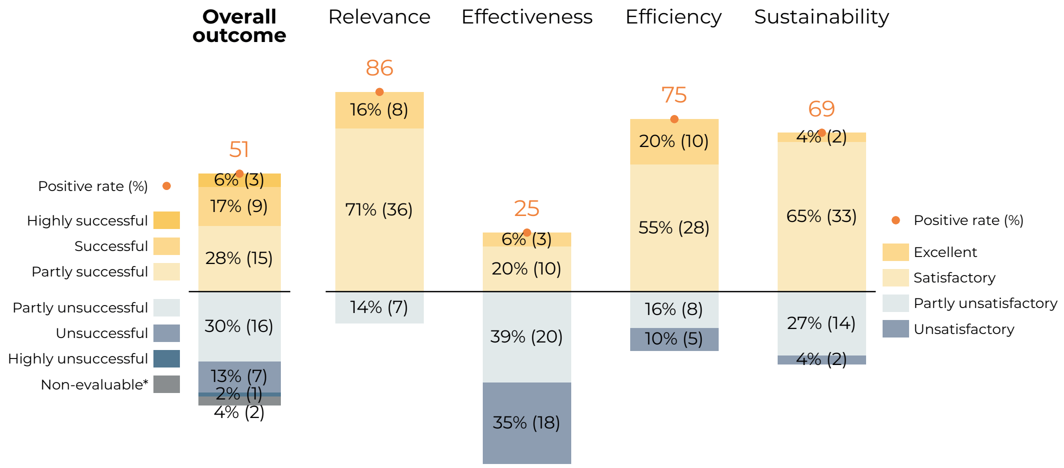


Figure 2.9

IDB Invest operations ratings, 2023

Source: OVE.

Note: * Includes two SAT operations without XSRs, thus rated negative in overall outcome (*N* = 53 operations). These are not included in core criteria ratings (*N* = 51). Percentages may not add up to 100 due to rounding.

2. Core criteria ratings

2.17 Relevance was the highest-rated core criterion, with 44 of 51 operations (86%) achieving positive ratings.[17] The 44 operations rated positive in relevance had objectives that were aligned with country needs and IDB Group priorities, and they had an adequate vertical logic. In contrast, the 7 operations rated negative suffered from a weak vertical logic at project approval, in combination with other issues that included design shortcomings such as a currency mismatch or an inadequate

16 In the DEO 2023 (document GN-3173), Management reported all four SAT operations separately from the regular cohort of operations.

17 For core criteria, only validated operations are considered, thus the two SAT operations without XSRs are excluded from the total hereafter.

client selection (5 operations), misalignment with country realities (2), and misalignment with IDB Group priorities at the time of evaluation (1 operation that supported investments in oil and gas). These factors are not mutually exclusive.

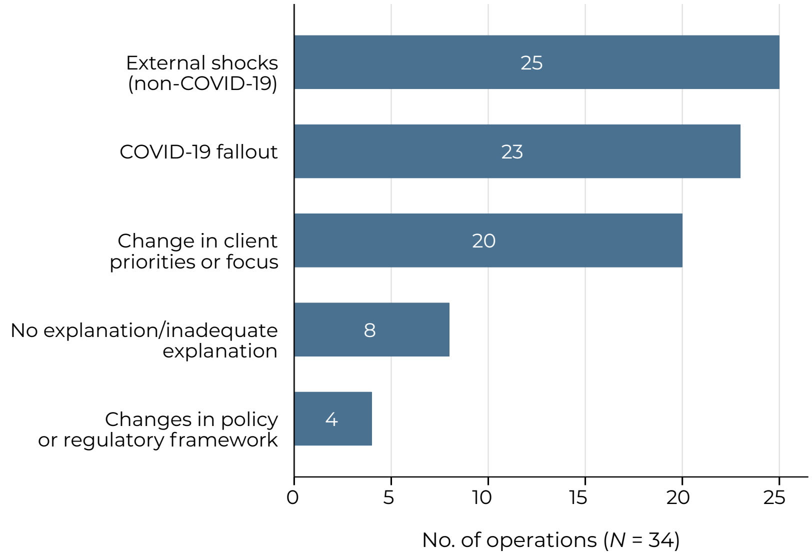
2.18 Effectiveness ratings were positive for 13 of 51 operations (25%); underachievement was the main factor behind these results. The 13 operations rated positive achieved more than half of their objectives, attribution of outcomes to the operation was plausible, and none of their objectives was rated "unsatisfactory." The 38 operations rated negative were mainly from the Financial Institutions (FI) segment (61%) and the Corporates (CO) segment (34%). Nearly all FI operations in the 2023 cohort had negative effectiveness ratings (23 of 24), while about two-thirds of CO operations did so (13 of 21). Overall, among the 38 operations rated negative in effectiveness, 22 (58%) were affected by the combination of underachievement (evidence that the operation did not achieve most of its development objectives) and poor M&E (insufficient evidence to demonstrate achievement of objectives due to measurement issues). In turn, 12 (32%) were exclusively affected by underachievement, and 4 (11%) exclusively by poor M&E.

2.19 According to the XSRs, the most frequent reason for underachievement was external shocks. Figure 2.10 summarizes the reasons for underachievement according to the XSRs. The most frequently cited reason was external shocks (other than COVID-19), such as changes in market conditions (14 operations mentioned changes in market prices, supply and demand, or increased competition), followed by macroeconomic shocks (9 operations), sociopolitical unrest (6), and climate events (5). Operations citing market conditions tended to be spread across multiple countries, but a majority of those citing macroeconomic shocks were concentrated in a few. Sociopolitical unrest affected operations in five countries. Climate events affected agriculture operations in three countries.

Figure 2.10

Reasons behind underachievement according to XSRs

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



Note: Numbers do not add up to 34 because reasons are not mutually exclusive.

2.20 The COVID-19 fallout was the second most frequently cited reason for underachievement, at times leading to a change in client priorities. Some FI clients who were focusing before 2020 on higher-risk portfolios (e.g., microenterprises, immigrants, green lending) changed their focus to lower-risk portfolios when the pandemic hit, given the uncertainty. Other FI clients saw their portfolios shrink with the pandemic, as demand for financial products decreased. In a few cases, the government's COVID-19 response expanded access to cheaper credit options, making the FI client's products less attractive. Lockdowns also affected clients whose activities depended heavily on mobility (e.g., construction, service industry). As for the remaining reasons for underachievement, there were 8 operations whose XSRs lacked an adequate explanation for at least one of their outcome indicators with unmet targets, and 4 operations that were affected by policy or regulatory changes (e.g., changes in banking regulations, government subsidies, or public spending).

2.21 In 8 FI operations, the changes in client priorities were related to a shift away from the target segments; and in 5, there had already been previous ineffective projects with the same clients. The analyses presented in the XSRs indicate that in 8 of the 20 operations where client priorities changed, effectiveness was affected because the clients' business strategy shifted away from the target segments. Box 2.2 describes the shifts in the 8 clients, which account for one-third of all FI operations in this cycle. Most cases involved a shift from micro, small, and medium-sized enterprises (MSMEs) or SMEs toward less risky segments, such as bigger companies. OVE reviewed the Annual Supervision Reports (ASRs) of these 8 operations. All of them detected that the target portfolio was not growing as expected years before the XSRs were prepared, and most of them detected a shift away from the target segment. Thus, supervision mechanisms worked well in detecting the shift. However, OVE found no evidence of IDB Invest undertaking follow-up measures (e.g., to nudge the clients into reorienting their lending towards the original target segments). Moreover, OVE also found that in 5 FI operations rated negative in effectiveness, there had been previous operations with the same clients that were also rated negative in effectiveness at the time. IDB Invest has an action plan to improve FI performance, including by selecting clients with a strong strategic focus on the target segment and using knowledge from previous operations to inform new ones (document GN-3173). However, to date the action plan does not include specific measures to deal with cases where shifts in targets segments are identified.

Box 2.2. FI operations where client priorities shifted away from the target segments

- **Operation 1.** The objective was to contribute to financing, including longer-term financing, for SMEs in agriculture. According to the XSR, due to the contracting economy and the COVID-19 crisis, the client adopted a more conservative strategy, reducing operations with smaller assets and shifting toward shorter-term loans.
- **Operation 2.** The objective was to increase the SME portfolio, especially in rural areas. Per the XSR, the client showed limited commitment to the SME segment, and evidence suggests that it adopted a more conservative strategy even before project approval, shifting toward less risky clients.
- **Operation 3.** The objective was to grow the SME factoring and leasing portfolios. Per the XSR, even before COVID-19 and a social crisis in the country, evidence suggests that the client was shifting toward bigger clients.
- **Operation 4.** The objective was to increase access to green financing. In the face of a deteriorating economic situation that made green financing more challenging, the XSR states that the client adopted a more conservative strategy possibly focused on established clients rather than on expanding to new areas like green financing.
- **Operation 5.** The objective was to grow the MSME portfolio. The XSR ascribes the fall in the microfinance portfolio to a change in the client strategy after COVID-19 and the political uncertainty in the country. This involved a shift from very small sub-borrowers to less risky ones, as the very small were considered inefficient from an operational point of view.
- **Operation 6.** The objective was to increase access to credit for microentrepreneurs. The XSR identified an increase in the average loan size that suggests a shift in focus toward less risky segments and presents data indicating that the share of loans to high-income borrowers increased over time.

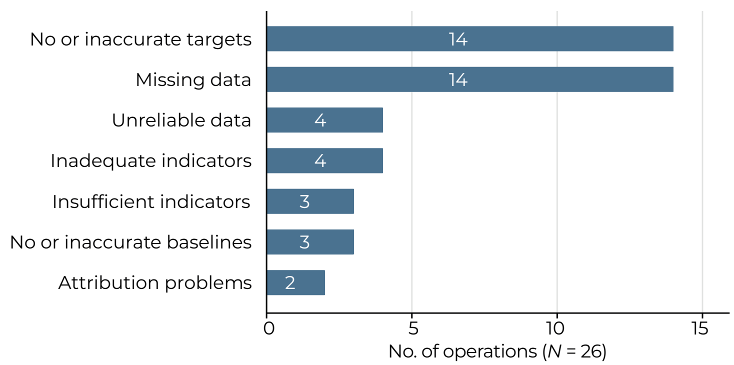
- **Operation 7.** The objective was to grow the outstanding SME portfolio. Per the XSR, the number of SME loans decreased after COVID-19, as the client took a more conservative approach and focused on sub-borrowers less affected by the pandemic.
- **Operation 8.** The objective was to increase access to credit for SMEs in agribusiness. The XSR mentions that, after COVID-19 and changes in policy and market conditions, the client changed its strategy in order to prioritize asset quality and restrict origination to sub-borrowers with higher credit ratings, affecting the SME portfolio growth. Data on average loan size also points to a shift in focus toward bigger sub-borrowers.

2.22 Effectiveness ratings were also affected by M&E issues. Among the 26 operations with M&E issues, 14 had either no targets at approval or targets that were based on inaccurate data or were too ambitious (Figure 2.11). Another 14 operations were affected by missing data as a result of clients failing to provide data for all relevant years or sufficiently disaggregated to know whether the intended beneficiaries were reached. Other less common M&E issues were unreliable data, inadequate or insufficient indicators to measure project objectives, lack of or inaccurate baselines, and attribution problems.

Figure 2.11

Types of M&E issues, IDB Invest operations

Source: OVE.



Note: Numbers do not add up to 26 because M&E issues are not mutually exclusive.

2.23 While M&E issues affected the ratings, IDB Invest's efforts helped reduce the magnitude of the problem. OVE identified that the 12 operations affected exclusively by underachievement and not by M&E did have, initially, several M&E challenges that were tackled by IDB Invest's Development Effectiveness Division (DVF). DVF addressed several M&E shortcomings by: (i) correcting baselines or targets of outcome indicators with proper justification, (ii) collecting additional data from external sources to complement XSR analyses, (iii) consulting with the clients to request missing data, and (iv) proposing adjustments to results matrices that improved evaluability. In addition, when loan documents did not provide a clear definition of project objectives, XSRs clearly spelled out how these were derived. These actions were systematically documented in the XSRs and provided a more complete picture of project performance.

2.24 Efficiency ratings were positive for 38 of 51 operations (75%). Financial and economic benefits of these 38 operations exceeded project costs, considering the time-value of money. In turn, 12 of the 13 operations rated negative had both financial and economic efficiency shortfalls, and 1 only economic. The most common factors affecting the efficiency of these operations were that the financial or economic returns of the operation were lower than the cost of capital (9 operations) and that FI operations faced a lower-than-expected growth in their target portfolios or could not reach the intended beneficiaries (9). Some operations also experienced higher-than-expected nonperforming loan (NPL) rates (6), as well as complications derived from COVID-19 (5). Three operations did not provide sufficient information, preventing a full efficiency assessment.

2.25 Sustainability ratings were positive for 35 of 51 operations (69%). These 35 operations achieved results that were likely to continue over time and complied with E&S safeguards requirements. In turn, of the 16 operations rated negative, 7 were exclusively affected by an uncertain continuity of results, 5 exclusively by E&S performance, and 4 by both factors. Continuity of results was at risk due to an uncertain environment (7 operations)—including macroeconomic shocks, changing market conditions, and the political context—as well as an unstable financial standing of the client (7) and shifts in client priorities (5). E&S performance was affected by shortfalls in fulfilling requirements (6), such as lack of reporting, unmet commitments, and little progress in implementing E&S action plans. Examples of social safeguards issues include some operations (4) that failed to mitigate health and hygiene risks or were found noncompliant with labor laws, and examples of environmental safeguards issues (2) include clients that did not have a complete environmental management system.

2.26 All in all, among the 51 operations, 27 achieved a positive overall outcome rating, of which 12 had positive ratings in all four core criteria. These 12 operations have a potential to generate learning for future operations. Figure 2.12 shows one bullet point for each of these 12 operations, summarizing the objectives that they achieved (among all project objectives defined at approval) and the countries where they were implemented. Annex IV provides further details on these operations and their achieved objectives.

Figure 2.12

IDB Invest operations with positive ratings in all core criteria

Source: OVE, based on XSR validations (see details in Annex IV).

IDB Invest operations

Corporates segment

- Expanded telecommunications coverage (BO)
- Increased a client's MSME portfolio (BR)
- Increased broadband access and reduced emissions (BR)
- Increased access to housing (ME)
- Strengthened leather value chain and exports (PR)
- Strengthened agricultural supply chain and trade (PR)
- Increased jobs, trade, and energy generation (UR)

Infrastructure and Energy segment

- Increased generation and reduced emissions (AR)
- Improved port productivity and created jobs (EC)
- Increased generation and reduced emissions (ES)
- Increased trade and jobs (PN)

Financial Institutions segment

- Increased a client's SME portfolio (EC)

Note: AR: Argentina. BR: Brazil. BO: Bolivia. EC: Ecuador. ES: El Salvador. ME: Mexico. PN: Panama. PR: Paraguay. UR: Uruguay.

3. Noncore criteria ratings

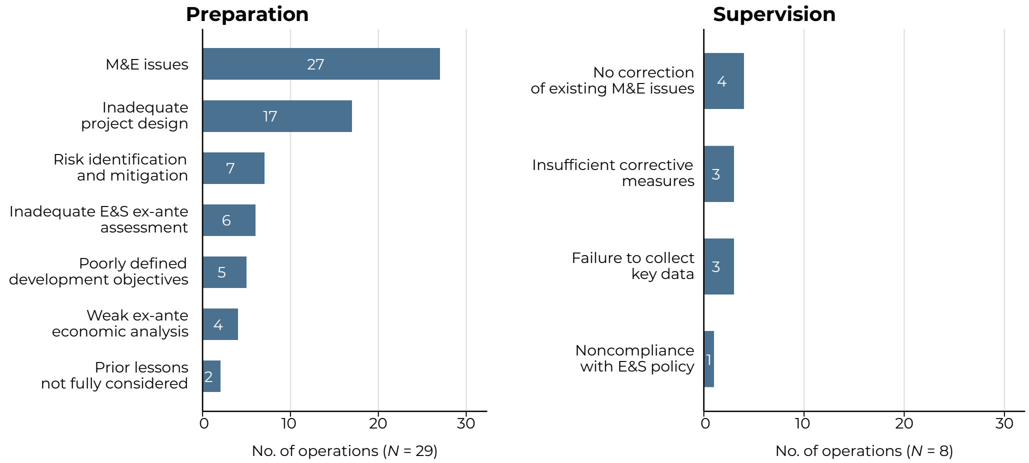
2.27 Of the 51 operations, 46 (90%) were rated positive in additionality and 43 (84%) in investment profitability. On additionality, the 46 operations rated positive provided the client with sufficient financial or nonfinancial additionality in pursuing the project. Among the 5 rated negative, none provided the client with financial additionality, and 3 of them also failed to provide nonfinancial additionality. On investment profitability, the 43 operations with positive ratings had gross profits that met or exceeded the benchmarks set at approval, while the 8 with negative ratings did not, due to prepayments or cancellations.

2.28 Work quality was the lowest-rated noncore criterion, with positive ratings for 21 of 51 operations (41%); most shortcomings were observed during preparation. This criterion rates the quality of IDB Invest's work during the project preparation phase (screening, appraisal, and structuring) as well as during supervision. In 22 of the 30 operations rated negative (73%), there were performance shortfalls exclusively during preparation, while in 7 (23%) they occurred both during preparation and supervision, and in 1 (3%) only during supervision. Figure 2.13 summarizes the most common shortfalls during each stage. M&E issues predominated during preparation, typically due to deficiencies in the results matrices and indicators. These were followed by other design problems, such as a weak analysis of country realities or needs assessment, an inadequate instrument for the client's needs, or an inadequate client selection. During supervision, existing M&E problems were not corrected in a timely manner, ultimately affecting final evaluation. There were also insufficient corrective measures (for example, when a loan got renewed despite clear underperformance) and insufficient efforts to collect key data from the client.

Figure 2.13

Factors behind negative work quality ratings, IDB Invest operations

Source: OVE.



Note: Numbers do not add up to 29 or 8 in the graphs, respectively, as shortcomings are not mutually exclusive.

C. IDB Invest operations with Mini-XSRs

2.29 This section presents the validation results of an additional 52 IDB Invest operations for which XSRs were not prepared when due. In the past, there were several challenges to identifying the moment when operations had reached EOM and were thus ready for XSR preparation. During 2020–2021, IDB Invest created an automated system that calculates the EOM date for all operations. This system allowed identifying 78 operations that had already reached EOM or exited the portfolio during 2015–2019 but that went undetected at the time, and their XSRs were therefore overdue. For 26 of them, IDB Invest still had a relationship with the clients and, based on the data it was able to collect, it prepared XSRs that were validated by OVE.[18] However, the remaining 52 operations had already been repaid or prepaid by the clients, and IDB Invest no longer had a relationship with them or data collection was exceedingly challenging.[19] As an exceptional measure, IDB Invest and OVE agreed on a simplified methodology—the "Mini-XSR"—to account for these operations. Mini-XSRs were prepared in two batches, one in 2022 (with 24 operations) and one in 2023 (with 28 operations). OVE's 2022 report presented the validation results of the first batch; this report does so for the full set of 52 operations for which Mini-

18 Of these 26 operations with full XSRs, 15 were validated in 2022 and 11 in 2023.

19 For 41 of the 52 operations, IDB Invest no longer had a relationship with the client. For the remaining 11, IDB Invest was still in contact with the client, but these operations were old and data collection was too challenging. As a result, IDB Invest agreed with OVE on preparing a Mini-XSR instead of a full XSR for these 11 cases, too.

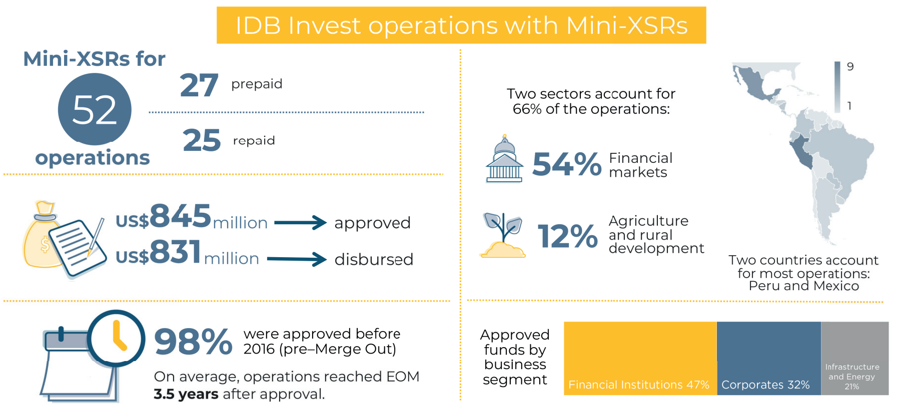
XSRs were prepared. Going forward, no additional Mini-XSRs are expected, as all operations will have full XSRs. Figure 2.14 shows key characteristics of the 52 operations with Mini-XSRs.

2.30 Results are presented separately from the regular cohort of operations, as the Mini-XSR methodology is not comparable to that of XSRs. Mini-XSRs assess all core and noncore criteria except for additionality. Unlike XSRs, the rating scale in Mini-XSRs is only “satisfactory” or “unsatisfactory,” with a third rating of “lack of data” in exceptional cases where available data was insufficient to assign a rating. The detailed methodology can be found in Annex V.

Figure 2.14

Characteristics of IDB Invest operations with Mini-XSRs

Source: OVE, with data from IDB Invest (2016, 2023a, 2023b, 2023c, 2023d).



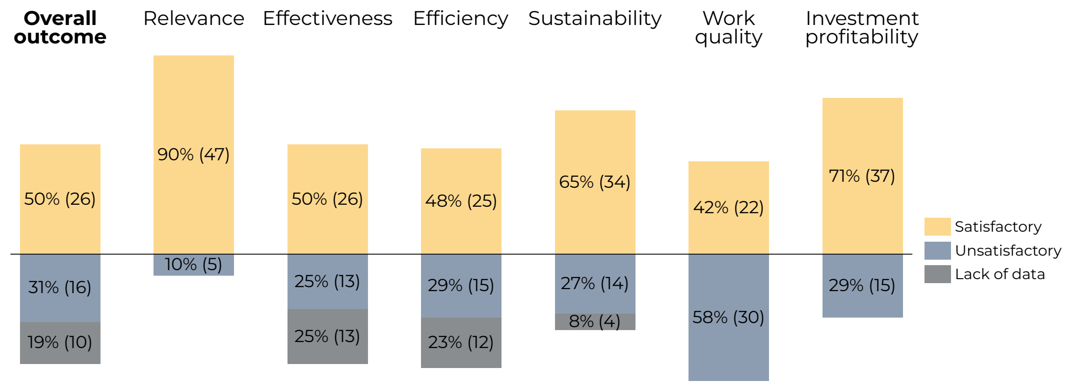
Note: Approval and disbursement data up to December 31, 2022.

2.31 Of the 52 operations, 26 (50%) had “satisfactory” overall outcome ratings. There were 16 operations (31%) rated “unsatisfactory” and 10 (19%) rated as “lack of data.” Figure 2.15 shows ratings for overall outcome, core, and noncore criteria.

Figure 2.15

Ratings of IDB Invest operations with Mini-XSRs

Source: OVE.



Note: *N* = 52 operations.

2.32 Of the 52 operations, 47 (90%) were rated "satisfactory" in relevance. Among the 5 operations with "unsatisfactory" ratings, 4 had design shortcomings such as an inadequate selection of client, instrument, or loan structure. In addition, 1 of these was also not aligned with the CS at approval. For the remaining operation, context conditions changed during implementation, and project design was not appropriately adapted to remain relevant.

2.33 Of the 52 operations, 26 (50%) had "satisfactory" effectiveness ratings. The 13 operations (25%) rated "unsatisfactory" faced underachievement, as there was evidence that they did not meet the targets of the outcome indicators under most of their objectives. In most cases, this meant that the operations failed to grow the target portfolios (4 operations), reach the intended beneficiaries (4), create jobs (3), or increase production (2) at the expected levels. For the 13 operations (25%) classified as "lack of data," the most frequent reason was that appropriate achievement data was nonexistent (8). In other cases, there was limited data that either pointed to inconsistencies (3) or only measured progress for a portion of the project's life (1), which proved too inconclusive to assign a rating. In one case, the lack of clearly defined objectives at design prevented the collection of relevant achievement data. Most of these issues stemmed from the fact that objectives were not adequately defined at approval and not sufficiently monitored during supervision.

2.34 Among the 52 operations, 25 (48%) had "satisfactory" efficiency ratings. "Unsatisfactory" ratings for 15 operations (29%) were due to a combination of reasons, most frequently the client's low financial performance (13), followed by high NPL rates (5) as well as below-target portfolio growth (5). For 12 operations (23%) rated as "lack of data," appropriate data to measure efficiency was nonexistent (9), or the limited available data was inconsistent or unreliable (3).

2.35 Out of 52 operations, 34 (65%) had "satisfactory" sustainability ratings. Nearly all 14 operations rated "unsatisfactory" were affected by a low likelihood that their results would continue over time (13), often because the client's financial standing worsened, in some cases leading to a shift away from the target segments. Several other operations had shortcomings in E&S safeguards compliance (7). The 4 operations (8%) rated as "lack of data" had insufficient (3) or nonexistent (1) data to assign a rating.

2.36 Of the 52 operations, 22 (42%) achieved "satisfactory" work quality, the lowest-rated criterion. Among the 30 (58%) operations rated "unsatisfactory," 25 faced problems during preparation (screening, appraisal, and structuring) and 24 during supervision. During preparation, most operations had problems in their M&E frameworks (15 operations), followed by an inadequate project design (14) and inadequate risk mitigation strategies (6). During supervision, the most frequent shortcoming was that key monitoring data was not collected (22, including 4 for which some ASRs were not prepared), followed by noncompliance or underreporting on E&S safeguards (3).

2.37 Of the 52 operations, 37 (71%) were rated “satisfactory” in investment profitability. In turn, all 15 (29%) operations with “unsatisfactory” ratings yielded less than 60% of the expected interest income for IDB Invest. In 10 of these cases, it was unclear whether IDB Invest contractually required a prepayment penalty, whether such penalty was paid, and whether it was included in the investment profitability calculation.

03

Performance Over Time, 2018–2023

3.1 This section compares 2023 results with previous validation cycles. The period of analysis is 2018–2023. Given changes in the PCR methodology between the 2020 and 2021 cycles, precautions are taken for data analysis of IDB operations by clearly segmenting the strictly comparable periods in the figures below as well as by controlling for the change in methodology in the regression analyses conducted. Table 3.1 shows which methodology was applied to each validation cycle.

Table 3.1. Changes in the evaluation methodology used by OVE

Validation cycle	Methodology for IDB operations	Methodology for IDB Invest operations
2018	PCR Guidelines 2018 (document OP-1696-1)	XSR Guidelines 2018
2019		
2020		
2021	PCR Guidelines 2020 (document OP-1696-6)	
2022		
2023		

Source: OVE.

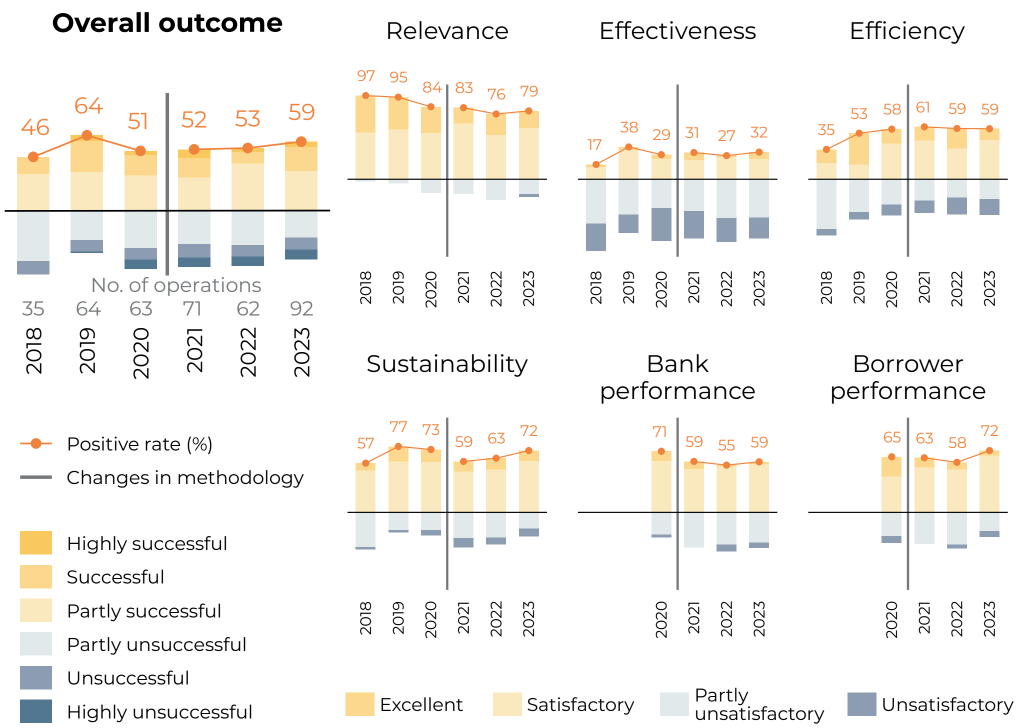
A. IDB performance, 2018–2023

3.2 After oscillating around 52% over the past three years, the positive rate in overall outcome ratings rose to 59% in 2023. Figure 3.1 shows time series data for all 387 IDB operations validated since 2018. Relevance ratings had a downward trend during 2018–2020 (the first comparable period), which continued during 2021–2022 but slightly reversed in 2023. However, relevance ratings remain lower than in 2021. Effectiveness has consistently been the criterion where operations have performed lowest, with an average positive rate of 30% in the most recent comparable period (2021–2023). Efficiency ratings, after an upward trend in 2018–2020, have remained relatively stagnant in 2021–2023, averaging 60% positive. In turn, sustainability ratings increased over the past two years. Bank performance has remained relatively stable in 2021–2023, averaging 57%, while borrower performance showed a marked improvement in 2023 compared to previous years.

Figure 3.1

IDB operations ratings over time, 2018–2023

Source: OVE.



3.3 Performance is heterogeneous across sectors and regions, though not so across lending instruments. As the IDB Group continues to accumulate performance data every year, there is a richer evidence base for analyzing differences in performance within the institution. Considering all 387 operations with PCRs validated during 2018–2023, Figure 3.2 shows the positive rates in overall outcome for each sector, region, and lending instrument.[20] OVE conducted regression analyses (detailed in Annex IV) to test whether differences in performance across sectors, regions, and lending instruments are statistically significant, while controlling for the change in methodology. Performance is measured as the positive rate in overall outcome. Findings show the following:

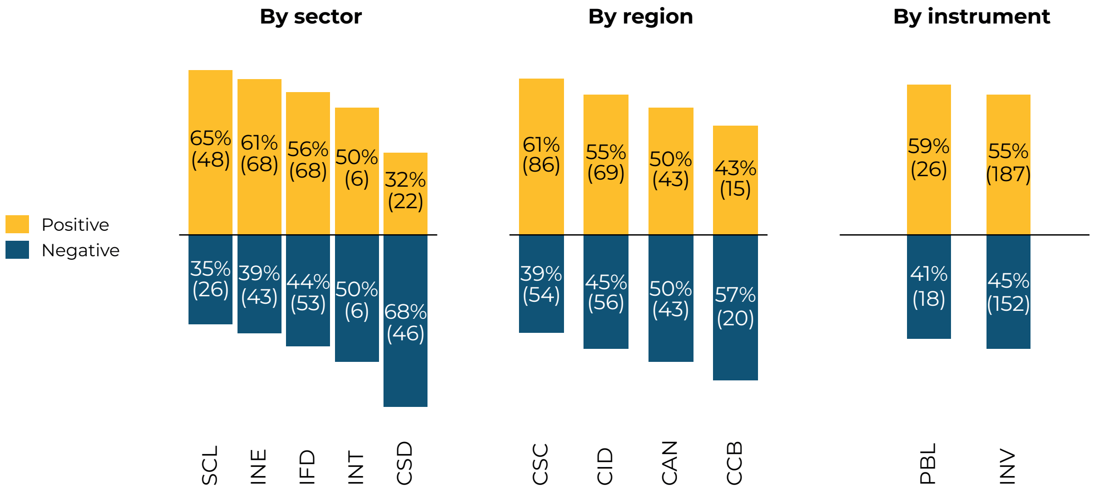
- **Sectors:** Social Sector (SCL) operations had the strongest performance and Climate Change and Sustainable Development Sector (CSD) operations the weakest. Performance among the rest of the sectors does not significantly differ.
- **Regions:** Operations in the Country Department Southern Cone (CSC) scored significantly higher. Among the rest of the regions, performance does not significantly differ.
- **Instruments:** Performance of investment operations (INV) is not significantly different from performance of PBLs.

20 Annex IV presents trends over time for each group.

Figure 3.2

Overall outcome ratings for all IDB operations validated in 2018–2023

Source: OVE.



Note: Number of operations in parentheses. Out of 387 operations validated by OVE, sector data excludes 1 operation shared between the Infrastructure and Energy Sector (INE) and CSD; region data excludes 1 regional operation covering countries from more than one region; and instrument data excludes 4 operations that were a combination of INV and PBL. INV includes all modalities of investment loans as well as investment grants. PBL includes multi-tranche PBLs and PBP series.

B. IDB Invest performance, 2018–2023

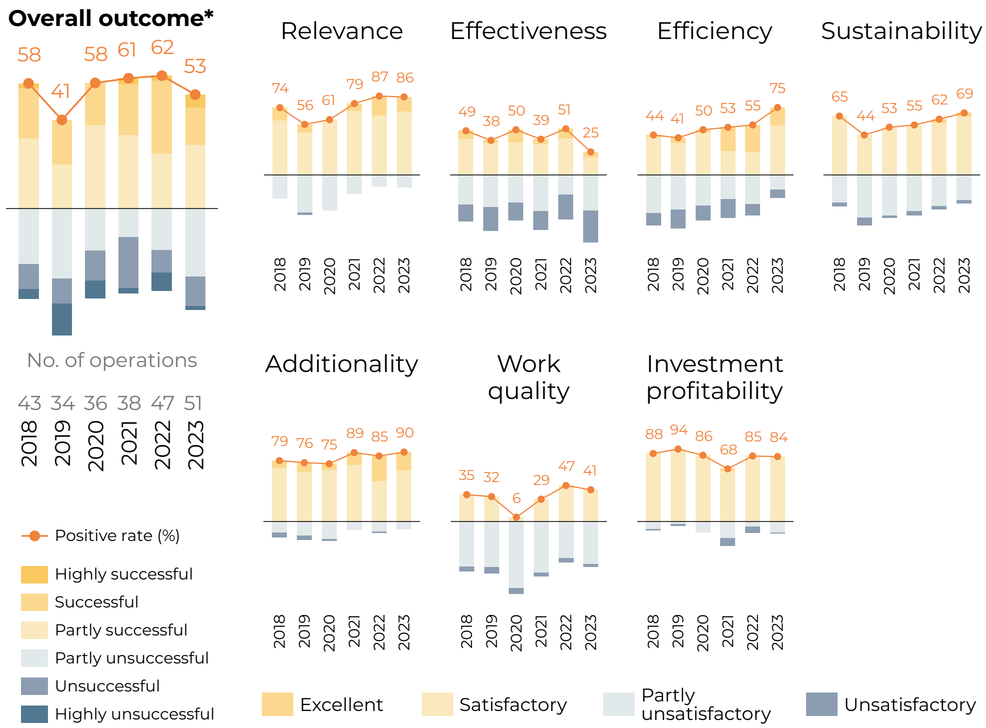
3.4 After oscillating around 60% over the past three years, the positive rate in overall outcome ratings dropped to 53% in 2023 (excluding SAT operations without XSRs). Figure 3.3 shows time series data for all 249 IDB Invest operations validated since 2018. For this analysis, the data excludes SAT operations without XSRs to avoid introducing comparability issues with prior years.[21] Relevance ratings had an upward trend since 2019, but this came to a halt in 2023. Effectiveness has performed lowest among core criteria, with a zigzagging trend that fell to its lowest value in 2023. By contrast, efficiency and sustainability ratings have continued their upward trend since 2019. Additionality reached its highest value in 2023 with a positive rate of 90%, continuing a trend with similarly high values in recent years. Work quality has varied considerably but has remained consistently lowest among noncore criteria, never achieving a positive rate higher than 47%. Investment profitability has been relatively variable, averaging 84% over the period of analysis.

21 SAT operations without XSRs started to be assigned a negative rating and counted toward the total number of operations for overall outcome ratings in 2022. However, they have been excluded from the analysis of performance over time to avoid introducing comparability issues with prior years. This corresponds to 4 operations in 2022 and 2 in 2023. In turn, SAT operations that prepared XSRs (2 in 2023) were validated and are therefore included, as every year all validated operations are included, posing no comparability issue.

Figure 3.3

IDB Invest operations ratings over time, 2018–2023

Source: OVE.



Note: * Excludes SAT operations without XSRs (4 in 2022 and 2 in 2023).

3.5 Performance is heterogeneous across business segments and regions, and somewhat heterogeneous between pre– and post–Merge Out operations.[22] Considering all 249 operations with XSRs validated during 2018–2023, Figure 3.4 shows the positive rates in overall outcome for each business segment, region, and for pre– and post–Merge Out operations.[23] The regression analyses conducted for IDB Invest (detailed in Annex IV) test whether differences in performance are statistically significant. Performance is measured as the positive rate in overall outcome. Findings show the following:

- **Business segments:** The Infrastructure and Energy segment (IE) scored highest, and FI lowest. The CO segment performance is higher than FI's but only in some of the tests performed, so the finding is not robust.
- **Regions:** There is no statistically significant difference among regions except for the Caribbean, which performed lowest. However, the data only includes a very small number of operations in the Caribbean (6).

22 The Merge Out was a process that consolidated the private sector lending activities of the IDB and the Inter-American Investment Corporation (IIC) into IDB Invest. It took effect on January 1, 2016. Differences among lending instruments are not presented for IDB Invest, given that a large majority (90%) of validated operations are senior loans.

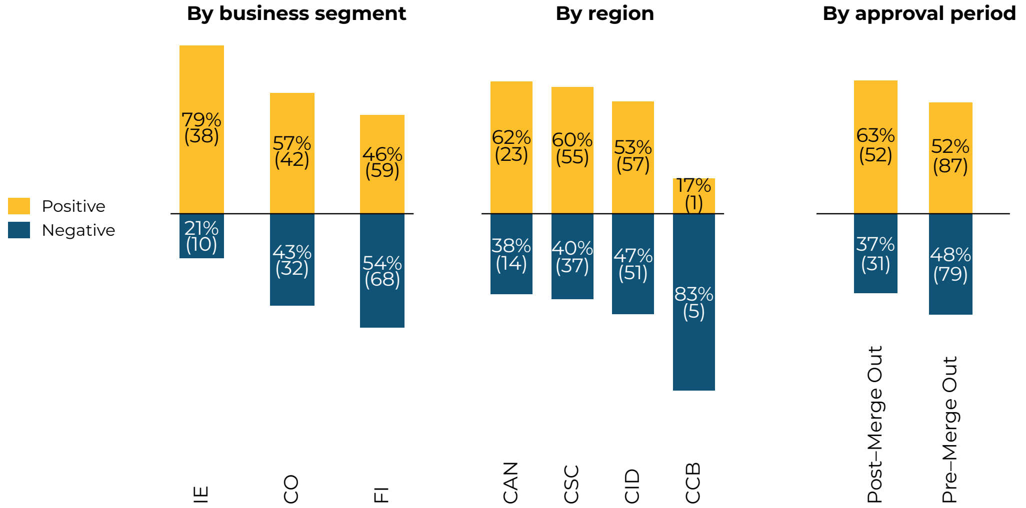
23 Annex IV presents trends over time for each group.

- **Pre– and post–Merge Out operations:** Post–Merge Out operations scored higher than pre–Merge Out operations but only in some of the tests performed, so the finding is not robust.

Figure 3.4

Overall outcome ratings for all IDB Invest operations validated in 2018–2023

Source: OVE.



Note: Number of operations in parentheses. Of 249 IDB Invest operations, region data excludes 6 regional operations covering countries from more than one region. Excludes SAT operations without XSRs (4 in 2022 and 2 in 2023).

04

Quality of Self-evaluations

A. Quality of PCRs

4.1 About half of the PCRs were of satisfactory quality. OVE assigns a rating to the quality of the self-evaluation report by Management. This rating covers compliance with PCR Guidelines, completeness of the analyses, quality of evidence provided, quality of the analyses, documenting changes in design and results matrix, candor, and lessons, among other aspects (see Annex VI). In this cycle, OVE assessed the quality of 78 PCRs corresponding to 92 operations.[24] Of the 78 PCRs assessed, 38 (49%) obtained a positive rating. This percentage remained virtually unchanged compared to last year, when 48% of PCRs were rated positive. PCRs rated negative (40) had issues with the quality (36) and completeness (33) of the analyses presented. Table 4.1 shows some of the most common shortcomings identified by OVE in the quality of the PCRs reviewed, which are similar to those found last year. In response to OVE's recommendation in 2022 for Management to strengthen the quality of PCRs, Management prepared an action plan that includes four main actions: (i) to strengthen the quality assurance instrument used by the Strategic Development Effectiveness Division (SPD/SDV) to review PCRs prior to submission to OVE; (ii) to ensure that PCR teams report on the reasons for underachievement; (iii) to strengthen SDV's quality assurance instruments for economic analysis; and (iv) to provide training to IDB operational staff, in collaboration with OVE. These actions were deemed partially relevant by OVE given shortcomings identified in the definition of the specific activities to be undertaken. OVE underlines the importance of addressing these shortcomings and promptly implementing and completing the actions.[25]

Table 4.1. Most common shortcomings of PCRs rated negative

Type of shortcoming	Number of PCRs (out of 40)
Related to effectiveness, such as: insufficient analysis of how changes during implementation affected the project's ability to achieve its objectives, failure to address deficiencies in results matrices, weak attribution analyses, calculation errors or unclear calculation methods, problems with the data used to measure indicators.	31 (78%)
Related to efficiency, such as: questionable assumptions in economic analyses, unreliable data (sources, errors), no explanation of time and cost overruns, detailed economic analysis not provided in the annexes.	26 (65%)
Related to relevance, such as: weak analysis of the project's vertical logic, insufficient analysis of how project changes during implementation affected relevance.	24 (60%)

24 Management presented one PCR for 14 operations in Brazil (completed under the PROFISCO programmatic approach) and one PCR for 2 operations in Haiti.

25 See OVE's 2022 report on the Recommendation Tracking System (document RE-588).

Type of shortcoming	Number of PCRs (out of 40)
Related to lessons learned, such as: lessons lacking detail or depth, lessons insufficiently supported by evidence, lack of lessons for future operations, lessons not relevant.	16 (40%)
Related to Bank and borrower performance, such as: insufficient reporting on key performance issues.	13 (33%)
Related to sustainability, such as: lack of reporting on actions taken to mitigate risks to continuity of results, omission of key risks to continuity of results, lack of reporting on E&S safeguards implementation and results.	12 (30%)

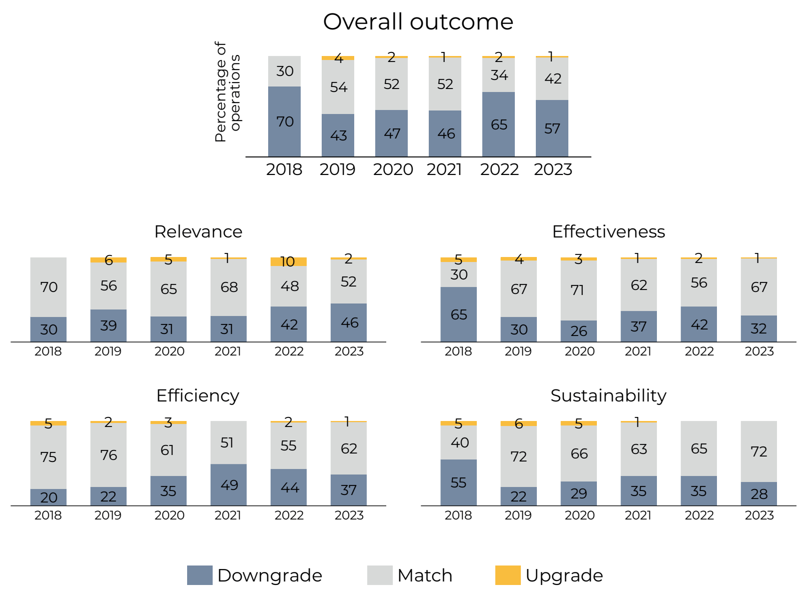
Source: OVE.
Note: The number of PCRs does not add up to 40, nor do the percentages add up to 100, because shortcomings are not mutually exclusive.

4.2 While differences between IDB Management's and OVE's ratings remain large, they narrowed compared to recent years for all core criteria but relevance. In the PCRs, Management self-assigns a rating for each criterion. OVE validates this rating, which can match Management's or result in a downgrade or an upgrade. In 2023, relevance showed the largest discrepancy between Management and OVE since 2018, as OVE downgraded the relevance ratings in 46% of the validated operations (Figure 4.1). For the other three core criteria, however, differences decreased compared to the past two or three years (although there have been smaller differences in earlier years). The overall outcome showed a similar pattern: ratings were downgraded for 57% of the operations—a smaller difference than last year's 65%, though still larger than in most of the earlier years.[26]

Figure 4.1

Differences between IDB Management's and OVE's ratings, 2018–2023 (comparable rating scales only)*

Source: OVE.



Note: Percentages may not add up to 100 due to rounding. * Excludes operations for which Management used a numeric rating scale not comparable to OVE's: data only reflects 20 out of 35 validated operations in 2018, 54 out of 64 in 2019, 62 out of 63 in 2020, and all validated operations thereafter (71 in 2021, 62 in 2022, and 92 in 2023).

26 In overall outcome, out of 92 operations, 52 ratings were downgraded (of which 36 were by one point in the six-point scale, 15 by two points, and 1 by three points), 39 were a match, and 1 was upgraded (by one point).

B. Quality of XSRs

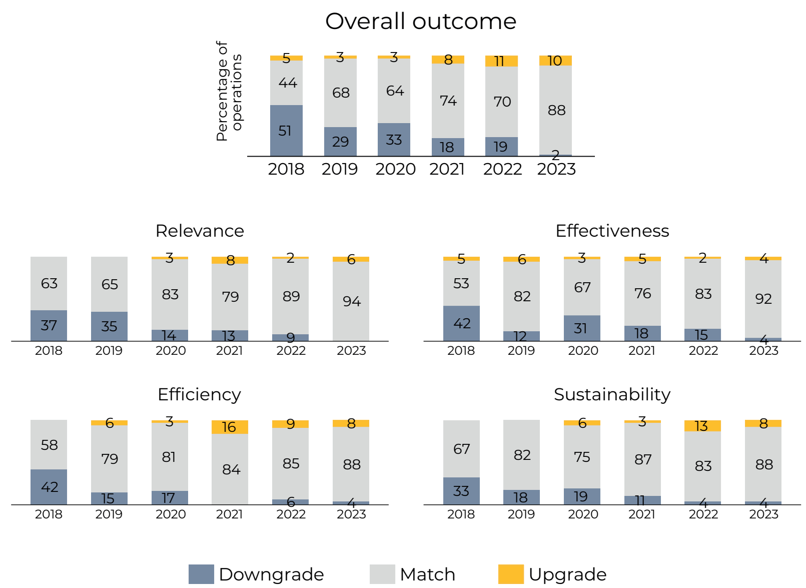
4.3 For the first time, all validated XSRs were of satisfactory quality. OVE assesses the quality of XSRs using the same rating criteria as those used for PCRs (see Annex VI). In 2023, out of the 51 XSRs delivered, OVE rated the quality of all of them as positive, with 43 of them rated as "good" and 8 as "excellent." XSRs closely followed the guidelines, cited sufficient data to support the findings, captured results in a balanced and complete way, were clear and candid, and ratings and lessons were evidence-based. Since 2019, quality of XSRs has been above 80% positive every year, peaking at 92% in 2020. While there was a downward trend in 2021 (89%) and 2022 (81%), the 100% positive in 2023 shows an important recovery.

4.4 Ratings by IDB Invest Management and OVE continued to converge. For the first time since 2018, Management's self-assigned relevance ratings for all 51 validated operations matched OVE's or were upgraded by OVE (Figure 4.2). Effectiveness also saw the smallest difference yet, with only 4% of operations receiving downgrades, compared with last year's 15%. Likewise, efficiency and sustainability both received downgrades for only 4% of the operations—some of the lowest discrepancies observed. As a result of this increasing convergence in ratings, the overall outcome showed only a very small difference, with 2% of operations receiving downgrades, while 88% were a match and 10% were upgraded by OVE. This makes the 2023 cycle the one where overall outcome ratings by IDB Invest Management and OVE have been closest to date.[27]

Figure 4.2

Differences between IDB Invest Management's and OVE's ratings, 2018–2023

Source: OVE.



Note: Percentages may not add up to 100 due to rounding.

27 In overall outcome, out of 51 operations, 45 were a match, 5 were upgraded (by one point in the six-point scale), and 1 was downgraded (by one point).

05

Conclusions and Recommendations

A. IDB operations

5.1 OVE validated the PCRs of 92 IDB operations, of which 59% achieved a positive overall outcome rating. While this fell short of the 70% target set in the 2020–2023 CRF, it is the highest positive rate observed in three years.

5.2 Among core criteria, relevance was the highest rated with 79% positive. Operations rated negative in relevance were mostly affected by weaknesses in their vertical logic. OVE identified that 4 of the 92 operations experienced changes that were arguably substantial and fundamental enough to have warranted a reformulation approved by the Board of Directors. A review and update of the regulations on project reformulation was planned, but Management postponed it and expects to address it in the context of the reform to the Development Effectiveness Framework under the new IDB Group's Institutional Strategy. OVE underlines the importance of promptly proceeding with these actions.

5.3 Effectiveness was the lowest-rated criterion, with 32% positive. This was mostly due to underdelivery of the outputs that the operations were expected to produce, which then led to not achieving most project objectives. M&E issues also affected effectiveness ratings, where attribution problems predominated. However, many PCRs did not provide sufficient explanation of why objectives were not achieved. Efficiency was 59% positive, and the operations rated negative were mostly affected by the low quality of economic analyses. Sustainability was 72% positive, with the negatively rated operations facing an uncertain continuity of results and subpar E&S performance.

5.4 Among noncore criteria, Bank performance was 59% positive. Most identified shortcomings occurred during the project preparation phase. Borrower performance was 72% positive; operations with negative ratings were mostly affected by the capacity of the executing units.

5.5 Trends over time show that, after oscillating around 52% over the past three years, the positive rate in overall outcome ratings rose to 59% in 2023. Performance is heterogeneous across sectors and regions, but not across lending instruments.

5.6 Only 49% of the PCRs were of satisfactory quality. This percentage remained virtually as low as last year, indicating that multiple shortcomings are still present. OVE underlines the importance of promptly completing and implementing Management's ongoing action plan to improve PCR quality. While OVE and IDB

Management continue to have large differences in the ratings assigned to the operations, these differences mostly narrowed in 2023 compared to recent years for all core criteria but relevance.

B. IDB Invest operations

5.7 OVE validated the XSRs of 51 IDB Invest operations. Considering an additional 2 SAT operations without XSRs, 51% achieved a positive overall outcome rating. This fell short of the 65% target set in the 2020–2023 CRF.

5.8 Among core criteria, relevance was the highest rated with 86% positive. Operations with negative relevance ratings were affected by vertical logic weaknesses and other design shortcomings. Effectiveness was the lowest rated, with 25% positive. This was mostly due to underachievement, driven by external shocks, the COVID-19 fallout, and changes in client priorities, particularly among FI operations where 8 clients shifted away from the target segments. OVE found no evidence that follow-up actions were taken in this regard. M&E issues also affected effectiveness ratings (where inaccurate targets and missing data predominated), although IDB Invest's efforts helped reduce the magnitude of the problem. Efficiency was 75% positive, with operations rated negative mostly affected by low financial or economic returns and slow portfolio growth (FI operations). Sustainability was 69% positive, with the negatively rated operations facing an uncertain continuity of results and subpar E&S performance.

5.9 Among noncore criteria, additionality was 90% positive. Investment profitability was 84% positive, and work quality was 41% positive. Most shortfalls among operations with negative work quality ratings occurred during project preparation.

5.10 Trends over time show that, after oscillating around 60% over the past three years, the positive rate in overall outcome ratings dropped to 53% in 2023 (excluding SAT operations without XSRs). Performance is heterogeneous across business segments and regions, and somewhat heterogeneous between pre– and post–Merge Out operations.

5.11 All XSRs were of satisfactory quality. This is the first validation cycle where all self-evaluations by IDB Invest received a positive rating by OVE, exhibiting recovery from last year's 81% positive. In addition, IDB Invest Management's and OVE's ratings continued to converge, with some of the smallest differences observed yet.

5.12 Considering these findings, OVE recommends the following.

For IDB Management:

1. **Ensure: (i) the overall quality of PCRs, including the quality of their economic analyses, by strengthening quality assurance instruments and their application; (ii) that PCRs systematically report on the reasons for underachievement; and (iii) that operational staff receive adequate training, working in collaboration with OVE.** As in 2022, OVE's report for the 2023 validation cycle found issues in the quality of PCRs, including the lack of sufficient explanation of why objectives were not achieved in the effectiveness section and the low quality of economic analyses in the efficiency section. Management should address these issues as part of the action plan prepared in response to OVE's 2022 recommendation on the matter, incorporating OVE's feedback on the plan's relevance.

For IDB Invest Management:

1. **Define measures in the action plan for FI operations for when project supervision finds that the client has shifted away from the target segment and ensure their systematic implementation.** Based on the finding that, in most cases, IDB Invest detected when clients shifted away from the target segment during supervision, but there is no evidence that measures were taken to address such shifts, Management should define measures for dealing with these cases in its action plan for FIs and ensure that they are implemented systematically.

References

AUG (Office of the Executive Auditor). 2022. *SG Loans Reformulation Process*. https://idbg.sharepoint.com/teams/ez-AUG/AUGIDB/SG%20Loans%20Reformulations%20Process%20-%20Audit%20Report.pdf

IDB (Inter-American Development Bank). 2006. *Multiple Works Programs*. Document PR-202.

——. 2010. *Argentina. Proposal for a Conditional Credit Line for Investment Projects (CCLIP) and first individual loan under the line for the Water and Sanitation Program for Urban and Suburban Centers (PAyS)*. Document PR-3557.

——. 2010. *Haiti. Proposal for a nonreimbursable financing for the project "Support to the Shelter Sector Response Plan"*. Document PR-3539.

——. 2013. *Haiti. Proposal for a nonreimbursable financing for the Sustainable Management of Upper Watersheds of Southwestern Haiti – Macaya National Park*. Document PR-4029.

——. 2015. *Substantial and Fundamental Changes to Operations*. Document OA-430.

——. 2016. *Update of the Project Completion Report (PCR) Guidelines*. Document OP-1242-5.

——. 2018. *Proposal to Amend the Development Effectiveness Framework for Sovereign Guaranteed Operations. Approved version*. Document OP-1696-1.

——. 2019. *Implementation Guidelines for the Environment and Safeguards Compliance Policy. Revised version*. Document GN-2208-27.

——. 2020. *Implementation of the Adjustments to the Development Effectiveness Framework Tools for Sovereign Guaranteed Operations*. Document OP-1696-6.

——. 2023a. *Data Marketplace. Table LEDW.ESG_ODS_OPER* [data set].

——. 2023b. *Data Marketplace. Tables LEDW.OPER_ODS_OPER, LEDW.OPER_ODS_ OPERCTY, LEDW.OPER_ODS_OPERUNIT, LEDW.OPER_ODS_OPEVENT, LEDW.SL_ FINBAL_PROD_RPT, LEDW.SL_VW_CASHFLOWS, LEDW.SL_VW_OPER_PRODUCTS* [data set].

IDB Group. 2015. *XSR Guidelines for IDBG Private Sector Operations. Final Version for Pilot. January 2015.* XSR Guidelines 2015.

——. 2018. *XSR Guidelines for IDBG Private Sector Operations.* XSR Guidelines 2018.

——. 2023. *Development Effectiveness Overview 2023.* Document GN-3173.

IDB Invest. 2016. *IDB-IIC Combined Historical Disbursements up to 2015* [data set].

——. 2023a. *Azure SQL Database: iic.Maestro Transaction* [data set].

——. 2023b. *Azure SQL Database: metrics.Operational Approvals RPT* [data set].

——. 2023c. *Azure SQL Database: metrics.Operational Use of Funds RPT* [data set].

——. 2023d. *DVF's XSR Hub: Full Report for OVE 07/25/2023* [data set].

OVE (Office of Evaluation and Oversight). 2017. *IDB and IIC Project Performance: OVE's Review of 2016 Project Completion Reports and Expanded Supervision Reports*. Document RE-520.

——. 2018. *OVE's Review of 2017 Project Completion Reports and Expanded Supervision Reports*. Document RE-530-2.

——. 2019. *OVE's Review of Project Completion Reports (PCRs) and Expanded Supervision Reports (XSRs): 2018/2019 Validation Cycle*. Document RE-544.

——. 2020. *OVE's Review of PCRs and XSRs: The 2020 Validation Cycle*. Document RE-552.

——. 2021. *OVE's Review of PCRs and XSRs: The 2021 Validation Cycle*. Document RE-565.

——. 2022. *OVE's Review of Project Completion Reports (PCRs) and Expanded Supervision Reports (XSRs): The 2022 Validation Cycle*. Document RE-575.

——. 2023. *Management's Implementation of OVE Recommendations: IDB Group's Evaluation Recommendations Tracking System 2022.* Document RE-588.

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/evaluacion

linkedin.com/showcase/idb-ove

@BID_evaluacion