



Project Evaluation

IDB Group Project Performance: The 2024 Validation Cycle

Copyright © 2025 Inter-American Development Bank (IDB). This work is subject to a Creative Commons Attribution 4.0 International Public License CC BY 4.0 (<https://creativecommons.org/licenses/by/4.0/legalcode>). The terms and conditions indicated in the URL link must be met and the respective recognition must be granted to the IDB.

Any and all disputes arising under this license that cannot be settled amicably shall be resolved in accordance with the following procedure. Pursuant to a notice of mediation communicated by reasonable means by either you or the licensor to the other, the dispute shall be submitted to non-binding mediation conducted in accordance with the World Intellectual Property Organization (WIPO) Mediation Rules. Any dispute that cannot be settled amicably shall be submitted to arbitration pursuant to the United Nations Commission on International Trade Law (UNCITRAL) rules. The use of the IDB name for any purpose other than the respective recognition and the use of the IDB logo are not authorized by this license and require an additional license agreement.

Note that the URL link includes terms and conditions that are an integral part of this license.

The opinions expressed in the work are those of its authors and do not necessarily reflect the views of the IDB, its Board of Executive Directors, or the countries they represent.



Inter-American Development Bank, 2025
Office of Evaluation and Oversight
1350 New York Avenue, N.W.
Washington, D.C. 20577
www.iadb.org/evaluation

RE-606
February 2025

Project Evaluation

**IDB Group
Project
Performance:
The 2024
Validation Cycle**

Office of Evaluation and Oversight

TABLE OF CONTENTS

[ACRONYMS AND ABBREVIATIONS](#)

[EXECUTIVE SUMMARY](#)

I.	INTRODUCTION	1
	A. Evaluation methodology and the role of OVE	1
	B. Scope of this report.....	5
II.	RESULTS FOR IDB AND IDB INVEST OPERATIONS VALIDATED IN 2024	5
	A. IDB results in 2024.....	6
	B. IDB Invest results in 2024.....	24
III.	PERFORMANCE BY SECTORS, SUB REGIONS AND INSTRUMENTS 2018–2024	37
	A. IDB performance.....	37
	B. IDB Invest performance	38
IV.	QUALITY OF SELF-EVALUATIONS AND DIFFERENCES BETWEEN MANAGEMENT’S AND OVE’S RATINGS.....	39
	A. Quality of PCRs	39
	B. Differences between IDB Management’s and OVE’s ratings	40
	C. Quality of XSRs.....	42
	D. Differences between IDB Invest Management’s and OVE’s ratings.....	42
V.	CONCLUSIONS.....	43

[REFERENCES](#)

- Annex I [Ratings by operation](#)
Annex II [OVE’s Validation notes of IDB operations](#)

The most recent editions of this report can be found below:

[2017 Report](#)
[2018 Report](#)
[2019 Report](#)
[2020 Report](#)
[2021 Report](#)
[2022 Report](#)
[2023 Report](#)

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

This report was prepared by José Ignacio Sémbler (Team Leader; acting OVE Corporate Cluster Leader), Bertha Briceño, Claudia Figueroa, Luisa Riveros, Analí Pérez, Marina Pupo Lafer, Fernando Barbosa, Jorge Luis González, Aarre Laakso, Julie King, and Melisa Wong with the guidance of Jozef Vaessen (OVE Principal Advisor) and under the overall direction of Marialisa Motta (OVE Director).

ACRONYMS AND ABBREVIATIONS

AP	Action Plan
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
CCLIP	Conditional Credit Line for Investment Projects
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
DELTA	Development Effectiveness Learning, Tracking, and Assessment
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)
M&E	Monitoring and evaluation
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
SAT	Special Assets Division
SCL	Social Sector
SME	Small and medium-sized enterprises
XSR	Expanded Supervision Report

ACKNOWLEDGEMENTS

The following OVE members participated in the validation process of Project Completion Reports (PCRs) and Expanded Supervision Reports (XSRs), which was led by Ana Maria Linares, Corporate Cluster Leader and Senior Advisor (until September 2024), and Bertha Briceño (Lead Specialist):

Anais Anderson	Monika Huppi	Analí Pérez
Josette Arévalo	Dana Michael King	Gabriela Pérez
Andre Averbug	Regina Legarreta	José Claudio Pires
Fernando Carlos Barbosa	Julián Loayza	Juan José Pompilio Sartori
Bertha Briceño	Odette Maciel	Luisa Riveros
César Bouillon	Stephany Maqueda	María Fernanda Rodrigo
Ernesto Cuestas	Roland Michelitsch	Carolina Romero
Claudia Figueroa	Rasec Niembro	José Ignacio Sémbler
Pablo Fleiss	Hans J. Paris	Alejandro D. Soriano
Federico Fraga	Roxana Pedraglio	Roni Szwedzki
Jorge Gallego	Lina Pedraza	Judy Twigg
Gunnar Fabian Gotz	Oliver Peña-Habib	Melisa Wong

The OVE team is grateful to all IDB Group colleagues who prepared the project self-evaluations. The OVE team would also like to extend a special thanks to 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 and for the close collaboration with OVE throughout the 2024 validation process.

EXECUTIVE SUMMARY

This report summarizes the 2024 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 results of its operations: 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 - *Project Completion Report (PCR): Principles and Guidelines* revised in 2020 (document [OP-1696-6, Annex 1](#)) for IDB and the [XSR Guidelines for IDB Group Private Sector Operations](#), revised in 2018 for IDB Invest. Self-evaluations and validations are intended to serve as accountability and learning tools. The accountability goal ensures that the IDB Group uses project resources effectively, efficiently, and sustainably to achieve approved objectives. The learning goal provides lessons to inform the design and implementation of both ongoing and new projects.

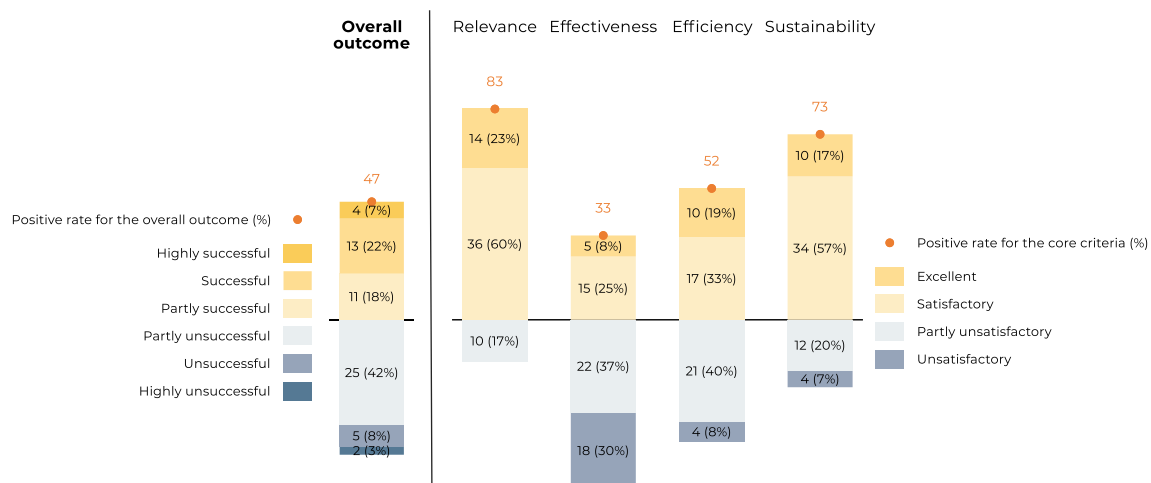
The IDB Group's project evaluation methodology (for both management self-assessment and OVE's validation) is objectives-based, which means that project performance is assessed against the objectives that the projects intended to achieve as defined in their design documents. Management and OVE assess overall project (outcome) performance 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. Projects' overall outcome ratings are calculated as a weighted average of the four core criteria ratings, where effectiveness has the highest weight (40 percent, versus 20 percent for the other three criteria).¹ In addition, both Management and OVE rate two non-core criteria for IDB operations (Bank performance and borrower performance) and three non-core criteria for IDB Invest operations (additionality, investment profitability, and work quality). Non-core criteria do not count toward the overall projects' outcome ratings. Finally, 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. 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 results in 2024

In 2024, OVE validated the PCRs of 60 IDB operations, including 51 investment loans, 7 policy-based loans, 1 combined operation (investment loan and policy-based loan), and 1 investment grant. Of these 60 operations, 47 percent achieved a positive overall outcome rating. This rate falls short of the 70 percent target in the 2020–2023 Corporate Results Framework (CRF). Relevance was the highest-rated core criterion, with 83 percent of operations rated positive, followed by sustainability (73 percent), efficiency (52 percent), and effectiveness (33 percent).

¹ For IDB's policy-based loans (PBLs), efficiency is not rated, and effectiveness weighs 60% instead.

IDB operations ratings, 2024



Source: OVE.

Note: $N = 60$ operations for all criteria except for efficiency ($N = 52$) because seven policy-based operations and one combined PBL and INV operation are not rated for this criterion. Percentages may not add up to 100 due to rounding.

Relevance was the highest-rated core criterion, with 83 percent of operations (50) rated positive, as they had objectives that were aligned with country development needs, government priorities, and IDB Group Country Strategies (CSs). These operations also had adequate project designs, including sound vertical logic. Operations rated negative (10) were mostly affected by weak vertical logic and designs that did not adequately account for country realities. The relevance ratings of some projects may have been affected by unforeseen circumstances (e.g. COVID-19 or shifts in national priorities), which could not be anticipated at approval or easily addressed during implementation. Since ratings, both for self-assessment and validation, are objectives-based as per the PCR guidelines, projects that were not restructured to adapt their objectives to changing circumstances, are rated based on their original objectives. Consequently, projects may receive negative ratings even if they financed activities and achieved results that were relevant during implementation. OVE's Development Effectiveness Framework evaluation (document [RE-583-2](#)) found that restructuring IDB projects is a cumbersome process, based on a benchmarking exercise that compared IDB's policies with those of other Multilateral Development Banks. While the decision to restructure a project ultimately rests with borrowing countries, internal complexity may lead IDB staff to adapt to changing circumstances while maintaining projects' original objectives.

Effectiveness was rated positive for 33 percent of operations (20). The operations rated positive achieved more than half of their objectives (none of which were rated "unsatisfactory") and attribution of outcomes to the operation was plausible. Operations rated negative (40) had a combination of unmet outcome targets and shortcomings in reporting. Unmet outcome targets were more frequent, mostly due to under-delivery of the outputs that the operations were expected to produce, which then led to not achieving most project objectives. The most common reasons for unmet outcome targets were shifts in government priorities, borrowers resource constraints, and project designs that did not consider country realities. The most common reporting shortcoming was missing data for outcome indicators to show progress on project objectives. In some cases, shortcomings in reporting (e.g. lack of data, attribution problems, inadequate indicators) can obscure a project's achievements. Providing accurate targets at project design and collecting and updating data at the time of project implementation and closure is essential to be able to

assess project effectiveness. However, when reporting shortcomings impede the assessment of achievement levels and result in a negative rating, tagging a project as ineffective could be misleading, as neither Management nor OVE can state whether the project has achieved (or not) its outcomes due to insufficient evidence.

Efficiency was rated positive for 52 percent of the investment operations (27). The operations rated positive provided a robust analysis demonstrating that the costs incurred to achieve project results were reasonable. Negatively rated investment operations (25) were mostly the result of PCRs lacking a cost-benefit analysis (CBA) or a cost-effectiveness analysis (CEA) when considered feasible or presenting analyses with caveats that made the calculations of economic rates of return unreliable. Due to the complexity of some projects in certain sectors, conducting a CBA or CEA is not always appropriate, including for projects with multiple components where many (unrealistic) assumptions need to be made about multiple benefit streams that are often difficult to measure or predict. Following the PCR guidelines, in such cases, Management could provide a justification for not conducting these analyses and resort to a cost and time overrun analysis (CTOA). This could lead to a satisfactory rating when the project classification performance remained satisfactory during implementation and no other factors affected efficiency. If no justification is provided or if Management attempts to conduct a CBA or CEA with unrealistic assumptions, the maximum rating allowed by the PCR Guidelines is partially unsatisfactory.

Sustainability was rated positive for 73 percent of operations (44). These operations achieved results that were likely to continue over time and adhered to the environment and social (E&S) standards set by the IDB Group. Operations rated negative (16) were affected by an uncertain continuity of results and/or an unsatisfactory performance in E&S safeguards.

The two non-core criteria for IDB operations, Bank performance and Borrower performance, were rated positive for 73 percent (44) and 58 percent (35) of validated operations, respectively. For the 44 operations that had a positive Bank performance, the IDB facilitated readiness for implementation through proper project preparation and approval and provided timely support to executing agencies during project implementation, ensuring fiduciary compliance. Shortcomings in Bank performance (16 operations) were identified during project preparation and supervision, including poor identification of risks and mitigation mechanisms, weaknesses in the design of their Monitoring and Evaluation (M&E) framework, and deficiencies in project monitoring. For the 35 operations with positive Borrower performance ratings, the borrower showed commitment, supported the quality of preparation and implementation of the operation, and complied with the covenants and agreements established with the Bank. Most of the 25 operations rated negative were affected by the limited capacity of the executing agencies.

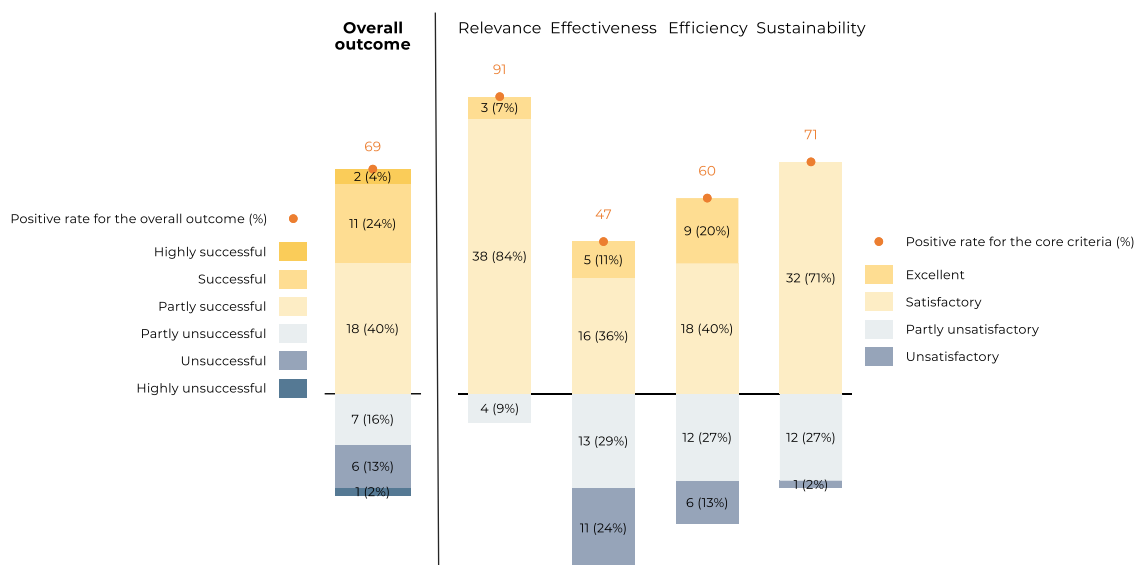
The signals raised during project supervision by the IDB's monitoring tool, the Progress Monitoring Report (PMR), were insufficient as early warnings regarding underperforming projects at completion. Many of the projects that received negative ratings at completion were never classified as "alert" or "problem" in the PMR, indicating that the current signals from the PMR often fail to identify such cases. These findings align with those of OVE's Development Effectiveness Framework evaluation (document [RE-583-2](#)). The shortcomings of the IDB's PMR classification are likely due to the fact that the PMR alert and problem status classification focuses on progress on the project's schedule, cost, and disbursement performance rather than on the project's progress toward achieving its outcomes. Progress towards achievement of project outcomes is a criterion that other

Multilateral Development Banks do consider in their classification and alert systems, together with other relevant criteria, such as project management, procurement, audit, and financial management.

B. IDB Invest results in 2024

In 2024, OVE validated the XSRs of 45 IDB Invest operations, including 31 senior loans, six bonds, two subordinated loans, two reverse factoring operations, one equity, one mezzanine, and two combined operations (one senior loan and subordinated loan, and one senior loan and reverse factoring). Of these 45 operations, 69 percent (31) achieved a positive overall outcome rating. This rate exceeds the target of 65 percent set in the 2020–2023 Corporate Results Framework. Relevance was the highest-rated core criterion, with 91 percent of operations rated positive, followed by sustainability (71 percent), efficiency (60 percent), and effectiveness (47 percent).

IDB Invest operations ratings, 2024



Source: OVE.

Note: Percentages may not sum to 100 due to rounding.

Relevance was rated positive for 91 percent (41) of the validated operations. Positively rated operations had objectives that were aligned with country needs and IDB Group priorities, and they had an adequate vertical logic. Those rated negative (4) suffered from design shortcomings such as inadequate client or instrument selection.

Effectiveness ratings were positive for 47 percent (21) of the validated operations. The operations rated positive achieved more than half of their objectives (none of which were rated “unsatisfactory”) and attribution of outcomes to the operation was plausible. Negative ratings (24 operations) are explained by a combination of unmet outcome targets and shortcomings in reporting, though the former was more common. The most frequent reason behind unmet outcome targets was external shocks, including the COVID-19 pandemic and market conditions. Effectiveness ratings were also affected by reporting shortcomings like missing data or inadequate indicators. As with the effectiveness ratings of IDB operations, those of IDB Invest operations may be adversely impacted by external circumstances and by reporting deficiencies. Circumstances that were unforeseeable at the time of project approval may not be taken into account easily during project

implementation (due, for example, to the difficulties related to restructuring projects), which may result in attributing negative ratings to projects that have achieved meaningful results.

Also, similarly to IDB but to a lesser extent, IDB Invest projects' effectiveness ratings can also be negatively affected by reporting shortfalls, on the basis of which it is not possible to affirm that results did not happen.

Efficiency was rated positive for 60 percent (27) of operations, which had financial and economic benefits exceeding their costs, considering the time-value of money. Negatively rated operations (18) were mostly affected by low financial or economic returns and subpar portfolio growth among financial intermediary operations. As with effectiveness ratings, the efficiency ratings of IDB Invest projects may be affected by unforeseen circumstances and inadequate reporting.

Sustainability was rated positive for 71 percent (32) of operations. Operations with positive ratings achieved results that were likely to continue over time and adhered to the E&S standards set by the IDB Group. Operations rated negative (13) were affected by an uncertain continuity of results and/or an unsatisfactory E&S performance.

The three non-core criteria for IDB Invest operations, additionality, work quality and investment profitability were rated positive for 100% (45 operations), 53% (24 operations) and 80% (36 operations), respectively. For operations with positive additionality, IDB Invest provided the client financial or nonfinancial support that made the project or investment happen, made it happen sooner than planned, and/or improved its development impact and sustainability. For operations with positive work quality (24), IDB Invest effectively carried out its pre-commitment work (screening, appraisal, and structuring) and monitored and supervised project risks, E&S performance, and development results following commitment. Most shortcomings for operations with negative work quality (21) were related to Monitoring and Evaluation (M&E), including issues with the results matrices and indicators, and problems with project designs (selection of instruments and clients). Finally, operations with positive investment profitability (36) had gross profits that met or exceeded the benchmarks set at approval. Operations rated negatively on investment profitability (9) did not meet gross profit expectations due to prepayments or cancellations.

IDB Invest uses the Development Effectiveness Learning, Tracking, and Assessment (DELTA) to measure an operation's potential impact, both at approval and during supervision. During supervision, the DELTA is used to provide alert of problem flags. Unlike IDB operations, most IDB Invest operations with negative ratings at completion had an "alert" or "problem" classification in the DELTA at supervision, making the DELTA a good early warning system. However, given the DELTA's focus on outcomes, IDB Invest maintains projects in "Alert" or "Problem" until the factors that are affecting performance are solved (either because external conditions are improved or corrective measures are implemented). Consequently, the signals raised by the DELTA tended not to be resolved during execution.

C. Performance by sectors, sub-regions, and instruments (2018-2024)

OVE also compared the average overall outcome ratings of the last six validation cycles (2018-2024) to identify patterns of performance of relevant groups of IDB and IDB Invest operations. The analysis covered 741 validated operations for the period 2018-2014, including 447 for IDB and 294 for IDB Invest.

The proportion of IDB operations with positive ratings in overall outcomes during 2018-2024 differs across sectors and regions but not across lending instruments. Half or more of the operations validated in the 2018–2024 period in the Social (SCL), Infrastructure and Energy (INE), Institutions for Development (IFD), and Integration and Trade (INT) sectors achieved positive overall outcomes on average. The Climate Change and Sustainable Development (CSD) sector was an exception, with a 28 percent positive rating. By region, the Southern Cone (CSC) and Central America, Haiti, Mexico, Panama, and the Dominican Republic (CID) had average positive outcome rates above 55 percent, while the Andean Group (CAN) and the Caribbean (CCB) were below 48 percent. Differences across lending instruments are, on average, minor (57 percent of the 51 PBLs vs 54 percent of the 391 investment loans).

The proportion of IDB Invest operations with positive ratings in overall outcomes differs across business segments, regions, and between pre- and post-merge-out operations. Seventy-six percent of 58 operations validated in the 2018–2024 period in the Infrastructure and Energy segment (IE) achieved a positive overall outcome on average, versus 58% of the 93 operations of the Corporate (CO) and 50 percent of the 143 Financial Institutions (FI) operations. The proportion of operations with positive ratings in overall outcomes in the Andean Group (CAN) and the Southern Cone (CSC) was above 60%, in Central America, Haiti, Mexico, Panama, and the Dominican Republic (CID) was 55%, while in the Caribbean (CCB) was 29%. However, this result should be interpreted with caution because the data only includes a small number of operations in the Caribbean (7). Finally, the proportion in post-merge-out-operations was 14 percentage points higher than in pre-merge-out operations (66 percent of the 125 pre-merge-out operations vs 52 percent of the 169 operations post-merge-out operations).

D. Quality of self-evaluations and differences between Management’s and OVE’s ratings

Fifty-one percent of IDB’s PCRs (30 out of 59 PCRs assessed) were of satisfactory quality. PCRs of good quality comply with the guidelines, provide quality evidence and analysis, document changes, and are candid, among other aspects. Most shortcomings of the PCRs with negative quality (29) were related to the quality and completeness of the analyses in the effectiveness and efficiency sections.

Differences between IDB Management’s and OVE’s ratings remain large. While OVE rated the overall outcome positively in 47 percent of the validated operations, Management did so in 80 percent of them. Among the four core criteria, the discrepancies between OVE and Management have been largest in the efficiency and effectiveness ratings. Management is implementing an Action Plan (AP) to improve the overall quality of PCRs. OVE is collaborating with Management on activities included in the AP, including joint training initiatives for operational staff and the recent establishment of a working group to analyze and address key issues related to the low quality of efficiency analysis.

For the second year in a row, all validated XSRs were of satisfactory quality. XSRs in this cycle closely followed the guidelines, cited sufficient data, and were balanced and complete, among other aspects. The quality of XSRs has remained over 80 percent positive since 2019.

Ratings by IDB Invest Management and OVE have been close in the past few years. While OVE rated the overall outcome positively in 69 percent of operations in 2024, Management did so in 71 percent of them. The discrepancy in ratings for the overall outcome decreased over time. A similar pattern is observed for the four core criteria.

I. 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. It is based on management's self-evaluations and OVE's validations.** Every year, IDB Group Management prepares self-evaluations of Project Completion Reports (PCRs) for IDB projects and Expanded Supervision Reports (XSRs) for IDB Invest projects.¹ 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.² Every year, OVE reviews and validates the self-evaluations prepared by management. Self-evaluations and validations are intended to serve as accountability and learning tools. The accountability goal ensures that the IDB Group uses project resources effectively, efficiently, and sustainably to achieve approved objectives. The learning goal aims at providing lessons to inform the design and implementation of both ongoing and new projects.³

A. Evaluation methodology and the role of OVE

1.2 **The IDB Group's project evaluation and validation methodology is objectives-based.** The evaluation and validation methodology follows the *Project Completion Report (PCR): Principles and Guidelines* revised in 2020 (document [OP-1696-6](#)) for IDB (public sector operations), and the *XSR Guidelines for IDB Group Private Sector Operations*, revised in 2018 for IDB Invest private operations. Management and OVE conduct their assessments comparing results against the objectives stated at project approval. This implies that project performance is assessed against the objectives that they intended to achieve as defined in their design documents.

1.3 **Management's self-evaluations rely on core and non-core criteria.** In both PCRs and XSRs, project performance is measured against 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 specific objectives were achieved; and (iv) sustainability of the results achieved through the mitigation of risks and compliance with environmental and social safeguards. (Table 1.1). Besides the core criteria, the PCR and XSR guidelines require the inclusion of non-core criteria. For PCRs, these are the performance of the IDB and its counterparts (i.e., Borrowers), while for XSRs non-core criteria are the additionality of IDB Invest, the outcome of the investment, and IDB Invest work quality. Table 1.1 presents each core criteria dimension and Table 1.2 presents non-core criteria.

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

² Factors determining EOM vary by project type (see details in [XSR Guidelines](#), Annex II).

³ See 2020 PCR Guidelines (document [OP-1696-6, Annex 1](#)) and [2018 XSR Guidelines](#).

Table 1.1. Core criteria for management self-evaluations of project performance

IDB operations	Core Criteria	IDB Invest operations
<ul style="list-style-type: none"> • Alignment of project development objectives with country development needs. • Alignment of project development objectives with IDB Group country strategy (during approval, implementation, and closure). • Alignment of project design with country realities (during approval, implementation, and closure). • Alignment of project design with project development objectives. • The extent to which the project's objectives and design remain aligned with strategic priorities and development needs at approval, during execution, and at closure. 	Relevance	<ul style="list-style-type: none"> • Alignment of project objectives with country development needs. • Alignment of project development objectives with IDB Group country strategy and corporate goals. • Alignment of project design with country realities. • Alignment of project design with project development objectives. • The extent to which the project's objectives and design remain relevant at the time of the evaluation.
<ul style="list-style-type: none"> • Extent to which project achieved each stated development objective, given project outputs produced. 	Effectiveness	<ul style="list-style-type: none"> • Extent to which project achieved each stated development objective, given project outputs produced.
<ul style="list-style-type: none"> • Extent to which project benefits exceed project costs or extent to which project benefits were achieved at less than expected or at reasonable costs and no other factors affected efficiency. 	Efficiency	<ul style="list-style-type: none"> • Financial performance: project contribution to company financial results and extent to which project process and business objectives were achieved. • Economic performance: extent to which project's economic benefits exceed costs of capital; project effects on key economic stakeholders.
<ul style="list-style-type: none"> • Safeguards performance, • Assessment of risks to continuation of project development results. 	Sustainability	<ul style="list-style-type: none"> • Safeguards performance. • Assessment of risks to continuation of project results.
Project Development Outcome		

Source: OVE, based on 2020 PCR Guidelines (document [OP-1696-6](#)), and [2018 XSR Guidelines for IDB Group Private Sector Operations](#).

Table 1.2. Non-core criteria for management's self-evaluation of project performance

IDB operations	Non-core criteria	IDB Invest operations
<ul style="list-style-type: none"> • Bank performance in ensuring quality at entry: the extent to which the Bank identified, facilitated preparation of, and approved the operation such that it was most likely to achieve planned specific development objectives. • Quality of Bank supervision: extent to which the Bank proactively identified and resolved threats to the achievement of relevant specific development objectives. 	Bank performance (IDB)/ Work quality (IDB Invest)	<ul style="list-style-type: none"> • IDB Group's work quality, taking into consideration precommitment work at entry, i.e. screening, appraisal, and structuring/underwriting. • IDB Group's work quality, taking into consideration its monitoring and supervision of the operation following commitment or guarantee issuance.

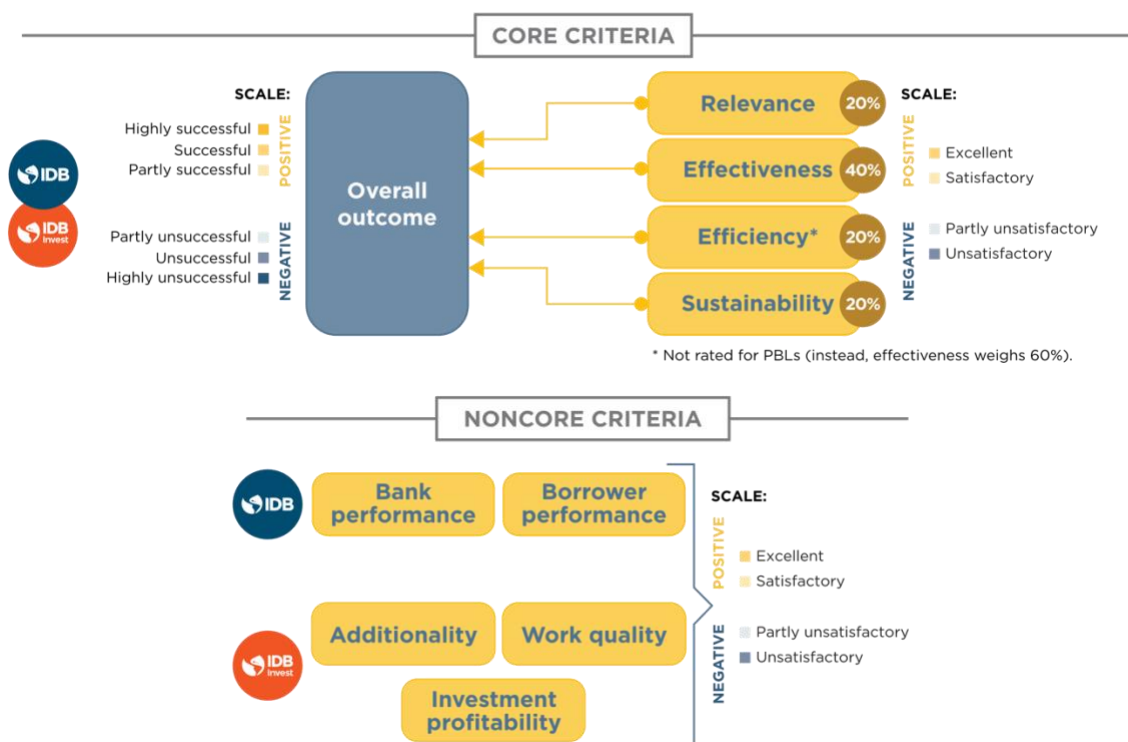
<ul style="list-style-type: none"> • Extent to which the Borrower ensured the quality of preparation and implementation of the operation, and complied with covenants and agreements, toward the achievement of development results. 	<p>Borrower performance (only IDB)</p>	
	<p>Additionality (only IDB Invest)</p>	<ul style="list-style-type: none"> • The assessment and rating are based on the counterfactual assessment of how the project would have (or would not have) proceeded without IDB Group's support. Ratings should be based on financial and nonfinancial additionalities achievements.
	<p>Investment profitability (only IDB Invest)</p>	<ul style="list-style-type: none"> • Investment's gross profit contribution (net of financing costs and loss provisions but before deducting administrative costs) with rating benchmarks set in relation to corresponding at-approval standards for minimally expected performance.

Source: OVE, based on 2020 PCR Guidelines (document [OP-1696-5](#)) and [2018 XSR Guidelines for IDB Group Private Sector Operations](#).

1.4 **Management assigns an overall rating to each project, which is the weighted average of the ratings of the four core criteria.** Following the PCR and XSR Guidelines, Management rates each core and non-core criterion on a four-point scale, ranging from “unsatisfactory” to “excellent”, and then assigns the overall outcome rating of each project, which is the weighted average of the ratings of the four core criteria. The weight of the effectiveness criterion is 40 percent, while the weights of relevance, efficiency and sustainability are 20 percent each.⁴ 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. Non-core criteria are rated on the same four-point scale as the core criteria, but do not count toward the project's overall outcome rating. Figure 1.1 provides an overview of the evaluation criteria and their respective rating scales.

⁴ For IDB's policy-based loans (PBLs), efficiency is not rated, and effectiveness weighs 60% instead.

Figure 1.1. Evaluation criteria and rating scales



Source: OVE, based on PCR and XSR Guidelines.

1.5 **OVE validates all PCRs and XSRs prepared by IDB and IDB Invest.** Once Management has submitted the PCRs and XSRs, OVE reviews them to ensure that they are substantiated by evidence and prepared in accordance with the PCR and XSR guidelines.⁵ OVE's validation process involves a review of project documentation, including approval documents, results matrices, and monitoring reports. Based on the PCR and XSR guidelines, so using the same methodology applied by Management, OVE assigns its own ratings (to core and non-core criteria and to overall outcomes) and prepares validation notes with its assessment for each validated operation. The IDB Group considers OVE's ratings final and uses them for corporate reporting in the IDB Group's Development Effectiveness Overview (DEO⁶ OVE also rates the quality of the self-evaluation reports (PCRs and XSRs) by Management on a four-point scale (excellent, good, fair, poor). The assessment covers adherence to relevant guidelines, completeness of the analyses, quality of evidence provided and quality of the analyses, documenting changes in the design and results matrix, candor, and lessons based on evidence. Table 1.3. presents the dimensions analyzed in the quality assessment conducted by OVE.

⁵ See 2020 PCR Guidelines (document [OP-1696-6, Annex 1](#)) and [2018 XSR Guidelines](#).

⁶ The DEO is an annual report by the IDB Group that shares the results of its work in the region.

Table 1.3. Dimensions analyzed in the analysis of quality of self-evaluations of project performance

Quality of self-evaluations	Dimensions analyzed
<p style="text-align: center;">PCR Quality (IDB) / XSR Quality (IDB Invest)</p> <p>Scale: Positive (Excellent, Good); Negative (Fair, Poor)</p>	<ul style="list-style-type: none"> • The extent to which the self-evaluations cover all criteria of project performance assessment and adhere closely to the Guidelines. • The extent to which the self-evaluation cites sufficient quantitative and qualitative data to support findings without redundancy or repetition. • The extent to which the self-evaluation captures the project’s results, discussing the achievement of individual objectives. • The clarity, consistency, and candidness of self-evaluation. • The consistency of the ratings with the provided evidence. • The extent to which self-evaluation identifies lessons based on the analysis and evidence provided.

Source: OVE.

B. Scope of this report

1.6 **This report presents the results of 105 operations validated by OVE in 2024.** Most validated PCRs and XSRs assessed a single project, although some covered a group of related projects with shared objectives.⁷ Throughout this report, the more general term *operation* is used to refer to both IDB and IDB Invest projects and to both validated projects and programs. Annex I lists all validated operations and their associated ratings. Annex II provides OVE’s validation notes for IDB operations only, as IDB Invest validation notes are confidential.

1.7 **This report includes five sections.** Following this Introduction, section II presents the 2024 results for the IDB Group and the factors that influenced the results. Section III briefly discusses IDB Group performance by sectors, sub-regions and instruments during 2018-2014. Section IV addresses the quality of PCRs and XSRs, and Section V presents overall conclusions.

II. RESULTS FOR IDB AND IDB INVEST OPERATIONS VALIDATED IN 2024

2.1 **This section presents the results for IDB and IDB Invest operations validated in 2024.** It presents ratings for overall outcomes, core criteria, and noncore criteria, and it includes an analysis of the main factors explaining project performance based on the information presented in the PCRs, XSRs, and OVE’s validation notes.⁸ 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. This section also presents some methodological issues that may affect Management’s self-assessments and OVE’s validations. The identification of these issues is based on specific examples of validated operations in 2024 and findings of other recent OVE’s work, including the Development Effectiveness Framework evaluation (document [RE-583-2](#)) and the Citizen Security evaluation (document [RE-560-1](#)). This is not a comprehensive

⁷ 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.

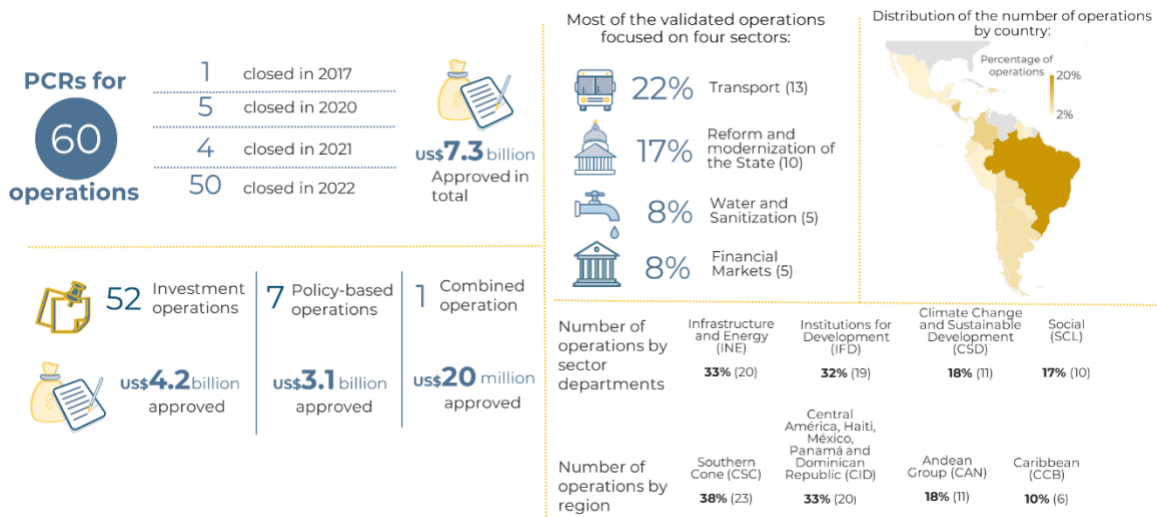
⁸ Such factors were derived using content analysis with computer-assisted qualitative data analysis software (CAQDAS). This method was applied to determine the presence of certain topics within OVE’s validation notes, the PCRs, and the XSRs, where the frequency of such topics is considered indicative of the main drivers of performance.

analysis of potential methodological issues or a full review of the PCR and XSR guidelines. OVE will continue to analyze these methodological issues – and other issues that may emerge in the future – in the context of its Development Effectiveness line of work.

A. IDB results in 2024

2.2 **In 2024, OVE validated the PCRs of 60 IDB operations.** The cohort of IDB operations validated in 2024 includes 51 Investment Loans (INV), 7 Policy-Based Loans (PBLs), 1 combined operation (INV and PBL), and 1 investment grant. Eighty-three percent of these operations (50) closed in 2022, while the remaining ones closed between 2017 and 2021. All validated operations were approved between 2009 and 2021, with an important concentration between 2013 and 2016 (68 percent, 41 operations). Most of the validated operations (52 percent) were focused on four sectors: transport (22 percent, 13 operations), reform and modernization of the State (17 percent, 10 operations), water and sanitation, and financial markets (8 percent, 5 operations each). The validated operations benefited 21 of the 26 IDB’s borrowing countries, with almost half of these operations focused on four countries: Brazil (20%, 12 operations), Nicaragua (10 percent, 6 operations), Colombia (8 percent, 5 operations), and Haiti (7 percent, 4 operations) (See Figure 2.1).

Figure 2.1. Characteristics of IDB operations with PCRs validated in 2024



Source: OVE, with data from IDB (2024).

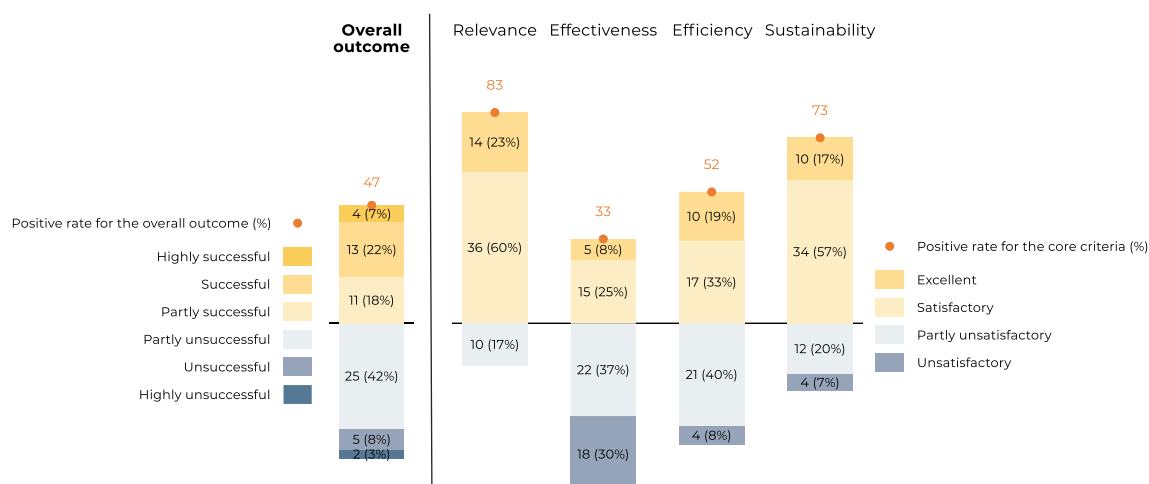
Note: Investment operations include all modalities of investment loans and investment grants.

1. Overall outcome ratings

2.3 **Of the 60 IDB operations validated in 2024, 28 (47 percent) achieved a positive overall outcome rating, with variations across different sectors and regions.** This rate falls short of the 70 percent target set in the 2020–2023 Corporate Results Framework (CRF). Fifty-four operations (90 percent of the total) fell in the middle four ratings (“unsuccessful,” “partly unsuccessful,” “partly successful,” and “successful”). Relevance was the highest-rated core criterion, with 83 percent of operations rated positive, followed by sustainability (73 percent), efficiency (52 percent), and effectiveness (33 percent). (Figure 2.2). The proportion of projects with positive overall outcome ratings in 2024 differs across sectors and

regions but not across lending instruments. However, these comparisons should be interpreted with caution because of the small number of operations in certain groups.⁹ While 70 percent of the 10 operations in the Social Sector (SCL), 58 percent of the 19 operations in the Institutions for Development (IFD), and 50 percent of the 20 Infrastructure and Energy (INE) validated operations had a positive overall outcome rating, all the operations of the Climate Change and Sustainable Development sector (11) had a negative rating. There are differences also by regions. Sixty percent of the 20 validated operations of Central America, Haiti, Mexico, Panama, and the Dominican Republic (CID), 50 percent of the 6 operations of the Caribbean (CCB), 44 percent of the 23 operations in the Southern Cone (CSC) and 27 percent of the 11 operations in the Andean Group (CAN) were rated positive. The overall outcome ratings are similar across instruments: 57 percent of the 7 PBLs versus 54 percent of the 52 INV.

Figure 2.2. IDB operations ratings, 2024



Source: OVE.

Note: $N = 60$ operations for all criteria except for efficiency ($N = 52$) because seven policy-based operations and one combined PBL and INV operation are not rated for this criterion. Percentages may not add up to 100 due to rounding.

2. Relevance

2.4 **The relevance of an operation is rated positively when its development objectives and design align with the IDB's strategic priorities and the country's development priorities throughout the operation's lifecycle.** Key relevance evaluation criteria include the alignment of project objectives with country development needs and priorities at approval, during execution, and at closure, as well as the alignment of project design with country realities and project development objectives (Table 1.1). Relevance also considers the operation's vertical logic (how well the chosen outputs and activities lead to the intended outcomes) at approval and throughout implementation.¹⁰ If adjustments to the project design occur during implementation, such as canceled outputs, or other

⁹ Section III briefly discusses IDB performance by sectors, sub-regions and instruments using a larger number of validated operations during 2018-2014.

¹⁰ 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.

changes in outputs, the PCR analysis must discuss the underlying reasons for the changes and provide a robust justification for the project's continued relevance to the country's development challenges.

- 2.5 **OVE rated the relevance of 50 out of 60 validated operations (83 percent) positive.** The objectives of the 50 relevant operations were aligned with country development needs, government priorities, and IDB Group Country Strategies (CSs). The objectives addressed key development challenges at approval and remained consistent with the sector-specific objectives outlined in the CS, as reflected in the CS results matrix, throughout the lifecycle of the project. These operations also had adequate project designs with clear, logical relationships between proposed activities, expected results, and objectives (sound vertical logic). The design of the 50 operations was well-suited to the local context, needs, and capacity and changes during implementation did not affect their vertical logic. These 50 operations benefited 21 countries, including Brazil (12 operations), Nicaragua (6), and Bolivia (3). These 50 operations were in 14 sectors. The sectors with the highest percentage of operations with positive relevance ratings include modernization of the state (10 operations out of 10), education (4 operations out of 4), health (4 operations out of 4), transport (11 operations out of 13), and financial markets (4 operations out of 5). This sector disaggregation is illustrative and should be interpreted with caution because of the small number of operations in certain sectors. Box 2.1 provides an example of a project with a positive relevance rating (2.1.a).
- 2.6 **Ten out of 60 validated operations had negative relevance ratings.** OVE rated negatively 10 operations due to a combination of four factors (not mutually exclusive) that are summarized below. Box 2.1 provides an example of a project with a negative relevance rating (2.1.b).
- i. All 10 operations negatively rated exhibited weak vertical logic. Seven operations had vertical logic shortcomings both at approval and during implementation, one only at approval and two only during implementation. Cancellations and other changes in projects were the main reasons affecting the vertical logic during implementation (9 operations). In two cases, nearly half of the resources were canceled. In another case, the Board approved a modification of the operation, which resulted in

reduced resources and the cancellation of outputs that were important to achieve most of the project's objectives that remained after the modification. For other six operations, key activities were removed.

- ii. In seven operations (70 percent), project design did not adequately account for country realities. Shortcomings included implementation arrangements that overlooked the executing agency's technical, managerial, and operational capacities; insufficient institutional coordination; and insufficient analysis of critical contextual factors at design that affected project implementation.
- iii. In four operations (40 percent), project objectives were not aligned with the CSs in place at the time of project approval, during implementation or at project closure. Of these, one was not aligned with the valid CS at the time of project approval, two were misaligned with the CS during the project's implementation, and the other one was not aligned at any stage—approval, implementation, or closure.
- iv. Two operations (20 percent) fell out of alignment with government priorities during execution due to strategic shifts by new authorities.

2.7 **The relevance ratings of some projects may have been affected by changes in circumstances that could not be anticipated at approval or easily addressed during implementation or after project closure.** Unforeseen circumstances such as the COVID-19 crisis or shifts in national priorities following changes in government, may have a negative impact on projects' relevance ratings. Since ratings, both for self-assessment and validation, are objectives-based as per the PCR guidelines, projects that were not restructured to adapt their objectives to changing circumstances, are rated based on their original objectives. Consequently, projects may receive negative ratings even if they financed activities and achieved results that were relevant during implementation. A 2024 OVE evaluation of the Development Effectiveness Framework (document [RE-583-2](#)) found that restructuring IDB projects is a cumbersome process, based on a benchmarking exercise that compared IDB's policies with those of other Multilateral Development Banks. While the decision to restructure a project ultimately rests with borrowing countries, internal complexity may lead IDB staff to adapt to changing circumstances while maintaining projects' original objectives. Box 2.1 provides an example of a project with a negative relevance rating affected by changes in circumstances that could not be foreseen at the time of project approval and that were not addressed through restructuring. (2.1.c.).

Box 2.1. Examples of relevance assessments

a) Argentina Public Health Services

A US\$150 million investment loan (AR-L1312), designed as a medium-term operation under a Conditional Credit Line for Investment Projects (CCLIP), aimed to improve the accessibility and coverage of public health services in Buenos Aires. The program sought to address important gaps in access and coverage of primary and emergency services in the province. Management and OVE both rated the project's relevance as Excellent. The program was aligned with the IDB Group CSs with Argentina in place at the time of project approval, during implementation, and at project closure, which prioritized strengthening the quality of health services for prevention and early detection of chronic non-communicable diseases. The program was not restructured. However, during the kickoff workshop, it was adjusted to address new priorities in response to the COVID-19 pandemic. The adjustments included changes that did not require restructuring the project, including the incorporation of new COVID-19-related indicators, accelerated execution, and new activities—investments in infrastructure and equipment necessary to handle the increased demand for emergency and intensive care services—to support the capacity for detection and treatment of COVID-19 cases within the framework of strengthening emergency services.^a

b) Colombia Public Transportation

A US\$320 million investment loan in Colombia (CO-L1091) aimed to modernize public transportation in four mid-size cities, improving bus service for passengers, mitigating climate change by reducing emissions, and contributing to the development of competitive and sustainable cities that would offer citizens safe mobility opportunities. The operation included components for (1) transport infrastructure, (2) fleet management and control systems, and (3) land acquisition, compensation plans, and resettlement of affected populations. Management rated the project's relevance Unsatisfactory, and OVE upgraded it to Partially Unsatisfactory because it was well aligned with the country's development needs and the Bank's strategies during the approval and implementation phases. A 42-month delay in project implementation resulted from issues that were largely related to the design of the project (such as a lack of final designs, which generated differences between the initial and final budgets of the project, and not providing clarity on responsibilities for financing the location and relocation of utility networks). Moreover, the project had deficiencies in its vertical logic. For example, the results matrix did not include results related to the construction or adaptation of yards and maintenance workshops envisioned in the loan proposal. In addition, the program performance indicators did not report results in the manner set out in the results matrix. Finally, cancellation of 55.8 percent of the loan amount arose largely because the project design did not sufficiently consider the realities of the country, particularly its implementation capacity. Examples included difficulties in pre-contractual and procurement processes carried out by city personnel without previous experience in similar projects. Ultimately, none of the cities fully implemented the envisioned transportation systems^b

c) Uruguay Tourism Development

This first US\$5 million operation, under a conditional credit line (CCLIP), aimed to boost tourism spending in the five departments comprising the Uruguay River Corridor, enhance natural capital management, expand tourism supply, and strengthen subnational tourism governance. The operation included components for (1) investments for the development of public tourist attractions to promote nautical, cultural, and ecotourism (like floating docks, ramps, a marina, and a visitor center); (2) technical and financial support for local entrepreneurship through a competitive fund; and (3) a tourism observatory to strengthen subnational tourism governance in the Corridor. Management rated relevance as Excellent, while OVE rated it as Partly Unsatisfactory. The project was aligned with the country's needs and the causal alignment between activities, outputs and outcomes was adequate in the project design. However, the cancellation of almost half of the resources of the program (48 percent) resulted in the cancellation of an important set of project activities and outputs that reduced the scope of the project, weakening its vertical logic. The early cancellation of funds was requested by the Ministry of Tourism under a new administration in 2020. The decision was driven by budget constraints, shifts in priorities following the change in government, and significant economic challenges in the region. The Uruguay River Corridor was in a very early development stage and faced challenges from Uruguay's limited nautical tradition and strong competition from Argentina's developed nautical tourism. The pandemic crisis further disrupted foreign tourist flows, exacerbated budgetary constraints, and prompted the shift in the government's strategic priorities. The unforeseen circumstances of the COVID-19 crisis and shifts in national priorities, not addressed through restructuring of the project (or adjustments in the result matrix), had a negative impact on OVE's rating of the project's relevance.^c

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

Note: ^a Program for the Strengthening and Integration of Health Networks in the Province of Buenos Aires (AR-L1312)

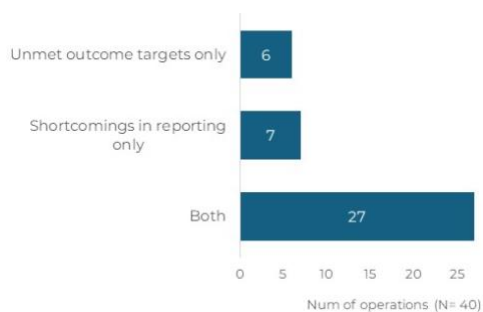
^b Strategic Public Transport Systems Program in Colombia (CO-L1091). ^c Tourism Corridor Development Program (UR-L1113).

3. Effectiveness

- 2.8 **Project effectiveness is rated positive when more than 50 percent of the project objectives are achieved.** To be rated positive, more than 50 percent of a project's specific objectives must be largely or fully achieved, with no specific objectives rated Unsatisfactory, and results must be attributable to the project. Project effectiveness is rated negative when these conditions do not apply—for example, when a specific objective is rated Unsatisfactory, when 50 percent or more of the project's specific objectives were partly or not achieved, and when a project's specific objectives have only a combination of Partly Satisfactory and Unsatisfactory ratings.¹¹
- 2.9 **Negative ratings are due to a combination of two factors: unmet outcome targets and shortcomings in reporting.** Unmet outcome targets are broken down into two key determinants: *output shortfalls* and *outcome shortfalls*. Output shortfalls occur when the project does not deliver the planned outputs (or not to their full extent) during its execution phase or when issues arise with the delivered outputs, such as poor quality and non-operational outputs. Outcome shortfalls occur when the project delivers on the planned outputs but not on the expected outcomes. The second factor, shortcomings in reporting, occurs when the evidence to demonstrate the achievement of project objectives is insufficient. These shortcomings occur when one or more of the following issues arise: (i) data are unavailable because they were not collected; (ii) data are available but are inadequate, insufficient to measure all dimensions of project objectives, or unreliable; and (iii) data are available but the progress shown on outcome indicators cannot be plausibly linked to the project.
- 2.10 **OVE rated effectiveness positive for 20 out of 60 validated operations (33 percent).** For these 20 operations, most objectives were achieved. For an objective to be considered achieved, the targets of outcome indicators must be largely met (at least 80 percent on average) and the observed results must be linked to the outputs delivered by the project. These 20 operations benefitted 11 different countries, including Nicaragua (5 operations), Argentina (2), Brazil (2), Colombia (2), the Dominican Republic (2), and Honduras (2). These 20 operations were in 9 sectors. The sectors with the highest percentage of operations with positive effectiveness ratings include private firms and SME development (3 operations out of 3), financial markets (4 operations out of 5), and transport (6 operations out of 13). This sector disaggregation is illustrative and should be interpreted with caution because of the small number of operations in certain sectors. Box 2.2 provides an example of a project rated positive on effectiveness (2.2.a). Forty out of the 60 validated operations (66 percent) were rated negative, mostly due to a combination of unmet outcome targets and shortcomings in reporting. Of the 40 operations with negative effectiveness ratings, 27 (68 percent) suffered from the combination of both unmet outcome targets and shortcomings in reporting, while 7 (18 percent) were due exclusively to shortcomings in reporting, and 6 (15 percent) were due only to unmet outcome targets (see Figure 2.3). Box 2.2 provides an example of a project with a negative effectiveness rating (2.2.b).

¹¹ Proposed Adjustments to the Development Effectiveness Framework Tools for Sovereign Guaranteed Operations. Approved version. Annex 1: Project Completion Report Principles and Guidelines (OP-1696-5).

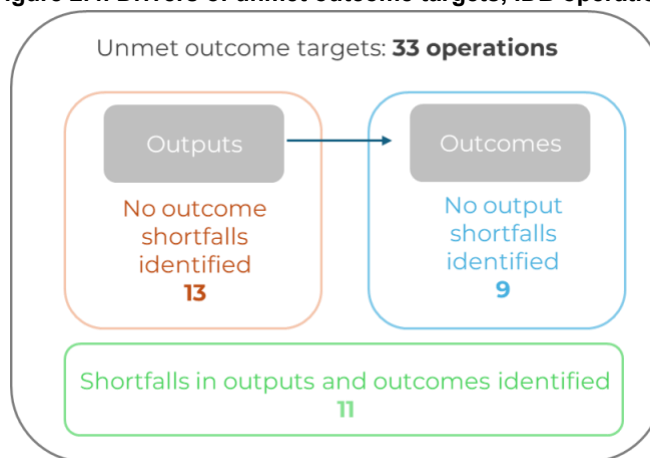
Figure 2.3. Factors that affected negative effectiveness ratings, IDB operations



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

2.11 **Thirty-three out of 40 operations with negative effectiveness ratings did not achieve their outcome targets.** A breakdown of the 33 operations with unmet outcome targets shows that output shortfalls were more frequent than outcome shortfalls (see Figure 2.4). Specifically, 13 operations (39 percent) experienced exclusively output shortfalls, 9 (27 percent) faced only outcome shortfalls, and 11 (33 percent) faced both output and outcome shortfalls. While in all these cases the expected outcome targets were not (fully) met, the distinction is helpful in determining where the causal chain broke down along the operation's vertical logic.

Figure 2.4. Drivers of unmet outcome targets, IDB operations



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

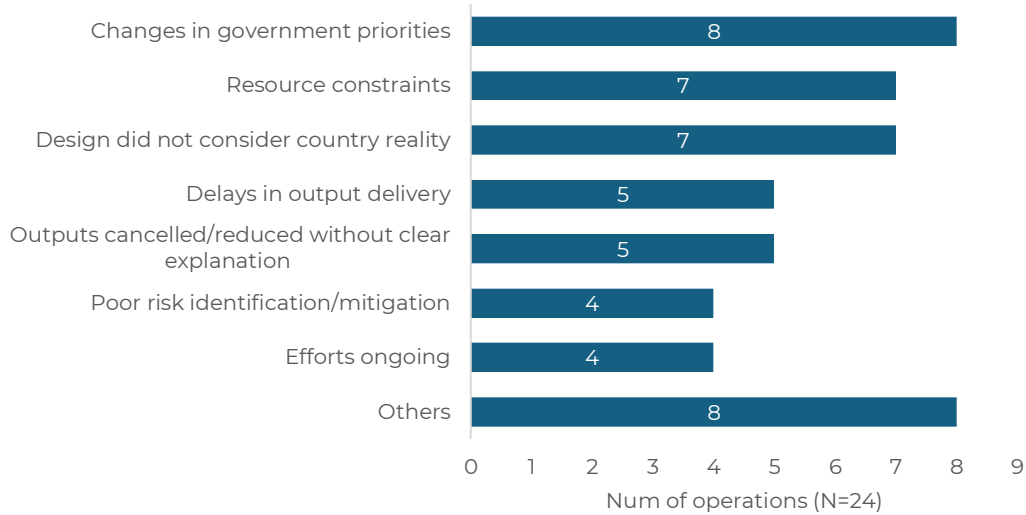
2.12 **Operations that experienced output shortfalls faced various challenges.** Among 24 operations that experienced output shortfalls, eight did not deliver the expected outputs or did not adjust them due to changes in governments' priorities. Seven operations faced resource constraints, with three being severely affected,¹² having over 50 percent of their financial resources canceled.¹³ Other frequent

¹² The other 4 projects did not experience such cancellations. Their canceled amount ranged from 0% to 5%. In one case, the local counterpart was reduced, and in another, co-financing never materialized.

¹³ On average, projects with positive effectiveness ratings experienced a cancellation of 1% of their approved amount while projects with negative effectiveness ratings experienced an average cancellation of 7%.

output shortcomings related to project designs not adequately accounting for the country's reality, delays in output delivery, and unclear explanations of output shortfalls. (See Figure 2.5).

Figure 2.5. Factors that affected output shortfalls, IDB operations

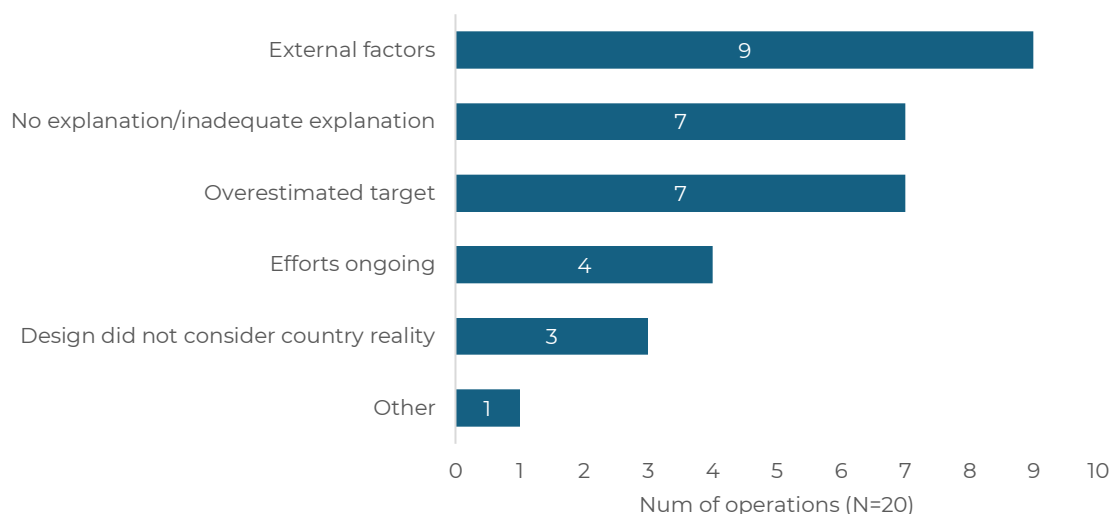


Source: OVE's categorization, based on PCR contents and OVE's validation notes.

Notes: Numbers do not add up to 24 because reasons are not mutually exclusive. The "others" category comprises eight projects, each facing one of six distinct issues, with no more than two projects sharing the same issue.

2.13 The most common reason for unmet outcome targets not related to output shortfalls was external shocks. Among 20 operations that experienced outcome shortfalls, nine were affected by external factors, which were primarily related to the COVID-19 pandemic that affected outcomes in areas such as transportation and other infrastructure, tourism, and fiscal management. Some countries also experienced social unrest, civil disturbances, and macroeconomic instability, further affecting project outcomes. As mentioned in the relevance section, such factors cannot be foreseen at project design and cannot be easily factored in during project implementation and may result in negative effectiveness ratings. In addition, there were seven cases where the reason for lower-than-expected outcomes was unclear, and another seven where the targets were overestimated. In four cases, the expected changes at the outcome level had not yet materialized but were expected to do so in the near future (See Figure 2.6).

Figure 2.6. Factors that affected outcome shortfalls, IDB operations



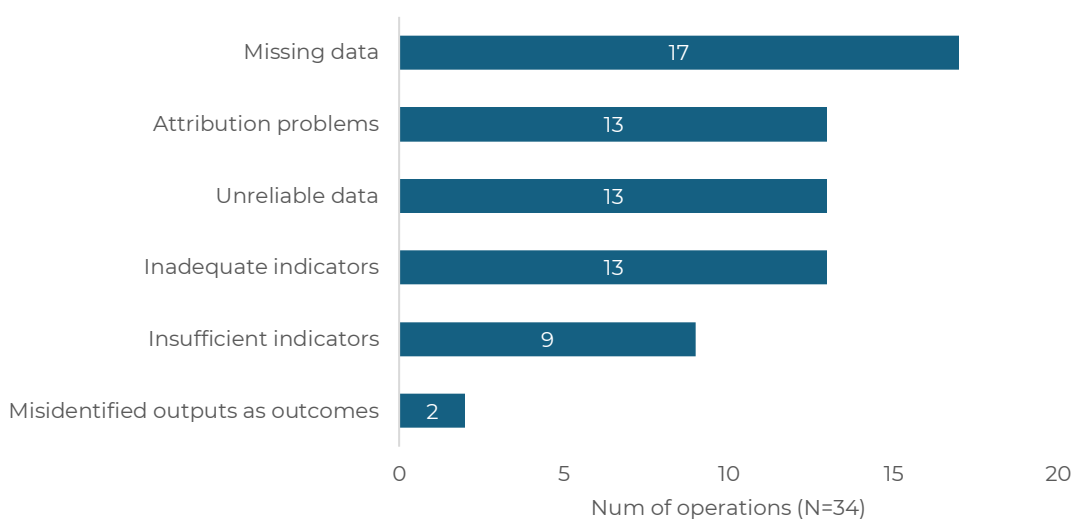
Source: OVE's categorization, based on PCR contents and OVE's validation notes.

Notes: Numbers do not add up to 20 because reasons are not mutually exclusive.

2.14 The most common reasons for shortcomings in reporting were missing data.

Seventeen of the 34 operations with reporting shortcomings had missing data. Ten of the 17 operations did not include updated values for one or more outcome indicators. For four of the 17 operations, the original outcome indicators were not reported though OVE considered that the original indicator should have been retained. In these cases, the original outcome indicators were not measured and were therefore assigned a value of zero when calculating the achievement ratio for the specific objective. Finally, three of the 17 operations provided outcome information only for a subset of interventions. Other frequent issues were attribution problems, unreliable data, and inadequate or insufficient indicators (See Figure 2.7).

Figure 2.7. Factors that affected reporting shortcomings, IDB operations



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

Note: Numbers do not add up to 34 because the reporting shortcomings are not mutually exclusive.

- 2.15 **In some cases, shortcomings in reporting (e.g. lack of data, attribution problems, inadequate indicators) can obscure a project's achievements.** Providing accurate targets at project design and collecting and updating data at the time of project implementation and closure is essential to be able to assess project effectiveness. However, when reporting shortcomings impede the assessment of achievement levels and result in a negative rating, tagging a project as ineffective could be misleading, because neither Management nor OVE can state whether the project has achieved (or not) its outcomes. OVE's evaluation of the Bank's support to Citizen Security (document [RE-560-1](#)) illustrates this point as it finds that while most closed projects had negative effectiveness ratings due to shortcomings in reporting to demonstrate achievement of their original objectives, they did contribute to achieve important outcomes.
- 2.16 **Management incorporated additional indicators to evaluate the effectiveness of more than half of the operations.** The 2020 PCR guidelines¹⁴ allow for incorporating outcome indicators between approval and the start-up plan, and during execution, to strengthen the empirical evidence base behind the (potential) achievement of objectives. The PCR Guidelines also state that if specific development objectives were approved without associated metrics, the PCR (i.e., at project closure) should try to reconstruct a measurement of those specific objectives to attempt to ascertain the results of the intervention. Accordingly, OVE reviewed the added indicators to determine whether they should be considered in evaluating project effectiveness. In this cycle, 55 percent (33) of the validated operations included one or more additional indicators (16 indicators during execution and 74 at closure). OVE accepted 13 of the 16 indicators (81 percent) added during project execution and 32 of 74 (43 percent) indicators added after

¹⁴ Proposed Adjustments to the Development Effectiveness Framework Tools for Sovereign Guaranteed Operations. Approved version. Annex 1: Project Completion Report Principles and Guidelines (OP-1696-5)

the execution was completed. OVE issued a clarification during the 2024 cycle, explaining the specific circumstances when outcome indicators can be added after 95 percent of resources have been disbursed (i.e., at closure). The clarification specifies that new outcome indicators can be added at closure (i.e., after 95 percent of the resources have been disbursed) only in the following cases: i) the added indicator is a proxy for an existing indicator that was not measured or whose source of information was discontinued, or ii) there were no outcome indicators to measure a specific objective of the project. The addition of a high number of outcome indicators close to projects' closure could be due to (among other factors), an insufficient focus on monitoring results during execution.

Box 2.2. Examples of effectiveness assessments

a) Nicaragua Atlantic Coast Road Connectivity Project

The Atlantic Coast Road Connectivity Project (NI-L1087) is an example of a project rated positively on effectiveness. This was a US\$61.5 million investment loan implemented by the Ministry of Transport and Infrastructure in Nicaragua.^a The project aimed to achieve two specific objectives: (i) reduce travel time and (ii) lower transportation costs along the La Gateada–Naciones Unidas–Bluefields corridor. Key activities included paving, constructing drainage systems, investing in road signs, and installing protective barriers, among others. The project received an Excellent effectiveness rating by both Management and OVE, meeting both objectives and achieving all associated outcome indicators, including improving road surface quality and reducing travel times and transportation costs.¹⁵

b) Brazil program for accelerating progress of education in Amazonas

The program for accelerating progress of education in Amazonas, Brazil (BR-L1328)^b is a US\$151 million investment loan with three specific objectives: (1) expanding the coverage of basic education;¹⁶ (2) improving the quality of basic education; and (3) increasing the management and monitoring capacity of the State Department of Education in the state of Amazonas. Activities included (1) constructing, adapting, or equipping 592 schools, and education centers; (2) implementing a learning acceleration and educational reinforcement program for 160,000 school students, and teacher-oriented programs and (3) redesigning administrative processes. Management and OVE rated the overall project as Partially Successful. Output shortfalls included adapting 3 of 20 planned extended education units, training 3,994 of 15,500 planned teachers, and cancelling programs for learning acceleration and teacher coaching that had aimed to serve 80,000 students and 135 schools respectively. The shortfalls were caused by a variety of factors, including delays due to COVID-19, changes in government priorities, and reallocation of resources to the health emergency. The project experienced two cancellations reaching 38.6 percent of the total approved amount, including counterpart financing. Still, the number of students attending school went up, academic performance improved in several subjects, and the integrated system for monitoring attendance and grades was implemented in all the State's Department of Education schools, which contributed to the overall project rating of Partially Successful.

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

Note: ^a Atlantic Coast Road Connectivity (NI-L1087). ^b Amazonas Education Development Acceleration Program (BR-L1328).

4. Efficiency

2.17 **The assessment of efficiency investigates the extent to which the costs incurred in achieving project results were reasonable.** The efficiency analysis applies to all investment operations and does not apply to policy-based loans. According to the PCR guidelines, to demonstrate how efficiently the project has used its resources, management should conduct either a cost-benefit analysis (CBA) or a cost-effectiveness analysis (CEA). In exceptional circumstances where

¹⁵ These outcome indicators were tracked through field measurements and modeling using HDM-4, a specialized software tool for transport projects. The results were considered attributable to the investments made.

¹⁶ Basic education includes early childhood education (ages 0 to 5 years), elementary (ages 6 to 14), and secondary education (ages 15 to 17).

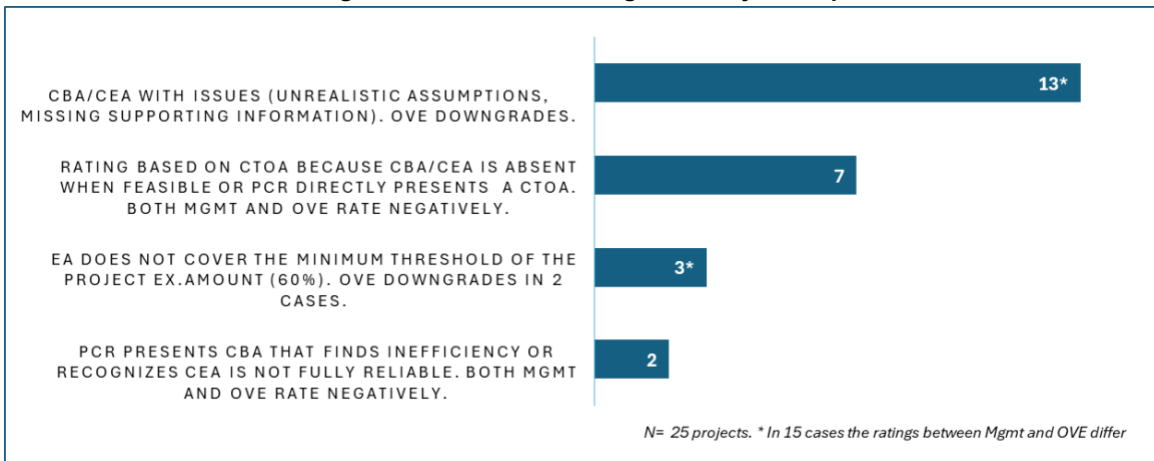
such an analysis is not feasible, management needs to provide a justification in the PCR and calculate the rating based on a cost and time overrun analysis (CTOA). In such cases, a maximum rating of satisfactory applies if the project performance ratings during implementation (PMR ratings) remained satisfactory. If OVE finds no clear justification for the absence of a CBA or CEA, OVE should assign a maximum of partially unsatisfactory if the PMR ratings of the project remained satisfactory during implementation. The PCR guidelines also prescribe that to be accepted, the economic analyses must present reasonable assumptions, data sources and a replicable methodology. If this is not the case, OVE can dismiss the analysis and calculate the efficiency rating based on the CTOA, assigning a maximum rating of partially unsatisfactory when the project did not have any “problem” or “alert” classification during its execution years. When the economic analysis is accepted by OVE, the rating would be positive if the economic rate of return (in case of CBA) is above 85 percent of the reference discount rate used (typically, 12 percent). If a CEA was conducted, the rating would be positive if the project results were achieved at a cost to or no more than 110 percent of the cost of alternative ways to achieve the results of the comparators.

- 2.18 **Of the 52 validated investment operations, 27 (52 percent) achieved a positive efficiency rating.** The 27 operations rated positive in efficiency included economic analyses supported by reasonable assumptions, data sources, and clear and reproducible methodologies. They also covered a significant amount of the project’s resources. Costs and benefits included in the analysis aligned well with the information provided in other sections of the PCR report. The magnitude of the benefits was consistent with that of the benefits presented in the effectiveness section and the benefits reflected well the specific objectives of the project. Also, the benefits were consistent with the scope of the projects and applied to the beneficiary population or the geographical areas where the projects intervened. The selected discount rates (opportunity cost of financing) aligned well with the literature and were justified. In sum, OVE considered the analysis of these positively rated operations reliable and the results of standard efficiency measures—such as economic rate of return (ERR) or benefit-cost ratio (BCR)—exceeded the thresholds defined in the PCR Guidelines. The 27 operations rated positive in efficiency benefited 12 different countries, including Brazil (7 operations), Nicaragua (6) and Dominican Republic (3). These 27 operations were in 13 sectors. The sectors with the highest percentage of operations with positive efficiency ratings include financial markets (4 operations out of 5), water and sanitation (4 operations out of 5), and transport (5 operations out of 13). This sector disaggregation is illustrative and should be interpreted with caution because of the small number of operations in certain sectors. Box 2.3 provides an example of a project with a positive efficiency rating (2.3.a).
- 2.19 **The efficiency ratings of the 25 operations rated negative are mostly the result of PCRs lacking a cost-benefit or cost-effectiveness analysis when considered feasible or presenting analyses with caveats that made the calculations of economic rates of return unreliable.** Of the 25 operations with negative efficiency ratings (Figure 2.8), 13 had an economic analysis that OVE dismissed due to unrealistic assumptions or missing supporting information. Seven operations were rated negatively by Management and OVE based on a cost and time overrun analysis (CTOA), since the economic analysis was absent without valid justification or because the PCR directly presented a CTOA. In three other cases, the cost-benefit or cost-effectiveness analysis did not cover the minimum

amount of project resources (60 percent) required to assess efficiency, and OVE downgraded two of them, while for the third one, the self-evaluation rated it negatively as well. The remaining two operations that were rated negative by Management and OVE either presented credible analyses that found that resources had not been used efficiently or recognized that the analysis presented could not be fully relied upon. Box 2.3 provides an example of a project with a negative efficiency rating (2.3.b).

- 2.20 **Due to the complexity of some projects in certain sectors, conducting a CBA or CEA is not always appropriate.** Conducting a CBA or CEA may not be appropriate in projects with multiple components where many (unrealistic) assumptions need to be made about multiple benefit streams that are often difficult to measure or predict. Following the PCR guidelines, in such cases, Management could provide a justification for not conducting these analyses and resort to a CTOA. This could result in a satisfactory rating when the project classification performance remained satisfactory during implementation and no other factors affected efficiency. If no justification is provided or if Management attempts to conduct a CBA or CEA with unrealistic assumptions, the maximum rating allowed by the PCR Guidelines is partially unsatisfactory. The PCR Guidelines are relatively strict on efficiency analysis requirements in comparison to MDB peers. The World Bank, for example, provides guidance on the aspects that should (ideally) be covered in the efficiency analyses that are part of project self-evaluations but allows Management some flexibility in deciding whether or not a CBA or CEA is appropriate and feasible. Box 2.3 provides an example of a project where a CBA was conducted (with many assumptions) while an exemption would have been justified (2.3.c).

Figure 2.8. Factors affecting efficiency, IDB operations



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

Box 2.3. Examples of efficiency assessments

a) Honduras cash transfer program

The Honduras cash transfer program (HO-L1093) is an example of positive efficiency rating. This was a US\$120.8 million investment loan that supported a cash transfer program to promote the accumulation of human capital – access to education and health- among minors in Honduran families living in extreme poverty.^a Management rated the efficiency of the program Excellent, and OVE concurred. The PCR carried out a cost-benefit analysis using data from an impact evaluation to estimate the impact on beneficiaries' future income due to increased use of education and health services. Using a real discount rate of 4 percent for the cash flows, the analyses found a benefit-cost ratio of 3.25¹⁷ and an internal rate of return¹⁸ above 12,9 percent in the base scenario. Sensitivity analysis confirmed the results as robust.^a

b) Health services in Buenos Aires

The health services in Buenos Aires project (AR-L1312) did not meet the minimum threshold as the cost-benefit analysis covered less than 60 percent of the total funded amount. This was a \$150 million results-based loan to strengthen health networks in the Buenos Aires province and ultimately improve the accessibility and effective coverage of public health services. The specific objectives were to increase the coverage of primary health care centers, improve the response capacity of emergency services, and reduce delays in dispensing cancer medications in the area. The program included many activities such as adopting standardized medical care models, providing medical equipment, implementing IT systems, and training medical staff. The program was approved in July 2019 and reached eligibility in April 2020, after the Buenos Aires province had declared a health emergency due to the COVID-19 pandemic (March 2020). In this context, adjustments were made to the amounts allocated to each component to include activities to support the emergency. At the start-up plan, new indicators related to the COVID-19 support were added to the specific objective of improving the response capacity of emergency services, and other indicator targets were adjusted to reflect the reallocation of resources. Management rated the efficiency of the project as Excellent, while OVE rated it as Partly Unsatisfactory because the cost-benefit analysis covers less than 60 percent of the total funded by the program.¹⁹ The project's cost-benefit analysis covered only 29.9 percent of the total funded by the program, or \$44.8 million. Following the PCR guidelines, OVE resorted to a cost and time overrun analysis (CTOA) based on PMR classifications.^c In response to the draft of this report, Management indicated some reasons why the CBA only covered a part of the total funding, including the difficulties of the PCR team in identifying reliable methods to estimate benefits for some oncology medications.

c) National Tourism Development Program in Bahia

The National Development Program in Bahia (BR-L1300) is an example of a complex project where conducting a CBA or CEA was not appropriate. The PCR team made efforts to assess the efficiency of a complex project, making several assumptions about benefits that were inherently difficult to measure. This was a US\$84.7 million investment loan, under the multiple works modality, aimed to generate formal employment, through tourism activity, in the "Todos os Santos" bay in the State of Bahia. The operation had multiple components, including investments to generate new nautical and cultural tourism options (e.g. restoration and construction of a network of boating facilities, urban renewal and landscaping of degraded areas), institutional strengthening, the preparation and implementation of a tourism marketing strategy and a solid waste management plan, socioenvironmental management activities. Management rated efficiency as Excellent, while OVE rated it as Unsatisfactory. The efficiency analysis was based on a complicated and unrealistic analysis based on a Computable General Equilibrium (CGE) model. The economic return was calculated assuming that its benefits result from impacts calculated through a differences-in-differences model. The CGE model was calibrated to reproduce an interregional Social Accounting Matrix for the state of Bahia, mainly constructed from an input-output matrix updated for the year 2019. The link between the differences-in-differences and computable general equilibrium models was established based on the elasticity of formal employment in hotels and accommodation relative to the production of these two sectors. The analysis included some inconsistencies and unrealistic assumptions (e.g. tourism spending values). This raised questions about the credibility of the results of the CGE analysis. Following the PCR Guidelines, OVE resorted to a cost and time overrun analysis (CTOA) based on semi-annual monitoring reports (PMR) classifications and downgraded management's efficiency rating.^c

Note: ^aSocial Protection System Support Program (HO-L1093) in Honduras; ^bProgram of Strengthening and Integration of Health Networks of the Province of Buenos Aires (AR-L1312).^cNational Tourism Development Program in Bahia (BR-L1300)

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

¹⁷ The Benefit-Cost Ratio (BCR) is the ratio of the sum of benefits (in present value) to the sum of costs (in present value). Ratios over one indicate that the project is efficient (higher ratios indicate higher efficiency).

¹⁸ The internal rate of return is the discount rate that makes the Net Present Value (compares all costs with benefits) to be zero, reflecting a break-even point where the flows of costs and benefits of the project are equal or balanced. Typically, if the ERR is larger than the recommended real discount rate the project would be efficient.

¹⁹ The PCR guidelines require that a cost-benefit analysis cover a minimum threshold of 60% of the total funded by the program according to the clarification of the PCR guidelines (FAQ file) agreed between OVE and SPD.

5. Sustainability

- 2.21 **An operation is rated positive on sustainability if no significant risks – technical, political, financial, economic, environmental, or social – threaten the continuation of achieved results.** The assessment evaluates these risks alongside government and stakeholder ownership (including the private sector and civil society), institutional support, governance quality, and the project's adaptability to changes in the operating environment. It also considers the probability and impact of risks beyond project completion, including those posed by natural or other disasters, and accounts for mitigation measures implemented during execution and residual risks that may persist. Environmental and social safeguards performance is assessed, including compliance measures, pending actions at closure, and planned IDB supervision activities.
- 2.22 **Out of 60 validated operations, 44 (73 percent) achieved positive sustainability ratings.** These 44 operations had a high likelihood of maintaining or improving the results achieved and satisfactory environmental and social (E&S) safeguards performance. Risks across technical, financial, economic, policy, institutional, and social dimensions that could compromise the continuity of results were accurately identified, and appropriate mitigation measures were implemented. Unmitigated risks associated with the projects were minimal or negligible. For projects that triggered environmental and social safeguards at approval, plans were implemented and monitored, with no unresolved issues reported at project closure. These 44 operations rated positive in sustainability benefited 20 countries, including Brazil (9 operations), Nicaragua (6), and Chile (3). These 44 operations were in 12 sectors. The sectors with the highest percentage of operations with positive sustainability ratings include financial markets (5 operations out of 5), modernization of the state (9 operations out of 10), and transport (10 operations out of 13). This sector disaggregation is illustrative and should be interpreted with caution because of the small number of operations in certain sectors. Box 2.4 describes an example of a project with a positive sustainability rating (2.4.a).
- 2.23 **The sustainability of 16 operations was rated negatively, mostly due to internal continuity risks.** Of the 16 operations rated negative in sustainability, five were affected by a combination of issues related to risk to continuity of results and unsatisfactory E&S performance, eight were only affected by issues concerning the continuity of results, and three only by E&S concerns. Thirteen operations faced internal continuity risks that had not been mitigated at project closure and whose impact was likely to be significant, such as insufficient financial resources, shortcomings in local institutional capacities, low commitment from key stakeholders, as well as external risks such as uncertain economic or political environments. The eight operations with unsatisfactory E&S performance (including operations that faced only E&S issues or a mix of E&S issues and continuity of results) had problems that remained unresolved at project closure and compromised their sustainability, including weak environmental management capacity, inadequate waste management, delays in resolving social conflicts or incomplete resettlement and compensation processes. Three of these operations did not include sufficient information on whether or how E&S issues had been adequately resolved at project closure. Box 2.4 provides an example of a project with unsatisfactory safeguards performance at project closure (2.4.b).

Box 2.4. Examples of sustainability assessments

a) Kingston water supply

The Kingston water supply project (JA-L1035) had a positive sustainability assessment. This was a \$133 million investment loan aimed at improving the efficiency, quality, and sustainability of drinkable water services in the Kingston Metropolitan Area and increasing access in selected urban centers of Jamaica.^a Management rated the sustainability of the program Excellent, and OVE concurred. At project closure, technical operation and maintenance of project-supported infrastructure were being conducted in accordance with engineering best practices, with the executing agency's capacity to do so having been strengthened by the project. OVE assessed positively the financial sustainability, as the revenue-to-operating cost ratio remained positive. Moreover, the project strengthened the executing agency's capacity to manage non-revenue water (NRW) reductions independently or through contractual arrangements. To sustain low NRW levels, the executing agency had initiated a follow-up performance-based NRW reduction contract in Portmore and was preparing a NRW reduction public-private partnership at the country level. E&S safeguards performance was rated as Satisfactory. All the works were undertaken in accordance with the E&S Management Plans, and no negative impacts were identified during the operational phase.

b) Brazil urban renewal

The Brazil urban renewal project (BR-L1055) had unsatisfactory performance at closure. This was a \$26.5 million investment loan aimed to improve the quality of life for residents of the Municipality of Niterói through urban renewal and social initiatives.^b The specific objectives were to improve city services in low-income neighborhoods, improve the city's transportation system, support the revitalization of decaying areas in the city center, and strengthen the institutional capacity of the city government. Activities included (1) financing comprehensive neighborhood improvement projects, (2) financing the implementation of an integrated system of "smart traffic signals," (3) redeveloping public spaces, and (4) training staff, purchasing equipment, and configuring IT systems for the municipality. Management rated the sustainability of the program Partly Unsatisfactory, and OVE concurred. After its closure in 2021, key issues persisted, particularly the resettlement and land regularization of 34 families affected by the project's municipal works. Frequent delays in document analysis led to repeated schedule revisions for land regularization, further complicating efforts to ensure long-term sustainability. COVID-19 also contributed to delays in resettlement and land regularization.

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

Note: ^aKingston Metropolitan Area Water Supply Improvement Program (JA-L1035) ^bUpdate BR-L1055: Niterói Urban Development and Social Inclusion Program (PRODUIS) (BR-L1386).

6. Noncore criteria ratings

- 2.24 **The PCR Guidelines also stipulate the assessment of two non-core criteria: Bank performance and Borrower performance.** The first concerns the quality of services provided by the Bank during the life of the project. The assessment of Bank performance takes into consideration two dimensions: (1) Bank performance in ensuring quality at entry; and (2) the quality of Bank supervision. The second non-core criterion concerns the effectiveness with which the Borrower discharged its responsibilities. The assessment focuses on the extent to which the Borrower ensured the quality of preparation and implementation of the operation and complied with covenants and agreements toward the achievement of development results. Non-core criteria are rated but do not count toward the project's overall outcome (See Figure 1.1).
- 2.25 **Bank performance was rated positive for 44 of 60 validated operations (73 percent).** The Bank adequately carried out the preparation and approval of the 44 operations rated positive to facilitate readiness for implementation, including adequate risk assessments and M&E plans. The Bank also provided timely support to the executing agencies during project implementation, including by ensuring adherence to fiduciary requirements. Sixteen operations were rated negative, including 14 with shortcomings in both preparation and supervision, one with shortcomings in preparation, and one in supervision.

A. Preparation. Among the 15 operations with preparation shortcomings, the most frequent problem was the poor identification of risks and mitigation mechanisms related to technical, financial, and institutional aspects. Most of these operations also presented weaknesses in the design of their M&E framework (14), including the selection of indicators that were inadequate to measure the expected outcomes (8), as well as weaknesses in their vertical logic (9).

B. Supervision. Among the 15 operations with supervision shortcomings, the most prevalent issue was deficiencies in project monitoring (12 operations), which usually included failure to correct M&E design issues (e.g., lack of baselines and targets for indicators) or to reflect project changes during execution (e.g., modifications in outputs) in the results matrix. Another frequent shortcoming concerned challenges that arose during implementation that were not addressed on time (10), including risks that materialized but were not mitigated. Most of these challenges stemmed from unclear definitions of responsibilities and ownership, or insufficient financial resources and technical capacity of the executing agencies.

- 2.26 **Borrower performance was rated positive for 35 of 60 operations (58 percent).** For these 35 operations, the borrower showed commitment, supported the quality of preparation and implementation of the operation, and complied with the covenants and agreements established with the Bank. Among the 25 operations rated negative, most were affected by the limited capacity of the executing unit (16 operations) due to high turnover and lack of appropriate skills. Administrative and operational challenges also affected the operations, as manifested by delays in delivering the outputs (16), cancellations and modifications of outputs (7), and coordination problems with other stakeholders (6). There were also challenges in complying with contractual agreements, especially on M&E responsibilities (12), fiduciary matters (including integrity and risk management; 5 operations), and E&S requirements (2 operations).

7. Projects Monitoring Classifications and Results at Closure

- 2.27 **Several projects with negative ratings at closure did not have “alert” or “problem” classifications in the Progress Monitoring Reports (PMRs).** Project implementation is monitored through the PMR, which should provide early warning signals to indicate when projects may require course correction. The PMR classifies project execution into three categories: “satisfactory”, “alert”, and “problem”. According to the PMR Guidelines, a satisfactory rating is assigned when the project is executed according to its plan; an alert rating, when there are delays and/or cost overruns in the project execution; and a problem rating when there are serious delays and/or cost overruns in executing the project. However, many of the projects that ultimately receive negative ratings at completion were never classified as “alert” or “problem” in the PMR, indicating that the signals from the PMR often fail to identify such cases. The analysis that follows considers the universe of projects that received a PMR classification during execution, that is 49 of the 60 validated operations. The remaining 11 operations were not eligible for a PMR classification during execution, according to the criteria stipulated by the PMR Guidelines (including 7 PBLs that have PMRs without classifications, 3 contingencies for natural disaster operations that do not have PMRs and 1 operation that did not have a classification as it closed the same year it became

eligible).²⁰ Considering Management ratings, thirty-five percent of the 23 operations with negative effectiveness ratings at completion never had an alert or problem classification in the PMR. Similarly, 33 percent of the 12 operations with negative overall outcome ratings by Management at completion did not have an alert or problem classification in the PMR. Forty percent of the 35 operations rated negative on effectiveness by OVE in this validation cycle did not have an alert or problem classification in the PMR. Similarly, 39 percent of the 28 operations with negative overall outcome ratings by OVE at completion did not have an alert or problem classification in the PMR. (Figure 2.9). This trend remains consistent when considering all operations validated between 2018 and 2024. These findings align with those of the Development Effectiveness Framework evaluation (document [RE-583-2](#)), which also found that the PMR signals were insufficient to predict the attainment of development results.

Figure 2.9. PMR classification and ratings at completion for IDB operations validated in 2024

		Management ratings at completion		OVE ratings at completion	
		Negative effectiveness	Negative overall outcome	Negative effectiveness	Negative overall outcome
Classification during execution	Satisfactory	35% (8)	33% (4)	40% (14)	39% (11)
	Alert or problem	65% (15)	67% (8)	60% (21)	61% (17)

Source: OVE based on data from IDB (2024). The universe corresponds to 49 operations with a PCR validated by OVE in 2024 that have at least one stage 2 PMR classification.

2.28 Most projects with negative ratings at closure that did have an alert or problem classification in PMRs had cleared their alert or problem status years before closure. IDB Project leaders tend to identify alerts or identify their projects as in problem status in the early stages of the project lifecycle and address alerts and problems years before project closure. Of the 49 operations with a PMR classification, only 1 had a problem classification at closure. The remaining 48 were classified as satisfactory in the last PMR, which contrasts with the rating at completion. This finding also highlights the disconnect between the PMR classification system and ratings at completion.

2.29 The IDB's PMR classification does not seem to capture well projects' likelihood of low ratings at closure. A key reason is likely to be that it focuses on schedule, cost and disbursement performance rather than on project's progress toward achieving outcomes. The IDB's alert and problem status classification during the project's execution (i.e. between eligibility and the disbursement of 95 percent of financial resources) depends on the value of a synthetic indicator. This indicator is calculated as a weighted average of schedule

²⁰ According to the PMR Guidelines, some projects do not have PMRs or PMR classifications. For example, policy-based loans (PBL), programmatic policy based-loans (PBP), loans based on results (PDL) and guarantees have PMRs but not a classification; investment grants of less than 3 million, contingency for natural disasters operations; immediate response facility for emergencies operations, and special development lending operations do not have PMRs. Specifically for this cohort, 7 are policy-based operations and do not have a classification, 3 are contingencies for natural disaster operations and do not have a PMR, and one operation reached operational closure by June of the same year when it reached eligibility, hence it never had a stage 2 classification.

and cost-performance indices and disbursement performance, measuring whether outputs are delivered within the expected budget and time frame, as well as how the project's disbursement ratio compares with the IDB's average disbursement profile for the country. Performance at closure depends however critically on the achievement of projects' outcomes, and not only on the delivery of outputs and disbursements. Progress towards achievement of project outcomes is a criterion that other Multilateral Development Banks do consider in their classification and alert systems, together with other relevant criteria, such as (among others) project management, procurement, audit and financial management (see OVE's Development Effectiveness Framework Evaluation (document [RE-583-2](#)) and an upcoming paper with a Comparative Analysis of Project Development Effectiveness Management Tools for Sovereign Guaranteed Operations).

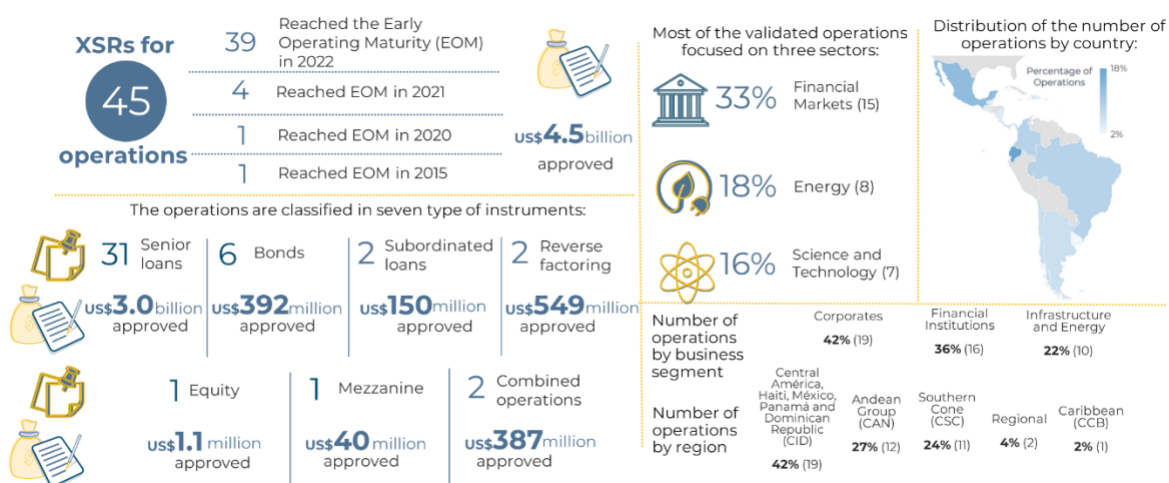
B. IDB Invest results in 2024

2.30 In 2024, OVE validated the Expanded Supervision Reports (XSRs) of 45 IDB Invest operations. The cohort of IDB Invest validated operations in 2024 is highly concentrated in some specific instruments and countries. In terms of instruments, the cohort includes 31 senior loans (69 percent), six bonds (13 percent), two subordinated loans (5 percent), two reverse factoring operations (5 percent), one equity (2 percent), one mezzanine (2 percent), and two combined operations (one senior loan and subordinated loan and one senior loan and reverse factoring, 4 percent). While the validated operations were in 15 of the 26 IDB Invest's borrowing countries, almost half of these operations benefited four countries: Ecuador (8 operations, 18 percent), Mexico (6 operations, 13 percent), Brazil (4 operations, 9 percent), and Colombia (4 operations, 9 percent).²¹ All validated operations were approved between 2011 and 2020, with most of them approved post-Merge Out²² (42 operations, 93 percent). Eighty-six percent (39) of the validated operations reached Early Operating Maturity (EOM) in 2022, while the remaining ones achieved EOM in 2021 (4 operations, 9 percent), 2020 (1 operation, 2 percent), and 2015 (1 operation, 2 percent). In terms of business segments, 19 operations (42 percent) were from the Corporates segment, 16 (36 percent) were from the Financial Institutions (FI) segment, and 10 (22 percent) were from the Infrastructure and Energy (INE) segment. More than 60 percent of the operations were focused on three sectors: financial markets (15 operations, 33 percent), energy (8 operations, 18 percent), and science and technology (7 operations, 16 percent) (Figure 2.10).

²¹ The number of operations validated by region in this cycle included one from the Caribbean (CCB), 11 from Southern Cone (CSC), 19 from Central America, Haiti, Mexico, Panama, and the Dominican Republic (CID), 12 from the Andean Group (CAN), and two regional operations (covering countries of more than one region).

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.

Figure 2.10. Characteristics of IDB Invest operations with XSRs validated in 2024



Source: OVE, with data from IDB Invest (2016, 2024a, 2024b, 2024c, 2024d).

Notes: Approval and disbursement data up to December 31, 2023. Operations in this validation cycle were expected to have reached EOM in 2022. For those that did so before, their EOM dates were not detected on time, or they were in Special Assets. The combined operations instrument category includes one senior loan and reverse factoring project, and one senior loan and subordinated loan project. Distribution of projects by country excludes two regional projects validated in this cycle that cover different countries.

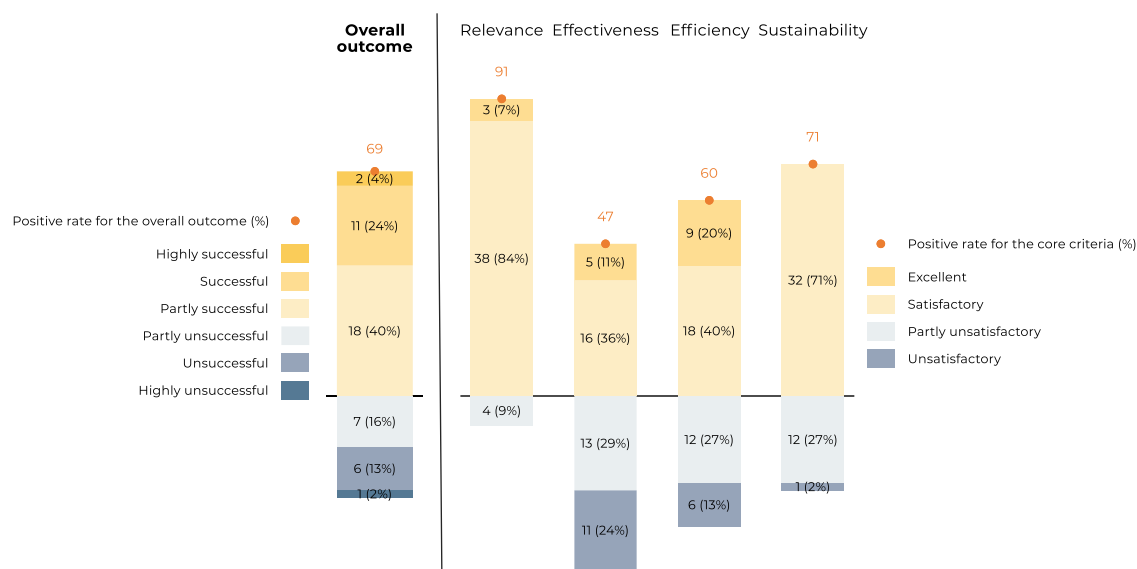
1. Overall outcome ratings

2.31 **Of the 45 validated operations, 31 (69 percent) achieved a positive overall outcome rating, with variations mainly across business segments and approval periods.** The 69 percent positive rate is above the target of 65 percent set in the 2020–2023 Corporate Results Framework (CRF). Forty-two operations (93 percent of the total) fell in the middle four ratings (“unsuccessful,” “partly unsuccessful,” “partly successful,” and “successful”). Relevance was the highest-rated core criterion, with 91 percent of operations rated positive, followed by sustainability (71 percent), efficiency (60 percent), and effectiveness (47 percent). (Figure 2.11). The proportion of operations with positive overall outcome ratings differs mainly across business segments and approval periods. However, these comparisons should be interpreted with caution because of the small number of operations in certain groups.²³ By business segment, 81 percent of the operations in the Financial Institutions (FI) segment were rated positive, while 63 percent of the Corporates operations and 60 percent of the Infrastructure and Energy (INE) operations were rated positive. There were no substantial differences in overall outcome ratings across regions.

²³

Section III briefly discusses IDB Invest performance by business segments, subregions, and approval period (pre- and post-merge out) using a larger number of validated operations during 2018–2014.

Figure 2.11. IDB Invest operations ratings, 2024



Source: OVE.

Note: Percentages may not sum to 100 due to rounding.

2. Relevance

2.32 **Relevance is rated positively when the project’s objectives and design align with the country’s development priorities and IDB Group’s priorities, both at approval and during implementation.** The relevance assessment focuses on the alignment of the project’s objectives and design with the country’s needs and development priorities, as well as with IDB Group’s priorities with respect to its private sector operations. Whenever the project’s conditions changed between approval and evaluation, the XSR should provide evidence of the extent to which the project was properly adapted to remain relevant.

2.33 **Forty-one of the 45 validated operations (91 percent) had positive relevance ratings.** These 41 operations had objectives that were aligned with country development needs, and IDB Group priorities, including IDB Group Country Strategies (CSs). These operations also had adequate project designs, with evidence supporting the cause-and-effect relationships between proposed outputs and outcomes (sound vertical logic). The analysis confirmed that the key assumptions underpinning the results chain held during implementation and that the project addressed the identified market failures. Changes during implementation did not affect the vertical logic of these operations. These 41 operations benefited 13 countries,²⁴ with most of them in Ecuador (8 operations), Mexico (6 operations), Brazil (4 operations), Colombia (4 operations), Argentina (3 operations) and Chile (3 operations). These 41 operations were in 10 sectors. The sectors with the highest percentage of operations with positive relevance ratings include energy (8 operations out of 8), agriculture and rural development (5 operations out of 5), financial markets sector (13 operations out of 15), and science and technology (6 operations out of 7). This sector disaggregation is illustrative and should be interpreted with caution because of the small number of operations

²⁴ The country disaggregation for each criterion excludes the two regional operations validated in this cycle covering more than one country.

in certain sectors. The 4 operations rated negative had design shortcomings such as inadequate client or instrument selection (3 and 2 operations, respectively), and weak vertical logic (2 operations). These factors are not mutually exclusive. Box 2.5 provides an example of an operation with a positive relevance rating (2.5.a) and an example of an operation with relevance shortcomings (2.5.b).

Box 2.5. Examples of IDB Invest relevance assessments

a) Access to finance for photovoltaic (PV) solar systems

An example of an IDB Invest project with a positive relevance rating is a senior loan to support a financial institution aimed at promoting access to finance for PV solar systems. The project aimed to expand the client's solar PV financing portfolio, extend the portfolio across eligible cooperatives, and finance the growth of the country's solar generation capacity. The project's objectives were aligned with the IDB Group's Institutional Strategy, the relevant IDB Group's Country Strategies (CSs), and the energy diversification goals of the country. The design of the operation featured a solid vertical logic, appropriate indicators to measure the project's progress, and advisory components to strengthen the company's sustainable finance capabilities, which include identifying and managing green and social projects and preparing Green Bond issuances. The client had an extensive rural presence, making it a relevant partner for supporting financing for eligible cooperatives with lower penetration in photovoltaic solar financing and promoting gender initiatives.

b) Increasing lending to SMEs and housing finance

An unsecured subordinated loan with the objective of increasing lending to SMEs and housing finance for lower-middle-income households provides an example of relevance shortcomings. Management and OVE concurred in rating the project's relevance as Partly Unsatisfactory. The project's relevance was weakened by an inappropriate selection of both client and instrument (loan structure). The client did not have the capacity at approval to disaggregate its portfolio and report SME loans or housing finance by income strata. (As the XSR notes, if a client does not have the capacity or experience to generate specific information that is indispensable for measuring a project's objectives, it should not be considered viable for that project, or technical assistance should be provided to the client to improve its reporting capacity).

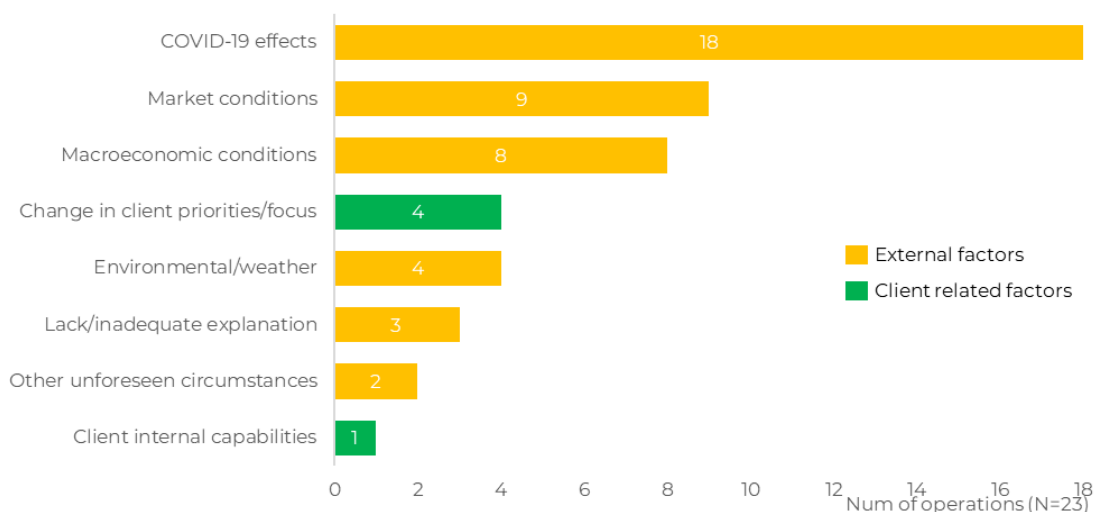
Source: OVE based on XSR contents and OVE's validation notes.

3. Effectiveness

- 2.34 **Similar to IDB operations, effectiveness at IDB Invest is rated positive when over half of the project development objectives are achieved.** To be rated positive, more than 50 percent of a project's development objectives must be achieved, with no development objective rated Unsatisfactory, and results must be plausibly attributable to the project. Project effectiveness is rated negative when the project did not achieve most of its development objectives, when one or more objectives were rated Unsatisfactory, and/or when achieved outcomes are not plausibly attributable to the project.
- 2.35 **Effectiveness ratings were positive for 21 of the 45 validated operations (47 percent).** The 21 operations rated positive achieved more than half of their objectives, and attribution of outcomes to the operation was plausible. For an objective to be considered achieved, its targets must have been largely reached (i.e. at least at 80 percent). The 21 operations with positive effectiveness ratings benefited 10 countries, with 6 of them in Mexico, 4 in Brazil, and 2 each in Chile and Colombia. These 21 operations were in 8 sectors. The sectors with the highest percentage of operations with positive effectiveness ratings include energy (5 operations out of 8), science and technology (4 operations out of 7), and financial markets (6 operations out of 15). This sector disaggregation is illustrative and should be interpreted with caution because of the small number of operations in certain sectors. Box 2.6a provides an example of an IDB Invest operation with a positive effectiveness rating.

- 2.36 **Unmet outcome targets were identified in 23 of 24 operations with negative effectiveness ratings.** Of the 24 operations with negative ratings, 17 were solely affected by unmet outcome targets, 1 by shortcomings in reporting only, and 6 by both unmet outcome targets and reporting shortcomings. The corporate (CO) and Financial Institution (FI) business segments had a higher proportion of projects with negative ratings, with 58 percent and 56 percent of operations receiving negative ratings, respectively. By contrast, 40 percent of the infrastructure and energy operations had negative ratings. Box 2.6 provides an example of an IDB Invest operation with a negative effectiveness rating (2.6.b).
- 2.37 **The most frequent reason for unmet outcome targets was external shocks, including the COVID-19 pandemic and market conditions.** The most frequently cited reasons for missing outcome targets were external (22 operations), including the pandemic (18 operations), followed by market conditions (9 operations). Of the 9 operations affected by market conditions, nearly half (4) experienced changes in direct or related demand, while the others encountered different types of challenges, including difficulties entering new markets, disputes over the product price, and product supply shortages, among others. Macroeconomic conditions influenced 8 operations in five countries. On the other hand, environmental or weather-related shocks, such as droughts, heavy rain, and landslides, affected 4 operations. Reasons related to changes in clients' strategies and operations affected achievement of outcome targets in 5 cases, including changes in the client's focus (4 operations) and client internal capabilities (1 operation). In the case of FI operations, clients usually adopted a more conservative lending strategy when exposed to adverse external conditions. Finally, 3 operations had one objective for which the reasons for not meeting their targets were either lacking or inadequate. (Figure 2.12).

Figure 2.12. Factors affecting effectiveness, IDB Invest operations

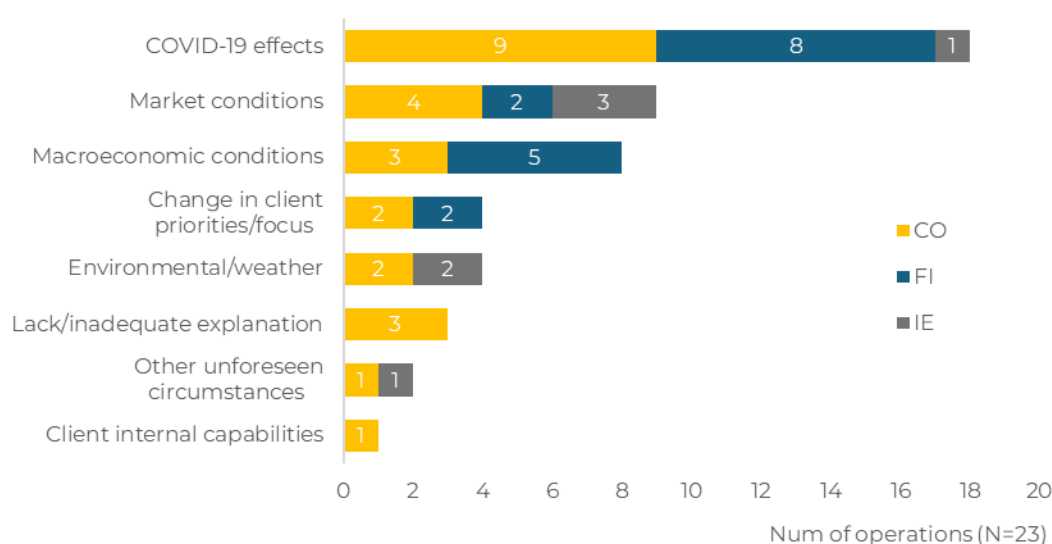


Source: OVE's categorization, based on XSR contents and OVE's validation notes.

Note: Numbers do not add up to 23 because reasons are not mutually exclusive.

2.38 **The reasons for unmet outcome targets vary depending on the business segment.** Corporates are the group with the highest percentage of operations (11 out of 19t) that faced external challenges as well as client-related changes and that in some cases did not provide adequate explanations for not meeting their targets. Financial Institutions had the second largest percentage of operations (8 out of 16) whose performance was substantially affected by external factors, such as the COVID-19 pandemic, economic recessions and high unemployment. Changes in client’s focus affected both Financial Institution operations (clients targeting less risky beneficiaries) and Corporates (clients changing their strategies). The performance of Infrastructure and Energy operations was solely affected by external factors, including the COVID-19 pandemic, market conditions, environmental or weather-related problems. (Figure 2.13).

Figure 2.13. Factors behind unmet outcome targets by business segment, IDB Invest operations



Source: OVE’s categorization, based on XSR contents and OVE’s validation notes.
Notes: CO= Corporates, FI= Financial Institutions, IE= Infrastructure and Energy. Numbers do not sum to 23 because reasons are not mutually exclusive.

2.39 **Reporting shortcomings contributed to negative effectiveness ratings for seven operations in the 2024 validation cycle.** The most frequent problem was missing data, affecting four operations, followed by two operations with inadequate indicators, one with overambitious targets, one with insufficient indicators to measure project objectives, and one with unreliable data. (The issues are not mutually exclusive.)

2.40 **As with the effectiveness ratings of IDB operations, those of IDB Invest operations may be adversely impacted by external circumstances and by reporting deficiencies.** As for IDB, IDB Invest effectiveness ratings may be negatively impacted by circumstances that were unforeseeable at the time of project approval and that often cannot be easily taken into account during project implementation (due, for example, to the difficulties related to project restructuring). Also, similarly to IDB, IDB Invest projects’ effectiveness ratings may also be negatively affected by reporting shortfalls. Tagging a project as ineffective due to shortcomings in reporting (e.g. lack of data, attribution problems, inadequate indicators) could be misleading because it is not possible to affirm that

the project did not achieve outcomes. Box 2.6 provides an example of an operation whose effectiveness rating was affected by several unforeseen circumstances (2.6.c).

Box 2.6. Examples of IDB Invest effectiveness assessments

a) Renewable energy

A senior energy loan provides an example of an IDB Invest operation with a positive effectiveness rating. The project aimed to (i) increase the supply from non-conventional renewable sources and (ii) avoid greenhouse gas emissions by providing long-term financing for the construction, operations, and maintenance of three wind power plants and associated facilities, with a combined installed capacity of 332 MW. The project received a satisfactory effectiveness rating by Management and OVE, achieving more than 80 percent of the targets set in the results matrix for both objectives on average. Specifically, the project generated an average of 763.24 MWh and achieved a reduction in greenhouse gas emissions of approximately 526,000 tCO₂eq/year considering the three plants. The project faced multiple difficulties during construction and one key lesson identified was the importance of strong partnerships, effective communication, and positive collaboration to successfully navigate challenges and uncertainties during project implementation.

b) Growing financial institutions' SME portfolios

An example of an operation with a negative effectiveness rating was a subordinated loan with a financial institution. Its objective was to promote growth in the portfolios of small and medium enterprises (SMEs) and women-led small and medium enterprises (WSMEs) held by a financial institution. Both Management and OVE rated the operation as Partly Unsuccessful and its effectiveness as Partly Unsatisfactory. The client's SME portfolio contracted in real terms between 2018 and 2023, even though its overall lending grew over the same period. The contraction in the SME portfolio was primarily driven by changes in the client's risk appetite policies, which led to a more conservative approach and a focus on more creditworthy SMEs, thereby limiting the pool of potential clients and hindering portfolio growth. Although the WSME portfolio grew in real terms between 2019 and 2022, it was not enough to offset the overall underperformance in the SME segment.

c) Production and export of sugar

An example of an operation whose rating was highly influenced by unforeseeable circumstances was a corporate senior loan with the purpose to increase the production and export of sugar and to support sugarcane growers. Both Management and OVE rated the overall operation as Unsuccessful and its effectiveness as Unsatisfactory. The project did not reach its outcome targets in terms of sugar production. The main cause of this shortfall was the 2019 drought in the country. This drought critically delayed the sugarcane harvest, affecting sugar production in 2019 and the client's subsequent sales and exports. Other external factors contributed as well, such as the COVID-19 pandemic which affected the sugar industry globally. Furthermore, a lengthy dispute with a farmer association, delayed harvests, reducing production.

Source: OVE based on XSR contents and OVE's validation notes.

4. Efficiency

2.41 **Efficiency is assessed by the extent to which the financial and economic benefits of the project exceed project costs, taking into account also the time-value of money.** Efficiency is rated considering the financial and economic performance achieved by the project. Specific benchmarks and metrics to measure efficiency consider the project type and business segment. Financial performance assesses the project's contribution to the company's financial results and the extent to which the project has delivered on the process and business objectives stated at approval. This is usually measured using metrics such as the Financial Rate of Return (FRR) or Return on Invested Capital (ROIC), benchmarked against the project's Weighted Average Cost of Capital (WACC). For Financial Institution operations, this performance is measured through the performance of the relevant sub-portfolio financed by the operation. Economic performance, on the other hand, evaluates the project's broader impact on the host economy beyond direct financial returns, i.e., if the benefits surpass the cost of capital after adjusting for economic

distortions. This involves comparing the economic rate of return (ERR) or the economic return on invested capital (EROIC) with the WACC, accounting for all the costs and benefits to key economic stakeholders. For Financial Institution operations, the rating is based on the economic contributions of the sub-borrowers, the evidence that the borrowers are repaying their loans (typically expressed as non-performing loans -NPLs-), or the comparison of the gross equity fund portfolio return with respect to its weighted average cost of capital (WACC). In addition, the XSR's efficiency assessment considers aspects of project design and implementation that either contributed to or reduced efficiency.

- 2.42 **Efficiency ratings were positive for 27 of the 45 validated operations (60 percent).** For the 27 operations rated positive, the financial or economic performance was at least satisfactory. This means that the project meets or exceeds its financial projections, with FRR or ROIC at least equal to the WACC, surpassing business objectives and demonstrating strong financial and operational performance. In the case of FI operations, there was evidence that the sub-portfolio had a positive effect on the intermediary's profitability. Projects were also positively rated for achieving economic benefits, such as generating employment, boosting productivity, or enhancing the financial viability of targeted sub-borrowers or investees. These 27 operations benefited 11 countries. Most of them were in Ecuador (6 operations), Mexico (4 operations), and Brazil (4 operations). These 27 operations were in 10 sectors. The sectors with the highest percentage of operations with positive efficiency ratings include industry (2 operations out of 2), financial markets (11 operations out of 15), science and technology (5 operations out of 7), agriculture and rural development (2 operations out of 5), and energy (2 operations out of 8). This sector disaggregation is illustrative and should be interpreted with caution because of the small number of operations in certain sectors. Box 2.7 provides an example of an operation with a positive efficiency rating (2.7.a).
- 2.43 **Among the 18 operations rated negative, 8 had both financial and economic efficiency shortfalls, 6 had only economic shortfalls, and 4 had only financial shortfalls.** The benefits of these projects fall below their respective benchmarks. In several cases, the operation's financial or economic returns were lower than the cost of capital (11 operations) (e.g. ERR or ROIC below the WACC). Also, three FI operations faced lower-than-expected growth in their target portfolios or could not reach the intended beneficiaries. Some operations also experienced higher-than-expected nonperforming loan (NPL) rates (2). Box 2.7 provides an example of an operation with efficiency shortcomings (2.7.b).
- 2.44 **The efficiency ratings of IDB Invest projects may be affected by inadequate reporting or unforeseen circumstances.** Three operations were rated negative because they did not provide sufficient information, preventing a full efficiency assessment. Also, three presented complications related to COVID-19 or other external factors that affected the efficiency of the operation. The example in Box 2.7 describes an operation whose efficiency rating was rated negative by both Management and OVE because the project's ROIC and EROIC were below the WACC as a result of a drought (2.7.c).

Box 2.7. Examples of IDB Invest efficiency assessments

a) Telecommunication infrastructure

An example of an IDB Invest operation with a positive efficiency rating is a senior loan aimed to expand the coverage of mobile broadband networks and to increase the productivity of telecommunication infrastructure by promoting infrastructure sharing (i.e., increasing the tenancy of towers). The operation achieved a Return on Invested Capital (ROIC) greater than the Weighted Average Cost of Capital (WACC), driven mainly by higher-than-expected revenues, and a higher debt-to-equity ratio. The Economic Return on Invested Capital (EROIC) of the operation was greater than the WACC, suggesting economic value creation.

b) Clean energy generation

An example of an IDB Invest operation with a negative financial and economic efficiency rating is a senior loan aimed to increase the supply of energy from clean energy sources, to displace GHG emissions and to generate employment during its implementation. Both management and OVE rated the operation's financial and economic efficiency as Unsatisfactory because the returns of the project were lower than its WACC. The operation achieved a negative Financial Rate of Return (FRR) attributable mainly to lower-than-expected production levels. High inflation rates also affected the FRR. Similarly, the Economic Rate of Return (ERR) was negative, given the lower production level.

c) Agricultural value chain

An example of an IDB Invest operation whose efficiency rating was affected by unforeseen circumstances is a senior loan with the aim to enhance the agricultural value chain, increase exports, and create local jobs in a vulnerable area. The loan was intended to support capital investments and permanent working capital and refinance an existing loan for the client's agricultural and livestock activities. Both Management and OVE rated the operation's efficiency as Partly Unsatisfactory. Severe droughts in 2020 and 2022 had a profound negative impact on the client's operational performance, financial health, and ability to achieve its development and business objectives, significantly impacting both financial and economic outcomes. The project's Return on Invested Capital (ROIC) was below the Weighted Average Cost of Capital (WACC). The Economic Return on Invested Capital (EROIC) was below the WACC, indicating limited capital efficiency.

Source: OVE based on XSR contents and OVE's validation notes.

5. Sustainability

- 2.45 **Sustainability is an assessment of the conditions that could influence the continuation of the results already achieved and the achievement of future expected results during the lifecycle of the project.** Sustainability is assessed based on the likelihood and impact of threats to the continuation of outcomes beyond Early Operating Maturity (EOM) and the success of mitigation measures in the project's design or execution. This assessment covers environmental, social, technical, financial, economic, and company governance aspects. Environmental and social (E&S) performance is measured against the IDB Group E&S standards. To be rated positive, the project should be in material compliance with E&S requirements at approval, including the implementation of the Environmental and Social Action Plans (ESAPs).
- 2.46 **Sustainability ratings were positive for 32 of the 45 validated operations (71 percent).** These 32 operations achieved results that were likely to continue over time and adhered to the E&S standards set by the IDB Group. Factors influencing project and company sustainability – such as financial and economic sustainability, governance, market, and technical considerations – were accurately identified and mitigated. Additionally, the client's overall profitability, adaptability, and prospects for sustainability and growth were assessed as satisfactory. These projects also properly implemented ESAPs to address potential gaps with respect to IDB Group E&S standards, whether anticipated during project design or identified later. No significant natural resource or other disasters, nor operational, sectoral, or country-

specific risks, were anticipated to disrupt the continuation of results. These 32 operations benefited 13 different countries, with most of them in Ecuador (8 operations), Mexico (5 operations), and Brazil (3 operations). These 32 operations were in 11 sectors. The sectors with the highest percentage of operations with positive sustainability ratings include financial markets (12 operations out of 15), agriculture and rural development (4 operations out of 5), science and technology (4 operations out of 7), and energy (4 operations out of 8). This sector disaggregation is illustrative and should be interpreted with caution because of the small number of operations in certain sectors. Box 2.8 provides an example of an operation with a positive sustainability rating (2.8.a).

- 2.47 **The sustainability of the 13 operations with a negative rating was affected both by uncertainty surrounding the continuity of results and by weak E&S performance.** Among the 13 operations rated negative, 5 were exclusively affected by the uncertain continuity of results, 3 exclusively by E&S performance, and 5 by both factors. Continuity of results was at risk due to unstable financial standing of the client (6 operations), an uncertain environment (2)—including changing market conditions and political factors—as well as operational challenges (2) and compliance risks (1). E&S performance was affected by shortfalls in fulfilling requirements (5 operations), including delays in implementing E&S action plans, lack of timely reporting on critical events by the client, and missed report submissions. Three cases with unresolved E&S safeguard issues faced challenges in environmental management, including deficiencies in air quality monitoring, biodiversity protection, and unmitigated risks to local communities and ecosystems. Two operations were subject to MICI investigations. Box 2.8 provides an example of an operation with uncertain continuity of results (2.8.b).

Box 2.8. Examples of IDB Invest sustainability assessments

a) Expanding healthcare

A senior loan to support a non-profit organization expand healthcare capacity (private and public) in response to the COVID-19 crisis provides an example of an IDB Invest operation with a positive sustainability rating. The objective of the operation was to enhance healthcare services through innovation and knowledge dissemination and strengthen the health sector supply chain. At closure, continuity of results was expected due to the client's strong financial position and governance. The client exhibited robust creditworthiness and maintained high liquidity reserves in cash and securities despite COVID-19 pandemic challenges. A stable institutional framework and an internal income base further mitigated external risks. The E&S performance of the operation was classified as satisfactory. The client demonstrated commitment and adequate compliance with the E&S Action Plan, as well as adherence to international standards, effective planning, and regular training programs that enhanced organizational capacities to deal with E&S issues.

b) Solar photovoltaic plant

A senior loan aimed at increasing the supply of energy from low-carbon sources and reducing greenhouse gas (GHG) emissions provides an example of an IDB Invest operation with a negative sustainability rating. The operation consisted in contributing to the financing of the construction and operation of a solar photovoltaic plant. Both Management and OVE rated the operation's sustainability as Partly Unsatisfactory. Although environmental and social sustainability was rated Satisfactory, the project faced an array of financial, operational, regulatory, and compliance risks. In May 2020, IDB Invest downgraded the project's financial risk ratings and lower-than-expected spot prices due to the COVID-19 pandemic, impacting the project's cash flow. In December 2023, IDB Invest placed the project on its watchlist due to operational and financial challenges within the project—including governance issues, financial risks, and strained relationships with contractors—that affected sustainability by undermining the long-term viability of the project.

Source: OVE based on XSR contents and OVE's validation notes.

6. Noncore criteria ratings

- 2.48 **Besides the core criteria, the XSR Guidelines also require the assessment of three non-core criteria: additionality, investment profitability, and work quality.** Non-core criteria are rated but do not count toward the project's overall outcome (See Figure 2.1). For IDB Invest, additionality refers to the key financial and non-financial inputs it brings to make the project or investment happen, make it happen sooner than it would otherwise, and/or improve its development impact and sustainability.²⁵ The assessment of investment profitability considers the investment's gross profit contribution (net of financing cost and loss provision but before deducting administrative costs) with rating benchmarks set in relation to corresponding at approval standards for minimally expected performance.²⁶ Finally, the work quality assessment takes into consideration (1) pre-commitment work at entry, i.e., screening, appraisal, and structuring / underwriting, and (2) the monitoring and supervision of the operation following commitment / guarantee issuance.
- 2.49 **All 45 validated operations were rated positive on additionality, and 36 of them (80 percent) were rated positive on investment profitability.** All 45 operations provided the client with either financial or nonfinancial additionality. The 36 operations with positive ratings on investment profitability had gross profits that met or exceeded the benchmarks set at approval, while 9 operations missed the benchmarks due to prepayments or cancellations for loans and negative returns for one equity investment.
- 2.50 **Work quality ratings were positive for 24 of 45 operations (53 percent).** For the 24 operations rated positive, IDB Invest effectively carried out its screening, appraisal, and structuring work prior to approval and commitment of the investment. Moreover, IDB Invest adequately executed its responsibilities following the commitment of the investments, including the monitoring and supervision of project risks, E&S performance, and development results. Of the 21 operations rated negatively on work quality, 18 had shortfalls exclusively during preparation, two in both preparation and supervision, and one in supervision. The most frequent shortcomings affecting work quality during structuring phase were related to M&E aspects (14 operations) including deficiencies in the results matrices and indicators, as well as inadequate project designs (9 operations), including challenges with the selection of instruments and clients. During supervision, IDB Invest did not collect key data from the client in two of the three operations with shortcomings at this phase.

²⁵ Financial additionality refers to the terms and conditions offered by IDB Invest and the resources it mobilizes from third parties. Non-financial additionality refers to non-financial products and services provided by IDB Invest to maximize the impact of projects and reinforce client's capabilities (e.g. mitigating risks). It also includes generating knowledge and supporting innovation and capacity building. According to the XSR Guidelines, "the assessment of additionality considers IDB Group's value proposition in providing support to the project". "The assessment and rating are based on the counterfactual assessment of how the operation would have (or would not have) proceeded without IDB Group's support". For a positive rating (at least a Satisfactory rating), "it is clear that: a) the project would not have gone ahead without IDB Group or b) absent the IDB Group, i) it would have entailed an unfair or inefficient allocation of risks and responsibilities or ii) it would have been weaker in business, developmental, social or environmental terms".

²⁶ Specific standards and benchmarks are considered for loans, equity investments, guarantees, and operations with more than one type of investment.

7. Projects Monitoring Classifications and Results at Closure

2.51 **The signals raised by IDB Invest’s project monitoring reports were aligned with project performance at completion.** IDB Invest uses the Development Effectiveness Learning, Tracking, and Assessment (DELTA) tool for project approval and supervision²⁷ During supervision, projects are classified as “satisfactory,” “alert,” “problem”, or “insufficient information” based on the deviation from the ex-ante DELTA score. In contrast to the PMR used in IDB operations, most IDB Invest operations with negative effectiveness ratings at completion had an alert or problem classification in the DELTA at supervision. The analysis that follows considers the universe of projects that had a DELTA classification during supervision, that is 43 of the 45 operations in this cohort. The remaining 2 operations did not have a DELTA classification from supervision because they were part of the special assets portfolio and, therefore, did not follow conventional supervision procedures. Considering Management ratings, 89 percent out of 19 operations with negative effectiveness ratings at completion had an alert or problem classification in the DELTA during the period under supervision (Figure 2.14). Similarly, of the 11 operations with negative overall outcome ratings at completion, 91 percent had an alert or problem classification in the DELTA. Considering OVE ratings, 86 percent out of 22 operations with negative effectiveness ratings at completion had an alert or problem classification in the

²⁷ The DELTA measures an operation’s potential impact on a scale from zero to ten and factors in several analyses in the areas of development outcomes and (financial and nonfinancial) additionality (evaluation of IDB Invest, document [RE-577](#))

DELTA at supervision. Similarly, of the 12 operations with negative overall outcome ratings at completion, 92 percent had an alert of problem classification in the DELTA.

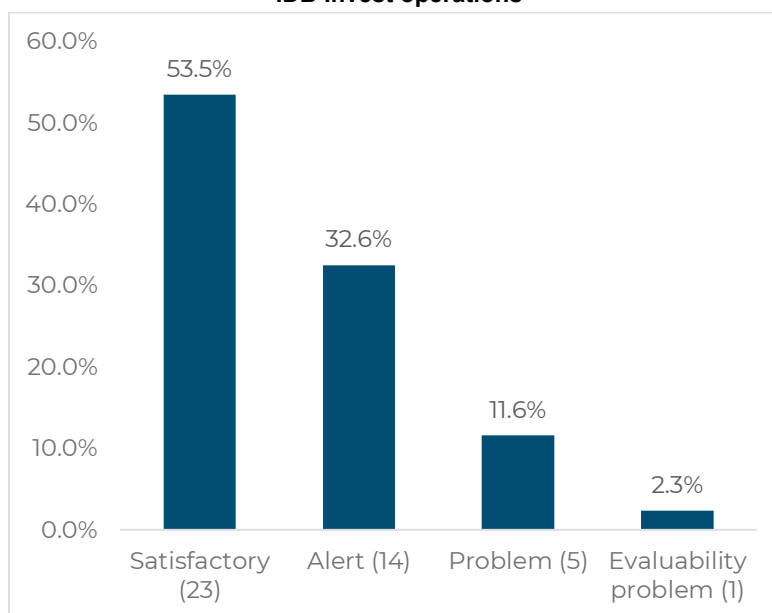
Figure 2.14. DELTA classification and ratings at completion for IDB Invest operations validated in 2024

		Management ratings at completion		OVE ratings at completion	
		Negative effectiveness	Negative overall outcome	Negative effectiveness	Negative overall outcome
Classification during execution	Satisfactory	11% (2)	9% (1)	14% (3)	8% (1)
	Alert or problem	89% (17)	91% (10)	86% (19)	92% (11)

Source: OVE based on data from IDB Invest (2024e). The universe corresponds to 43 operations with a XSR validated by OVE in 2024 that have at least one DELTA classification of satisfactory, alert, or problem. DELTA classification considers if the operation had an alert or problem during the period under supervision.

2.52 The issues raised by project monitoring reports tend not to be resolved during execution. The DELTA classification aligns more closely with results achieved at closure than the PMR. Given the DELTA’s focus on outcomes, IDB Invest maintains projects in “Alert” or “Problem” until the factors affecting project performance are solved (either because external conditions improved or corrective measures are implemented). Consequently, the signals raised by the DELTA often remain unresolved throughout execution. Of the 43 operations with a DELTA classification from supervision, 44 percent were classified as alert or problem in their last available supervision before 2024 (Figure 2.15).

Figure 2.15. DELTA at supervision classification– IDB Invest operations



Source: OVE based on data from IDB Invest (2024e).

Notes: The universe corresponds to 43 operations with a XSR validated by OVE in 2024 that have at least one DELTA classification of satisfactory, alert or problem. DELTA classification corresponds to the last available supervision before 2024.

III. PERFORMANCE BY SECTORS, SUB REGIONS AND INSTRUMENTS 2018–2024

3.1. **This section compares the average overall outcome ratings of the last six validation cycles to identify patterns of performance of relevant groups of IDB and IDB Invest operations.** The analysis includes 447 IDB and 294 IDB Invest validated operations for the period 2018–2024.²⁸ Results should be interpreted with caution for two reasons: (i) the ratings are sensitive to the limited number of operations in certain groups, and (ii) the results for IDB need to consider the adjustments in the PCR guidelines in 2020 that included, for example, additional guidance for specific project types, loan instruments, and efficiency analysis.

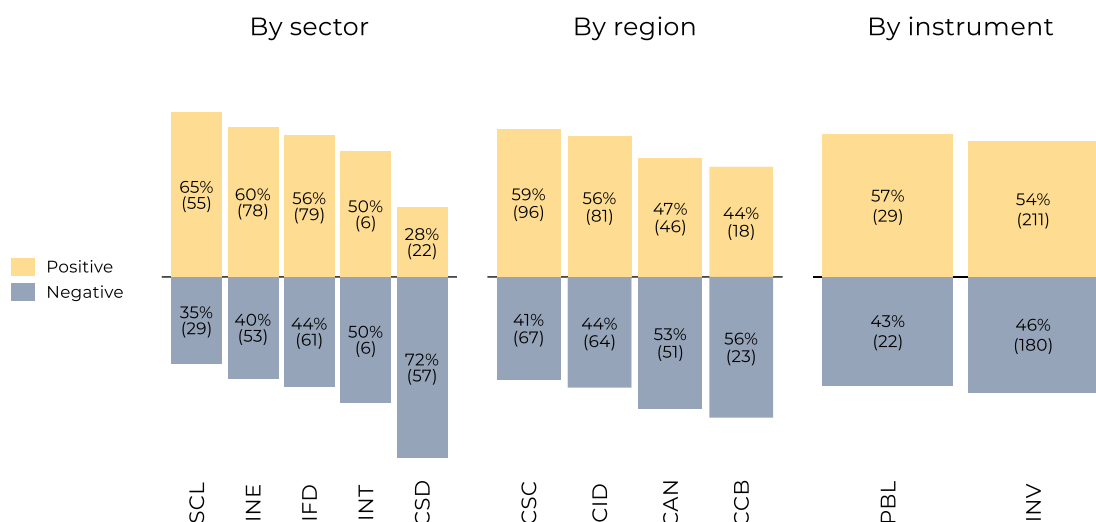
A. IDB performance

3.2. **The proportion of IDB operations with positive ratings in overall outcomes differs across sectors and regions but not across lending instruments.** Half or more of the operations validated in the 2018–2024 period in the Social (SCL), Infrastructure and Energy (INE), Institutions for Development (IFD), and Integration and Trade (INT) sectors achieved positive overall outcomes on average. The Climate Change and Sustainable Development (CSD) sector was an exception, with a 28 percent positive rating, mainly because this sector has had lower positive ratings in effectiveness and sustainability than other sectors. By region, the Southern Cone (CSC) and Central America, Haiti, Mexico, Panama,

²⁸ The analysis by region excludes one regional IDB operation and 15 regional IDB Invest operations.

and the Dominican Republic (CID) had average positive outcome rates above 55 percent, while the Andean Group (CAN) and the Caribbean (CCB) were below 48 percent. Differences across lending instruments are, on average, minor. (Figure 3.1).²⁹

Figure 3.1. Overall outcome ratings for all IDB operations validated in 2018–2024



Source: OVE.

Notes: Number of operations in parentheses. Acronyms by sector: SCL=Social Sector, INE=Infrastructure and Energy, IFD=Institutions for Development, INT=Integration and Trade, and CSD=Climate Change and Sustainable Development Sector. Acronyms by region: CSC=Southern Cone, CID=Central America, Haiti, Mexico, Panama, and the Dominican Republic, CAN=Andean Group, and CCB=Caribbean. Acronyms by instrument: PBL= Policy-Based loans and INV= Investment loans. Out of 447 operations validated by OVE during the period, sector data excludes one operation shared between the INE and CSD, regional data excludes one regional operation, and instrument data excludes five 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.

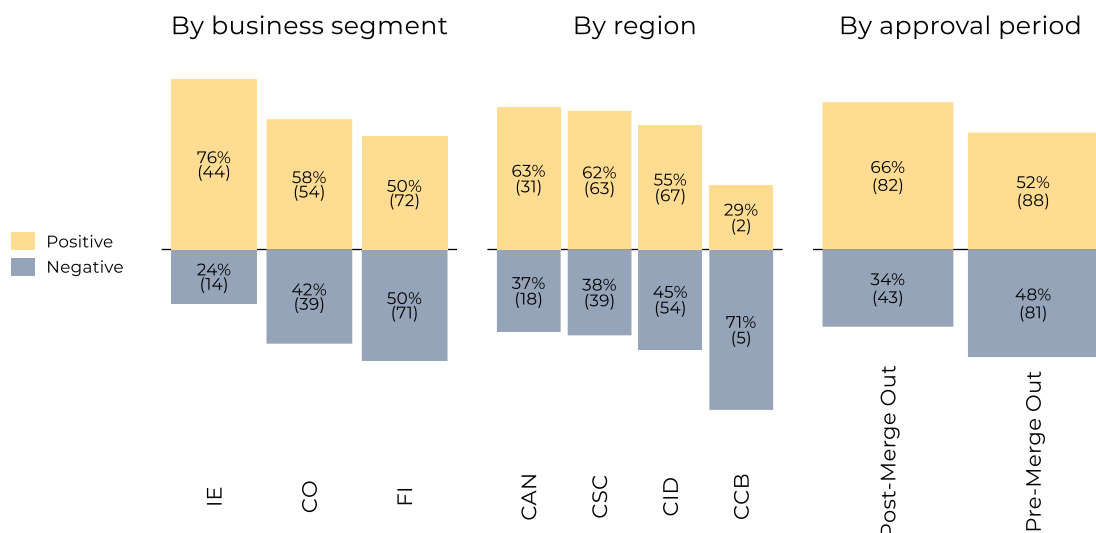
B. IDB Invest performance

3.3. **The proportion of IDB Invest operations with positive ratings in overall outcomes differs across business segments, regions, and between pre- and post-merge-out operations.** Seventy-six percent of 58 operations validated in the 2018–2024 period in the Infrastructure and Energy segment (IE) achieved a positive overall outcome on average, versus 58 percent of the 93 operations of the Corporate (CO) and 50 percent of the 143 Financial Institutions (FI) operations. The proportion of operations with positive ratings in overall outcomes in the Andean Group (CAN) and the Southern Cone (CSC) is above 60 percent, in Central America, Haiti, Mexico, Panama, and the Dominican Republic (CID) it is 55 percent, while in the Caribbean (CCB) it is 29 percent. However, this result should be interpreted with caution because the data only includes a small number of operations in the Caribbean (7). Finally, the proportion in post-merge-out-operations was 14 percentage points higher than in pre-merge-out-operations. (Figure 3.2).

²⁹

It should be noted that some groups have a small number of validated operations, such as the Integration and Trade Sector (12) or the Caribbean region (42).

Figure 3.2. Overall outcome ratings for all IDB Invest operations validated in 2018–2024



Source: OVE.

Note: Number of operations in parentheses. Acronyms by business segments: IE=Infrastructure and Energy, CO=Corporates, and FI=Financial Institutions. Acronyms by region: CAN=Andean Group, CSC=Sothorn Cone, CID=Central America, Haiti, Mexico, Panama, and the Dominican Republic, and CCB=Caribbean. Out of the 294 IDB Invest operations validated by OVE during the period, regional analysis excludes 15 regional operations. As in the case of the previous validations report (2023), for comparability purposes, SAT operations without XSRs are excluded from this analysis (4 in 2022 and 2 in 2023).

IV. QUALITY OF SELF-EVALUATIONS AND DIFFERENCES BETWEEN MANAGEMENT’S AND OVE’S RATINGS

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 and quality of the analyses, documenting changes in design and results matrix, candor, and lessons, among other aspects. In this cycle, OVE assessed the quality of 59 PCRs corresponding to 60 validated operations.³⁰ Of the 59 PCRs assessed, OVE rated 30 (51 percent) as positive.
- 4.2 **OVE identified several shortcomings in the quality of the 29 PCRs rated negative.** Shortcomings related to relevance (present in 59 percent of PCRs) included weak analysis of initial alignment with country realities and insufficient assessment of how implementation changes affected relevance. Those related to effectiveness (present in 76 percent of PCRs) included inadequate indicators (e.g., missing baselines or targets), poor attribution analysis, and insufficient explanations for target outcomes. Those related to efficiency (72 percent of PCRs) included unreliable assumptions, lack of sensitivity analysis, and missing information in cost-benefit or cost-effectiveness analyses. Those related to sustainability (41 percent) included gaps in reporting on the continuation of results, risk mitigation, and safeguards. Finally, shortcomings related to noncore criteria (24 percent) included insufficient reporting on the performance of the Bank and borrowers (Table 4.1). The shortcomings were similar to last year’s.

³⁰ Management presented one PCR for 2 operations in Haiti.

4.3 **In response to OVE’s recommendations, Management is implementing an Action Plan (AP) to improve the overall quality of PCRs.** The AP includes four main actions: (i) strengthen the quality assurance instrument used by the Strategic Development Effectiveness Division (SPD/SDV) to review PCRs prior to submission to OVE; (ii) ensure that PCR teams report on the reasons for underachievement (i.e., unmet outcome targets); (iii) strengthen SDV’s quality assurance instruments for economic analysis; and (iv) provide training to IDB operational staff.³¹ OVE is collaborating with SPD on activities included in the AP, including joint training initiatives for operational staff and the recent establishment of a working group to analyze and address key issues related to the quality of the efficiency analysis.

Table 4.1. Most common shortcomings in PCR quality

Type of shortcoming	Number of PCRs (out of 29)
Relevance: at design (weak analysis of alignment with country realities and of vertical logic) and during implementation (poor analysis of changes made in the results matrix and of how changes during implementation affected relevance).	17 (59%)
Effectiveness: quality of indicators (inadequate baseline, targets, data used, or calculations), quality of analysis (weak attribution analysis, misidentification of outputs as outcomes), insufficient explanations of achievement/underachievement of changes in outcome targets.	22 (76%)
Efficiency: Lack of credibility (questionable assumptions, unclear methodology, questionable comparators/control group, lack of sensitivity analysis, unreliable data), key information missing on CBA/CEA.	21 (72%)
Sustainability: key information missing on the continuation of results (e.g., on risks and actions to mitigate them, on the rationale for assumed continuation of results) and on safeguards (disclosure of relevant issues, report on implementation and results).	12 (41%)
Noncore criteria/Bank and borrower performance: insufficient reporting on key performance issues.	7 (24%)

Source: OVE.

Note: The number of PCRs does not add up to 29, nor do the percentages add up to 100, because shortcomings are not mutually exclusive.

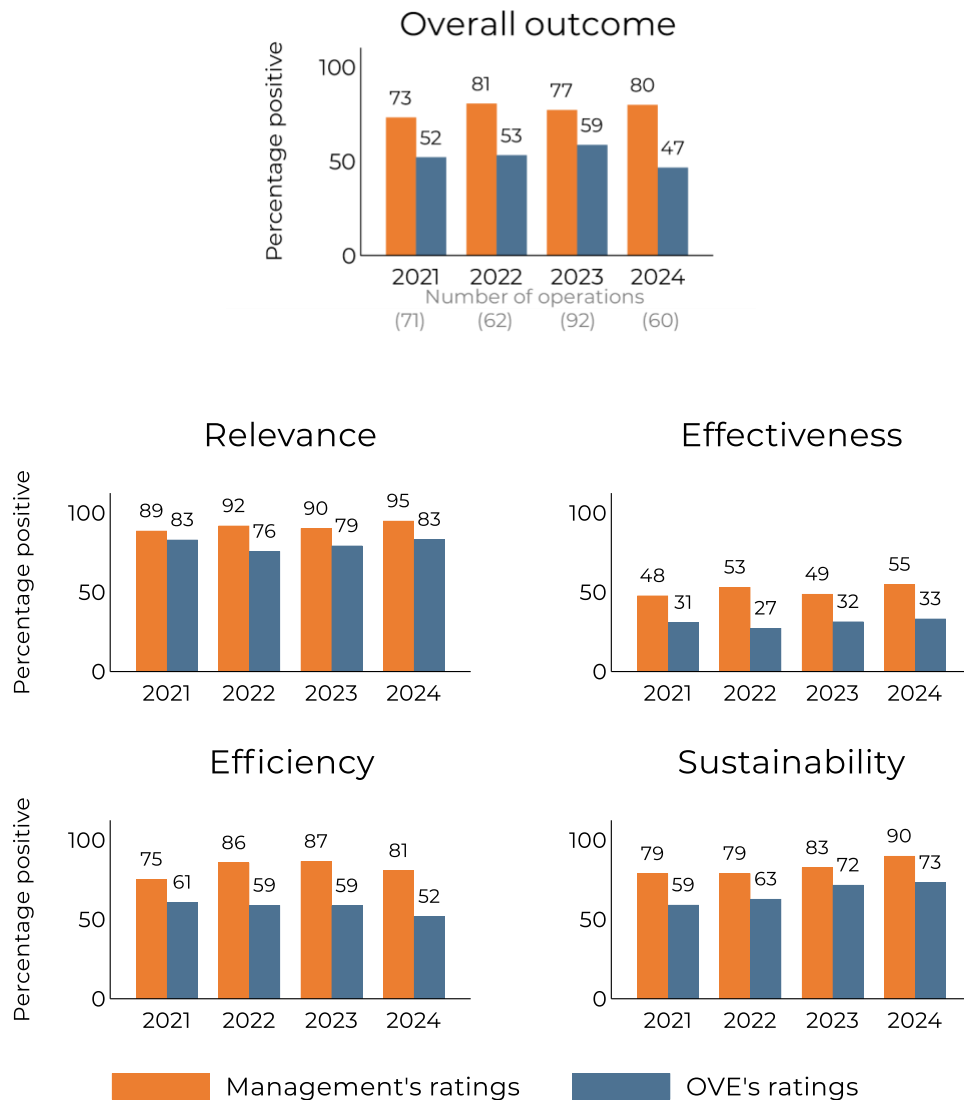
B. Differences between IDB Management’s and OVE’s ratings

4.4 **Differences between IDB Management’s and OVE’s ratings remain large.** OVE rated 47 percent of 60 operations as having a positive overall outcome, versus 80 percent rated positive by Management, a difference of 33 percentage

³¹ The recommendations provided by OVE across validation cycles consistently focus on enhancing the alignment and compliance of PCR guidelines. Initially, OVE emphasized the need to coordinate with Management on methodologies for self-evaluation, train PCR authors, and establish clear programming and submission timelines for PCRs. Over time, this focus expanded to include strengthening project performance measurement, improving the documentation and analysis of evaluation criteria, and integrating lessons learned. In later cycles, there was more emphasis on enhancing performance evaluation through improved operational guidance and quality assurance. The most recent validations report (document [RE-591](#)) emphasized improving the overall quality of PCRs. OVE’s 2024 evaluation of the DEF (document [RE-583-2](#)) recommended improving quality assurance of PCRs to ensure a sound analysis of factors explaining achievement of (or failure to achieve) project development results and to identify lessons learned for future operations.

points. This is the largest discrepancy in recent years: between 2021 and 2023, OVE rated an average of 55 percent of operations positive versus Management's average of 77 percent, a difference of 22 percentage points. Among the four core criteria, the discrepancies between OVE and Management have been largest in the efficiency and effectiveness ratings, with average differences of 25 and 21 percentage points during 2021–2024, respectively. (Figure 4.1).

Figure 4.1. Comparison of positive ratings between Management's and OVE's assessment for IDB operations over time



Source: OVE.

Notes: The graphs include only periods (2021-2024) where ratings by Management and OVE are comparable across all validated operations. During 2018-2020 Management prepared PCRs using 2014 guidelines that applied a 0–1 numeric rating scale, while OVE has used the current ordinal scales in all cycles since 2017. The number of operations per year is rated for all criteria except efficiency, given that policy-based loans (PBLs) and the combined PBL and INV are not rated for this criterion. The number of operations for the efficiency assessment is 69 in 2021, 56 in 2022, 75 in 2023, and 52 in 2024.

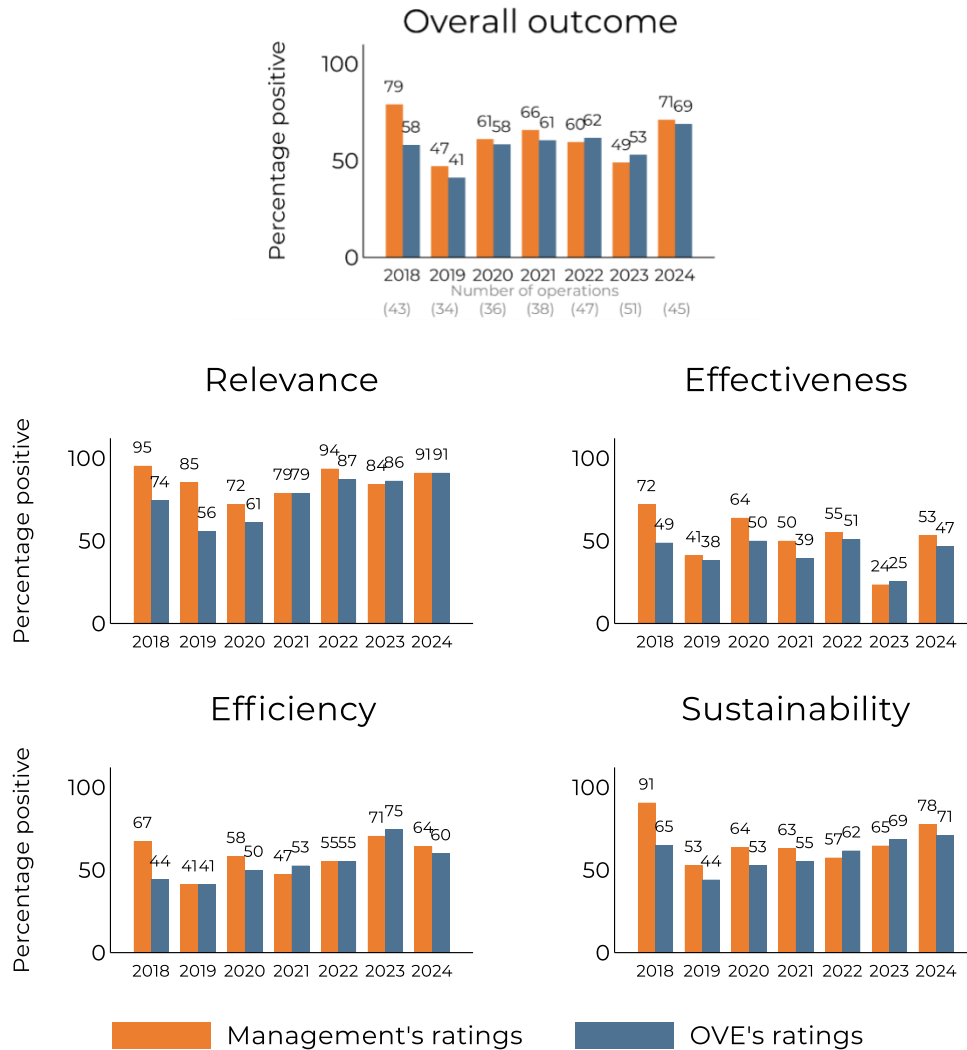
C. Quality of XSRs

4.5 **For the second year in a row, all validated XSRs were of satisfactory quality.** OVE assesses the quality of XSRs using the same rating criteria as those used for PCRs. In 2024, OVE rated the quality of all 45 XSRs delivered as positive, with 32 of them rated “good” and 13 “excellent.” Overall, XSRs closely followed the guidelines, cited sufficient data to support their findings, captured results in a balanced and complete way, and provided evidence-based ratings and lessons. Since 2019, the proportion of XSRs with a positive quality rating has been above 80 percent every year.

D. Differences between IDB Invest Management’s and OVE’s ratings

4.6 **Ratings by IDB Invest Management and OVE have been close in the past few years.** OVE’s and Management’s ratings of the overall outcomes of the 45 projects that closed in 2024 were very close, with management rating positive 71 percent of the operations and OVE 69 percent. The discrepancy in positive ratings for the overall outcome decreased from 21 percentage points in 2018 to just 2 percentage points in 2024. A similar pattern is observed for the four core criteria. In 2024, only one operation received a more critical downgrade due to Management’s inaccurate application of the XSR Guidelines. This operation shifted from a positive to negative overall outcome rating, moving from “partly successful” to “partly unsuccessful” (Figure 4.2).

Figure 4.2. Comparison of positive ratings between Management's and OVE's assessment for IDB Invest Operations over time



Source: OVE.

Note: For comparability across the years, the figure excludes SAT operations without XSRs (4 in 2022 and 2 in 2023).

V. CONCLUSIONS

- 5.1 **In 2024, OVE validated IDB's and IDB Invest's self-assessments of 105 projects for accountability and learning purposes.** Every year, IDB Group Management prepares self-evaluations of Project Completion Reports (PCRs) for closed IDB projects and of Expanded Supervision Reports (XSRs) for IDB Invest projects that have reached early operating maturity. OVE validates these self-evaluations prepared by Management to ensure that the IDB Group has used resources effectively, efficiently, and sustainably to achieve projects' approved objectives (accountability purpose) and to identify lessons to inform the design and implementation of ongoing and future operations (learning purpose). Both IDB Group Management and OVE assess project performance using the PCR and

XSR Guidelines. Projects' aggregate outcome ratings are based on the weight of 4 core criteria: effectiveness (40% weight) and relevance, efficiency and sustainability (20% weight each). Management and OVE also rate non-core criteria, including Bank and Borrower performance for IDB; and Additionality, Work Quality and Investment Performance for IDB Invest. Non-core criteria do not count towards the aggregate outcome rating. In 2024, OVE validated 60 IDB operations and 45 IDB Invest operations.

- 5.2 **OVE rated positive the overall outcome of 47 percent of the 60 IDB validated operations, which was below the corporate target. The overall outcome was driven mostly by low effectiveness and efficiency ratings.** The 47 percent overall outcome rate fell short of the 70% target set in the 2020–2023 Corporate Results Framework (CRF). OVE rated relevance positive for 83% of the validated operations (50), which had adequate designs and objectives aligned with country development needs, government priorities, and IDB Group Country Strategies. The operations that OVE rated negative (10) were mostly affected by weaknesses in their vertical logic. OVE rated effectiveness positive for 33% of the operations (20). These operations achieved more than half of their objectives with plausible attribution of outcomes. Most operations rated negative (40) were affected by a combination of unmet outcome targets (mainly of outputs) and shortcomings in reporting. Efficiency was positive for 52% of the investment operations (27). These operations included robust analysis demonstrating that the cost incurred to achieve project results was reasonable. The 25 investment operations rated negative were mostly affected by the low quality of economic analyses. Finally, OVE rated sustainability positive for 73% of the operations (44). These operations achieved results likely to continue over time and adhered to the E&S standards. The 16 negatively rated operations faced an uncertain continuity of results and subpar E&S performance.
- 5.3 **The signals raised by the IDB's alert and problem flag system in Project Monitoring Reports (PMR) were insufficient as early warnings of project performance at completion.** Many of the operations that received negative overall outcome ratings at completion were never classified in "alert" or "problem" status. The IDB's PMR classification focuses on schedule, cost, and disbursement rather than on project's progress toward achieving the projects' outcomes. Performance at closure depends however critically on outcomes achievement, a criterion that other Multilateral Development Banks do consider in their alert and problem classifications.
- 5.4 **OVE rated positive the overall outcome of 69% of the validated 45 IDB Invest operations, which was above the corporate target. The overall outcome was driven mostly by high relevance and sustainability ratings.** OVE's rating for validated IDB Invest operations exceeded the target of 65% set in the 2020–2023 Corporate Results Framework (CRF). OVE rated relevance positive for 91% of the validated operations (41). These operations had objectives that were aligned with country needs and IDB Group priorities, and they had adequate vertical logic. The operations with negative relevance ratings (4) suffered from design shortcomings, including inadequate client or instrument selection. OVE rated effectiveness positive for 47% of the validated operations (21). The operations rated positive achieved more than half of their objectives and attribution of outcomes to the operations was plausible. Negative ratings (24 operations) were explained by a combination of unmet targets and shortcomings in reporting, though the former

was more common due mainly to external factors, including the COVID-19 pandemic and market conditions. OVE rated efficiency positive for 60% of the operations (27). These operations had financial and economic benefits exceeding their costs, considering the time-value of money. The 18 operations rated negative were mostly affected by low financial or economic returns and slow portfolio growth (FI operations). OVE rated sustainability positive for 71% of the operations (32). These operations achieved results that were likely to continue over time and adhered to the E&S standards. Operations rated negative (13) faced an uncertain continuity of results and subpar E&S performance.

- 5.5 **The signals raised by the IDB Invest’s alert and problem flag system at supervision (DELTA) were good early warnings of project performance at completion.** Unlike IDB operations, most IDB Invest operations with negative ratings at completion had an “alert” or “problem” classification in the DELTA during supervision, making the DELTA a good early warning system. IDB Invest management, however, could not in most cases successfully address during project execution the signals raised by the DELTA.
- 5.6 **The analysis of 741 validated operations for the period 2018–2024, including 447 for IDB and 294 for IDB Invest, showed some patterns of performance.** The proportion of IDB operations with positive ratings in overall outcomes differs across sectors and regions but not across lending instruments. While half or more of the operations validated in the Social (SCL), Infrastructure and Energy (INE), Institutions for Development (IFD), and Integration and Trade (INT) sectors achieved positive overall outcomes on average, 28 percent (22) of the operations validated in the Climate Change and Sustainable Development (CSD) sector achieved positive overall outcomes. By region, the Southern Cone (CSC) and Central America, Haiti, Mexico, Panama, and the Dominican Republic (CID) had average positive outcome rates above 55 percent, while the Andean Group (CAN) and the Caribbean (CCB) were below 48 percent. Differences across lending instruments were, on average, minor (57 percent for PBLs vs 54 percent for investment loans). In the case of IDB Invest, the proportion of operations with positive ratings in overall outcomes differs across business segments, regions, and between pre- and post-merge-out operations. Seventy-six percent (44) of operations validated in the Infrastructure and Energy segment (IE) achieved a positive overall outcome on average, versus 58 percent (54) of the Corporate (CO) and 50 percent (72) of the Financial Institutions (FI) operations. The proportion of operations with positive ratings in overall outcomes in the Andean Group (CAN) and the Southern Cone (CSC) was above 60 percent, in Central America, Haiti, Mexico, Panama, and the Dominican Republic (CID) was 55 percent (67), while in the Caribbean (CCB) was 29 percent (2). Finally, the proportion in post-merge-out-operations was 14 percentage points higher than in pre-merge-out operations (66 vs 52 percent).
- 5.7 **The quality of IDB Invest’s self-assessment is high, while shortcomings persist for IDB.** For the second year in a row, all 45 validated XSRs were of satisfactory quality, while about half of IDB’s PCRs (30) were satisfactory. The most common shortcomings of PCRs were similar to last year’s and mostly related to the quality and completeness of the analyses in the effectiveness and efficiency sections. Management is implementing an Action Plan (AP) to improve the overall quality of PCRs with the collaboration of OVE.

- 5.8 **Ratings by IDB Invest Management and OVE have been close in the past few years, while differences between IDB Management's and OVE's ratings remain large.** Ratings by IDB Invest Management and OVE have been close in the past few years. While OVE rated the overall outcome positively in 69% of operations in 2024, IDB Invest Management did so in 71% of them. The discrepancy between IDB management and OVE remains high: while OVE rated the overall outcome positively in 47% of operations in 2024, IDB Management did so in 80% of them. Among the four core criteria, the discrepancies between OVE and IDB Management have been largest in the efficiency and effectiveness ratings.
- 5.9 **OVE's 2024 validation exercise shed light on some methodological issues that may affect management's self-assessments and OVE's validations.** Relevance and effectiveness ratings may be negatively impacted by circumstances that were unforeseeable at the time of project approval that often cannot be easily taken into account during project implementation (due, for example, to the difficulties related to project restructuring) or at the time of project evaluation (due to the self-evaluation and validation principles set out in the PCR/XSR guidelines). Consequently, projects may receive negative ratings even if they financed activities and achieved results that were relevant during implementation. Shortcomings in management's reporting (e.g. lack of data, attribution problems, inadequate indicators) can also negatively affect effectiveness rating (self-assessed or validated). When data and reporting shortcomings impede the assessment of outcomes, tagging a project as ineffective could be misleading. In this case, neither management nor OVE can affirm whether a project achieved – or did not achieve – outcomes due to lack of evidence. This also goes for efficiency. Due to the complexity of some projects in certain sectors, conducting a CBA or CEA is not always appropriate. Instead, CBAs or CEAs with strong caveats and unrealistic assumptions continue to be conducted, often resulting in OVE downgrading projects' efficiency ratings due to unreliable evidence. OVE will continue to analyze these methodological issues – and other issues that may emerge in the future – in the context of its Development Effectiveness line of work, in collaboration with Management and the IDB and IDB Invest Boards of Directors, with the aim to continue to improve its capacity to adequately fulfill its accountability and learning functions with the yearly validation exercise.

REFERENCES

- IDB (Inter-American Development Bank). 2024. *Data Marketplace. Table LEDW.ESG_ODS_OPER [dataset]*.
- 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](#)*.
- . 2024. *2024 Development Effectiveness Overview. Revised version. Document [GN-3227-1](#)*
- IDB Invest. 2016. *IDB-IIC Combined Historical Disbursements up to 2015 [dataset]*.
- . 2024a. *Azure SQL Database: iic.Maestro Transaction [dataset]*.
- . 2024b. *Azure SQL Database: metrics.Operational Approvals RPT [dataset]*.
- . 2024c. *Azure SQL Database: Maestro Project [dataset]*.
- . 2024d. *DVF's XSR Hub: Full Report for OVE 11/26/2024 [dataset]*.
- . 2024e. *Trusted Impact [dataset]*.
- 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. *IDB Group Project Performance: The 2023 Validation Cycle. Document [RE-591](#)*.
- . 2024. *Evaluation of the Development Effectiveness Framework. Document [RE-583-2](#)*.
- . 2024b. *Evaluation of IDB Support in the Area of Citizen Security in the Region. Document [RE-560-1](#)*.

Office of Evaluation and Oversight - OVE

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

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

 iadb.org/evaluation

 linkedin.com/showcase/idb-ove