Office of Evaluation and Oversight

Evaluation Synthesis Note

Comparative Analysis of Project Development Effectiveness Management Tools for Sovereign Guaranteed Operations of the AfDB, ADB, IDB, IFAD, and WB

IDB | IDB Invest

CONTENTS

This document was prepared by a team led by Claudia Figueroa, comprising Monika Huppi, Luisa Riveros, Lucero Vargas, Ingrid Garay, Melisa Wong, and Aarre Laakso with the guidance of Jozef Vaessen (OVE Principal Advisor) and Jose Ignacio Sembler (acting OVE Corporate Cluster Leader), and under the overall direction of Marialisa Motta (OVE Director). A first draft of the document was developed as background for OVE's Evaluation of the Development Effectiveness Framework, which was conducted under the guidance of Cesar Bouillon and Ana Maria Linares, and the direction of Ivory Yong Protzel.

ACRONYMS AND ABBREVIATIONS

ADB	Asian Development Bank
AfDB	African Development Bank
CRF	corporate results framework
DEF	Development Effectiveness Framework (IDB)
DEM	Development Effectiveness Matrix (IDB)
DEO	Development Effectiveness Overview (IDB)
DEPF	Development Effectiveness Policy Framework (IDB)
DMF	design and monitoring framework (ADB)
IDB	Inter-American Development Bank
IDEV	Independent Development Evaluation
IEG	Independent Evaluation Group (WB)
IEO	independent evaluation office
IFAD	International Fund for Agricultural Development
M&E	monitoring and evaluation
MDB	multilateral development bank
OECD–DAC	Organisation for Economic Co-operation and Development–Development Assistance Committee
OVE	Office of Evaluation and Oversight (IDB)
PBO	policy-based operation
PCR	project completion report
PMR	project monitoring report
SG	sovereign guaranteed operations
SI	Synthetic Indicator (IDB)
TOC	theory of change
WB	World Bank

ACKNOWLEDGMENTS

The Office of Evaluation and Oversight (OVE) team is grateful to staff of the development effectiveness offices and the offices of independent evaluation at the African Development Bank (AfDB), the Asian Development Bank (ADB), the International Fund for Agricultural Development (IFAD), the World Bank and the Inter-American Development Bank (IDB) for generously sharing their knowledge, experience and feedback. The report remains the sole responsibility of the IDB's Office of Evaluation and Oversight.

EXECUTIVE SUMMARY

This report compares the project development effectiveness management instruments for sovereign guaranteed (SG) operations of the Inter-American Development Bank (IDB) with those of four other multilateral development banks (MDBs). The IDB Office of Evaluation and Oversight (OVE) conducted this comparative analysis in the context of its 2024 evaluation of the IDB Development Effectiveness Framework (DEF) (IDB, 2024). OVE reviewed and finalized the analysis, which is presented in full in this report, as part of its 2025 Development Effectiveness evaluation line of work (IDB, 2025a), aimed at informing the ongoing institutional efforts to enhance the IDB's approach to development effectiveness. The IDB's DEF instruments at the project level include (i) an up-front assessment to enhance project evaluability through the Development Effectiveness Matrix (DEM); (ii) an instrument to report on and monitor progress during project implementation: the progress monitoring report; and (iii) a self-evaluation at project completion to evaluate project achievements: the project completion report (PCR) which is validated by OVE. The DEF instruments also include corporate reporting tools such as a corporate results framework (CRF) and the Development Effectiveness Overview (DEO). The report compares IDB's DEF instruments for SG operations with those of four comparators: the African Development Bank (AfDB), the Asian Development Bank (ADB), the International Fund for Agricultural Development (IFAD), and the World Bank.

OVE used a variety of instruments to carry out this comparative analysis. The analysis is based on a review of documents governing each institution's SG project development effectiveness management instruments, a review of a sample of these instruments (e.g., up-front project assessments, project monitoring reports, PCRs, and CRFs), as well as interviews with staff from comparator institutions' independent evaluation offices and staff from the offices responsible for development effectiveness. The report reflects the status of development effectiveness management instruments in the five institutions as of mid-2024.

The complement of tools used to manage SG project development effectiveness is similar across all five institutions, but their focus, design, and the methods they rely on vary across the institutions. At the project level, three core elements include (i) an up-front assessment of project design quality with varying focus on evaluability, (ii) an instrument to report on and monitor progress during project implementation, and (iii) a self-evaluation at project completion which is validated by the institution's independent evaluation office. At the aggregate level, information from the various project management tools feeds into corporate results monitoring and reporting. Although the tool kits are similar across the five institutions reviewed, the focus, detailed design, and methodology of each instrument vary across the institutions. Key differences include (among others): (i) the instruments used for project design and evaluability assessments; (ii) the different dimensions covered by project monitoring instruments and the criteria and methods used to classify projects into performing, potential, and problem projects; (iii) the level of approval authority for project restructuring; and (iv) the application of criteria to assess project performance at completion.

All comparator institutions assess project design at various stages of project preparation, but they use different instruments and criteria. Most institutions require that project design documents spell out the project's underlying intervention logic or theory of change (TOC) and link it to the project results framework. The specificity of these TOCs varies substantially across the institutions, with the World Bank placing more emphasis on this aspect than the other institutions. The IDB's DEM seeks to enhance project evaluability by assessing and scoring the vertical logic and evaluability of projects up-front. The DEM consists of four parts: strategic alignment, project evaluability, project risks, and IDB's additionality. A score is assigned to each evaluability dimension, which then leads to an overall evaluability score. A minimum score is required for the project to be presented to the Board. IFAD and the AfDB also have instruments that assess and score project evaluability up-front, which encompass more dimensions than the IDB's DEM. At the World Bank, the TOC and associated project results framework are reviewed, among other aspects, as part of project quality enhancement reviews, but there is no formal instrument akin to the DEM to assess and score project evaluability. The ADB's project design and monitoring framework (DMF) integrates elements of the TOC into the project results framework and is reviewed as part of quality enhancement reviews during preparation, but there is no explicit assessment of project evaluability akin to IDB´s DEM. The ADB assesses and scores certain implementation readiness aspects of infrastructure projects.

There is considerable variation in the way the five institutions monitor and assess the performance of projects under implementation. All five institutions require that their project teams periodically report on progress and performance of active projects through a well-defined reporting system of project monitoring reports (PMRs).[1] The scope of this reporting and the approach to assess and summarize project performance vary across the five institutions. All institutions' PMRs cover disbursement performance, progress on outputs, and compliance with environmental and social safeguards. All institutions except the IDB also assess performance on procurement (for investment operations), financial management and audits, and compliance with loan covenants. The AfDB, IFAD, and World Bank also assess project management and monitoring and evaluation (M&E) performance. The IDB tracks progress on specific evaluations if such have been planned and incorporated into the progress monitoring report system. All institutions include a rating of overall implementation progress, but the way this rating is derived varies considerably across the institutions. All except the ADB assess and rate progress toward achieving development objectives. The AfDB, IDB, and World Bank project monitoring systems also include a risk assessment. IFAD also assesses progress toward various institution-specific mandates.

All five institutions draw on performance indicators from PMRs to identify underperforming projects, but the factors taken into consideration for this purpose vary across the institutions. All institutions have established a three-point project classification system that separates performing projects from potential and actual problem projects. The ADB, AfDB, and IFAD project classification systems draw on a broad set of PMR ratings and criteria. The World Bank has the most encompassing system. It takes into consideration an extensive list of ratings from PMRs and other factors. The IDB classifies projects based on a system-generated indicator (called the Synthetic Indicator or SI). The SI reflects whether outputs are delivered within the expected budget and timeframe, as well as project disbursement performance. Unlike the classification system of most comparators, the SI does not take into consideration progress toward development objectives, nor other aspects taken into account by comparators such as project management, procurement, financial management, environmental and social safeguards or legal covenants. OVE evaluations have found that the IDB's project performance classification during implementation is not a good predictor of a project's rating at completion, as reflected in substantial disconnects between progress monitoring

---

1 The various institutions use different terminology for their monitoring reports. This report uses "project monitoring report" (PMR) as a generic term.

report ratings during implementation and project effectiveness and performance ratings at completion. The World Bank and IFAD monitor how promptly action is taken to resolve issues affecting projects in problem status.

All five institutions have policies and procedures in place that allow to restructure an underperforming project at the request of the borrower. Except for the IDB, all comparator organizations have updated their restructuring policy or procedures within the last several years. These updates have generally been undertaken with dual objectives: (i) to introduce more flexibility and streamline restructuring procedures to encourage proactive restructuring when an operation is off track; and (ii) to clarify criteria for restructuring and associated approval authority. All five institutions differentiate between two levels of restructuring, depending on the changes involved. The first level of restructuring involves changes in project designs that require approval by the respective institution's Board of Directors, while the second level involves changes for which the approval authority has been delegated to various levels of the institution's management. The World Bank's Board has delegated the most wide-ranging approval authority to management, including approval of modifications of the project development objectives. To incentivize early action on underperforming projects and ensure that projects are not simply restructured to reduce results targets in line with achievements toward the end of the project, the World Bank uses a split rating system when evaluating the performance of projects at completion. The split rating system measures project performance against the original project design as well as against the restructured design and weighs the overall performance by the share of disbursements before and after restructuring. The ADB, AfDB, and World Bank require that management regularly report to the Board on restructured projects.

All five institutions require the preparation of a self-evaluation at project completion and draw on the Organisation for Economic Co-operation and Development–Development Assistance Committee (OECD-DAC) criteria to evaluate their projects, but the number of criteria used and how they are applied vary across the five institutions. The ADB, AfDB, and IDB assess four core criteria: relevance, effectiveness, efficiency, and sustainability, and then derive an overall project performance rating from them for investment projects. The ADB and AfDB use the same core criteria for policy-based loans, whereas the IDB only uses relevance, effectiveness and sustainability for policy-based loans. The World Bank uses relevance, effectiveness and efficiency to derive an overall investment project performance rating and relevance and effectiveness to derive the overall performance rating for policy-based loans. It discusses risks to the continuation of development results (akin to sustainability) as an unrated non-core criterion. IFAD supplements the four OECD-DAC criteria (relevance, effectiveness, efficiency, and sustainability) with IFAD-specific criteria related to its institutional mandate, including innovation, gender equality and women's empowerment, scaling up, environment and natural resource management, and adaptation to climate change. IFAD's overall project performance rating is then derived from the assessment and ratings of each of the criteria.

The aspects covered under the relevance criterion vary somewhat across the five institutions. The AfDB, IDB, and IFAD use a fairly similar approach to assessing relevance, looking at the relevance of project objectives in view of country development needs, institutional and country strategy priorities, as well as the relevance of project design given aspired development objectives (vertical logic). They also look at whether project development objectives and design remained relevant throughout implementation or were adjusted to remain relevant to changing circumstances. The ADB supplements the aspects applied by the AfDB, IDB, and IFAD with an assessment of whether projects had any innovative effects and showed potential for scaling up. The World Bank limits the relevance assessment to examining the extent to which project development objectives align with the institution's country strategy priorities for investment loan operations. For policy-based operations (PBOs), the World Bank undertakes an assessment of the relevance of prior actions (policy conditions).

Although all institutions include an assessment of the extent to which aspired outcomes and objectives were achieved when evaluating the effectiveness criterion, how they assess the criterion varies across the five institutions. When rating project effectiveness, the IDB, IFAD, and World Bank focus on the achievement of intended outcomes and objectives as measured by the indicators in project results frameworks. The ADB and AfDB also include an assessment of outputs against their targets in their effectiveness rating. The ADB allows scoring effectiveness solely based on outputs when outcome data is missing but outputs were achieved, and the evaluation can provide a credible justification of why the outputs can be expected to have led or lead to the aspired outcomes. The ADB furthermore includes an assessment of the results of implementing the environmental and social safeguards measures and gender action plan in its effectiveness assessment. The IDB, IFAD, ADB, and AfDB have clearly defined numeric cut-offs (percent of targets achieved) for each effectiveness rating score. The World Bank provides general guidance on effectiveness ratings, allowing for more evaluator judgment when assigning the ratings for investment operations but not for PBOs.

The approach to measuring efficiency also varies across the five institutions, while assessment of sustainability is fairly uniform across the institutions that use this criterion. All institutions strive to include results of an economic analysis (cost-benefit analysis or cost-effectiveness analysis) as a measure of efficiency for investment operations. The IDB downgrades efficiency when no credible economic analysis is presented, but such an analysis was deemed feasible. The World Bank downgrades efficiency if such an analysis is lacking in the PCR but was included in the project appraisal document. The other institutions include additional criteria on efficiency (e.g. process efficiency, implementation timeliness, unit costs). The assessment of sustainability is fairly uniform across the four institutions which include this aspect among the core project evaluation criteria (ADB, AfDB, IDB, IFAD).

All five institutions derive the overall project performance (or outcome) rating from ratings assigned to the core criteria, but there is variation in how the core criteria ratings are aggregated to derive the overall outcome rating. The AfDB and IDB give higher weight to the effectiveness rating than to other core criteria when deriving the overall outcome rating. The World Bank does so for investment projects. The ADB and IFAD calculate the simple average of the core criteria ratings to determine the overall performance rating. The IDB calculates a weighted average of the core criteria ratings and supplements that with two additional decision rules. The AfDB and World Bank use a decision tree to derive the overall project performance rating from core criteria ratings.

The independent evaluation offices of all five institutions validate project self-evaluations. All evaluation offices except that of the AfDB validate 100 percent of project self-evaluations. The AfDB's evaluation office is required to validate a sample of at least 65 percent, but depending on the number of self-evaluations, it validates a higher share. Other than the IDB's OVE, all evaluation offices also carry out a limited number of independent project evaluations each year. The independent evaluation offices of all institutions report validated project ratings to the Board annually. The ADB, AfDB, and IDB also report validated ratings in management's annual development effectiveness report, which presents updates on each institution's CRF. The World Bank does so in its Organizational Effectiveness and Efficiency Dashboard.

Project results monitoring in all five institutions is part of a broader system of results monitoring. All five institutions monitor results at the corporate, portfolio and project level. All five institutions have a corporate results monitoring system that incorporates information on project performance and results, drawing on project performance ratings from PMRs and validated ratings of PCRs. IFAD and AfDB also report on quality at entry scores of newly approved projects. The ADB reports on the share of design- and procurement-ready infrastructure projects in its 2019–24 CRF. The IDB does not report on project DEM scores.

Including project and portfolio performance indicators in the CRFs can help focus managerial attention on project performance. Interviews with staff from comparator organizations carried out for this background paper suggest that when CRFs contain targets for project and portfolio performance and results delivery which are filtered down to the level of unit managers, such indicators are taken seriously. However, interviews also suggest that targets can lead to overly optimistic reporting and ratings in an effort to meet performance targets.

The governance structure for development effectiveness management is similar in all five organizations. In all institutions, the development of policies and guidelines governing quality at entry, project monitoring, and, in most cases, self-evaluation is the responsibility of a central department with input and feedback from operational departments. At the ADB, developing guidelines for PCRs is the responsibility of the independent evaluation office. In the other institutions, these guidelines are developed by management with inputs or feedback from the independent evaluation department. The AfDB's and IFAD's scored quality at entry assessments, which include an evaluability assessment, are coordinated by a central department that seeks inputs from various specialists such as financial management, safeguards and procurement, and sector specialists. At the IDB, the DEM is prepared by project teams and reviewed by a central department (Office of Strategic Planning and Development Effectiveness). At the World Bank, project results frameworks and the associated TOC are reviewed as part of the overall project quality assurance process at regional level, with inputs from regional M&E specialists. Similarly, at the ADB, the DMF is reviewed as part of the regular quality assurance process by management with inputs or feedback from the independent evaluation department. The preparation of PMRs and PCRs is the responsibility of project teams in all institutions, with managerial approvals at varying levels.

Reporting on project and institutional development effectiveness is similar across the five institutions. PMRs are made public, at least in part, by all institutions except the AfDB. The performance of the active portfolio is reported to the Board by all institutions. Project self-evaluations are shared with the Board for information and made public by all institutions, except at IFAD, where they are not shared with the Board prior to being made public. A report summarizing the results of validations of project self-evaluations is produced by the independent evaluation office of all institutions and sent to the Board annually. The Board discusses it in all institutions except the AfDB. All institutions produce annual updates of their CRFs, generally in the context of a development effectiveness review, which is also sent to and discussed by the Board. The Board of the ADB discusses the Annual Portfolio Performance Review and the Development Effectiveness Review simultaneously and the independent evaluation office's Annual Evaluation Review, which summarizes among other topics the results of that year's validations, within a few days of management’s two reports, therefore allowing for some checks and balances on management’s reporting.

I. INTRODUCTION

1.1 **The Inter-American Development Bank (IDB) introduced its Development Effectiveness Framework (DEF) to improve the development results of its operations.** In 2008, the IDB adopted the DEF "to increase the effectiveness of all of the Bank's products through: (i) setting clear standards and metrics for the evaluation of all development interventions; (ii) providing clear guidance to staff about analytical requirements for meeting the standards; (iii) aligning governance structures to comply with those set out as good practice standards; and (iv) establishing a results framework incorporated in the Corporate Performance Framework to monitor progress in key development effectiveness indicators" (IDB, 2008).

1.2 **The IDB's DEF encompasses three core project-level development effectiveness instruments that also provide information for the IDB's corporate results framework (CRF).** The DEF of the IDB encompasses three core project-level development effectiveness instruments: (i) an up-front assessment to enhance project evaluability through the Development Effectiveness Matrix (DEM); (ii) an instrument to report on and monitor performance during project implementation: the progress monitoring report; and (iii) a self-evaluation at project completion to evaluate project achievements: the project completion report (PCR). The latter is validated by the Office of Evaluation and Oversight (OVE), which independently assesses the project's achievements. The DEF also includes aggregate reporting tools: the CRF and the Development Effectiveness Overview (DEO). The CRF monitors the progress and results of the IDB Group's institutional strategy through a set of indicators and targets established every four years, including two indicators related to DEF instruments. The DEO is the IDB Group's annual report on development effectiveness. It reports on the progress on CRF indicators and includes lessons learned, results, and findings from PCRs and other sources of information. These reporting tools inform the Board and other stakeholders on the results of IDB operations at the aggregate level.

1.3 **The IDB aims to enhance its approach to development effectiveness.** After the Barranquilla Resolution of 2021 and the Washington Resolution of 2022,[2] the IDB's Board of Governors mandated the IDB's Board of Executive Directors to direct management to conduct in-depth analyses and initiate reforms to enhance development effectiveness, among other topics. In this context, IDB management is currently revising its approach to development effectiveness and the IDB Board of Directors has approved a new Development Effectiveness Policy Framework (DEPF) in early 2025.

1.4 **To help inform the ongoing discussion on the IDB's approach to development effectiveness, this report presents the results of a comparative analysis of the IDB's DEF as it applies to sovereign guaranteed (SG) operations to that of four other multilateral development banks (MDBs).** In 2024, OVE delivered the Evaluation of the DEF (IDB, 2024). As part of this evaluation, OVE undertook a review of project development effectiveness management instruments for SG operations in four comparator MDBs. The evaluation summarized key findings of this analysis. In the context of OVE's

2 Documents AG-7/21, AG-7/22.

discussions with the Board on OVE's 2025-2026 work program, it was agreed that OVE will complement the DEF evaluation with further evaluation work to inform the ongoing institutional efforts to enhance the IDB's approach to development effectiveness. As part of this line of work, OVE revised and finalized the review of the comparative analysis that this report presents in full. The four MDBs taken as comparators are the African Development Bank (AfDB), the Asian Development Bank (ADB), the International Fund for Agricultural Development (IFAD), and the World Bank.[3] The report is based on (i) a review of documents governing each institution's SG project development effectiveness management instruments, (ii) a review of a sample of these instruments (e.g., up-front project assessments, project monitoring reports, project self-evaluations and validations, CRFs), and (iii) interviews with staff from comparator institutions' independent evaluation offices and staff from the offices responsible for development effectiveness. The information was collected during the last quarter of 2023. Feedback from the comparator institutions was obtained during the second half of 2024, after which the report was revised and finalized. The report reflects the status of development effectiveness management instruments in the five institutions as of mid-2024.[4]

1.5 **This report is divided into five chapters.** Following this introduction, chapter B describes the development effectiveness management instruments used by the institutions at the project level. Chapter C explains how project results are monitored at the corporate level. Chapter D describes the governance associated with development effectiveness. Finally, Chapter E provides conclusions from the analysis.

[3] These four institutions were selected based on two criteria: (1) their independent evaluation offices of the institutions are members of the Evaluation Cooperation Group, and (2) the institutions have a consolidated development effectiveness performance management system focused on sovereign guaranteed operations.

[4] This report merely describes—and does in no way constitute an evaluation of—the instruments and aggregate reporting tools of the four institutions taken as comparators to the IDB.

II. PROJECT DEVELOPMENT EFFECTIVENESS MANAGEMENT INSTRUMENTS

2.1 **The complement of project development effectiveness instruments is similar across all five institutions but their focus, design, and the methods they rely on vary across the institutions.** The three core instruments that all five institutions use to manage project development effectiveness are: (i) an up-front assessment of project design quality with varying focus on project evaluability, (ii) an instrument to report on and monitor progress during project implementation, and (iii) a self-evaluation at project completion that is validated by the independent evaluation office. All five institutions use information from the project development effectiveness instruments to inform aggregate corporate results monitoring and reporting (for example, in their corporate results frameworks, CRFs). While the overall development effectiveness tool kit is similar across the five institutions, the focus, design, and methods applied to each instrument vary across the institutions. Key differences include: (i) the instruments used for project design and evaluability assessments; (ii) the different dimensions covered by project monitoring tools and the criteria and methods used to classify projects into performing, potential and actual problem projects; (iii) the level of approval authority for project restructuring; and (v) the application of criteria to assess project performance at completion. These and other differences are discussed in more detail in the rest of the report.

A. Project design and evaluability assessments

2.2 **All five institutions require that, at the design stage, each project have a clearly spelled out results framework, and most also call for a theory of change (TOC) that outlines the project intervention logic.** The African Development Bank (AfDB), the Asian Development Bank (ADB), the International Fund for Agricultural Development (IFAD), and the World Bank all require that, at the design stage, project documents spell out the project's underlying intervention logic, or TOC. The Inter-American Development Bank (IDB) is also in the process of introducing this requirement. The specificity of these TOCs varies substantially across the institutions, with the World Bank placing more emphasis on this aspect than the other institutions. The TOC is complemented by the project results framework which is needed to assess project development effectiveness and forms an integral part of the Board-approved loan documents.

2.3 **All five institutions review the quality of project design and evaluability at various stages of project preparation, but they use different instruments and criteria for their assessments**. At the IDB, this instrument is the Development Effectiveness Matrix (DEM), which seeks to enhance evaluability at design. While the quality assurance process of all four comparators includes a review of the project intervention logic and the results framework, only IFAD and the AfDB use an instrument similar to the IDB's DEM that assesses and scores the vertical logic and evaluability of projects up-front, though their instruments cover more aspects than IDB's DEM.

2.4 **The central focus of the IDB's DEM is project evaluability.** The IDB's DEM consists of four parts: a section on strategic alignment, a section on project evaluability, one on project risks, and one on IDB's additionality. In addition to reviewing the project's alignment with country development challenges and IDB's strategic priorities, the alignment section also reviews whether the project will contribute to any IDB development results monitored in the CRF. The evaluability section is the only DEM part that is scored numerically. It includes an assessment of the project logic, the quality of the economic analysis, and the monitoring and evaluation (M&E) framework and associated M&E plans. The project logic assesses the quality of the diagnostic, the proposed solutions, and the results matrix. The latter assesses, for example, whether each specific objective has one or more associated results achievable by project closing, and whether means of verification and data sources are clearly identified for each indicator. The economic analysis assessment reviews the results and quality of the cost-benefit or cost-effectiveness analysis. Finally, the monitoring and evaluation section assesses whether outputs have physical and financial targets; whether there is an evaluation plan to assess the effectiveness of the operation; and several elements regarding the relevance, methodology, and quality of the proposed evaluation. A score is assigned to each evaluability dimension which then leads to an overall evaluability score. The DEM score ranges from 0 to 10, and only projects with a score of 5 or higher can be submitted for Board approval.

2.5 **The AfDB's recently revised quality at entry assessment is broader than the IDB's DEM.** The AfDB's quality at entry assessment covers 13 criteria grouped into four dimensions: (i) strategic readiness, (ii) results readiness, (iii) implementation readiness, and (iv) cross-cutting priorities. Strategic readiness looks at strategic alignment, policy compliance, and adherence to environmental and social safeguards. The results readiness dimension is similar to the IDB's evaluability assessment and includes criteria such as project rationale and integration of prior lessons, quality of the results framework and the economic and financial analysis and positive effects not directly related to the purpose of the project (e.g., catalytic role or financial leverage, capacity building.). Implementation readiness assesses the adequacy of the implementation capacity and of the monitoring and evaluation capacity, adherence to financial management standards and compliance with procurement standards. Cross-cutting issues check whether a fragility lens has been used where applicable and whether climate change and gender have been mainstreamed into project design. Each criterion and dimension other than strategic readiness[5] is rated on a four-point scale guided by a detailed list of questions. The ratings of each of the four dimensions are informed by but not arithmetically derived from the ratings of criteria underneath it. No overall readiness score is assigned, nor is there a minimum threshold score that a project needs to pass to be submitted for Board approval.[6] However, if a project scores two or less (on a four-point scale) on any one dimension, a flag is assigned to that dimension. Since each project goes through the readiness review several times over the preparation cycle, flags indicate where design improvements are needed, though there is no requirement to eliminate all flags before a project is submitted for Board approval.

2.6 **IFAD's quality of design assessment instrument is also broader than the IDB's DEM.** IFAD's quality of design assessment tool, called the DEM+, comprises three sections based on the three divisions responsible for the reviews. The first section is focused on the project approach and technical and operational aspects, including an assessment of the project's strategic alignment and TOC, the likelihood of the project achieving its development objective, the results of the

---

5 Strategic readiness and the associated criteria are rated yes/no.

6 AfDB recently eliminated the average readiness score, as it found that the average obfuscated design problems, since a project could still receive an acceptable average score when it suffered from significant shortfalls in one or two areas. Instead, the flag system was introduced,

project financial and economic analysis, institutional aspects and implementation arrangements, complementary activities, exit strategy and scaling up, and implementation readiness. The second section focuses on risk mitigation, the alignment of the log frame with the TOC, the quality of the indicators, M&E planning, procurement, as well as environmental and social safeguards compliance and mainstreaming of key IFAD priorities (gender and climate change). The third section focuses on financial management arrangements. A six-point scale (from Highly Unsatisfactory to Highly Satisfactory) is used to rate most dimensions[7] and the scoring is accompanied by a qualitative "quality of design" assessment note. An overall quality of design score is assigned taking into consideration the scores of the dimensions and the quality of design assessment note. The overall quality of design score is not an average of the scores of the criteria assessed, rather, it is an overall assessment of the quality of project design. A project needs to reach an overall score of four (out of six) to be submitted for Board approval.

2.7 **The World Bank does not use a standardized quality at entry scoring instrument, but the project TOC and results framework are appraised as part of the quality enhancement reviews.** The World Bank does not use a standardized instrument like the IDB's DEM or AfDB's or IFAD's project design quality assessment instruments to check and score project evaluability or project readiness. The World Bank requires that loan documents explicitly outline a clear TOC underlying the project logic. Together with the results framework, the TOC is reviewed, among other aspects, as part of project quality enhancement reviews during the project preparation process, but there is no formal evaluability assessment per se.

2.8 **The ADB requires project design to include a project design and monitoring framework (DMF) and assesses implementation readiness for infrastructure projects.** The ADB's project DMF integrates elements of the TOC into the project results framework. The DMF is reviewed as part of quality enhancement reviews during preparation, but there is no explicit assessment of project evaluability akin to the IDB's DEM. For infrastructure projects, ADB also assesses certain implementation readiness aspects and assigns an implementation readiness score. The implementation readiness aspects include whether detailed designs for works have been completed, procurement readiness and, going forward, also financial management, institutional and safeguards readiness.

B. Project development effectiveness management during implementation

1. Project monitoring instruments

2.9 **All five institutions have a tool to monitor and report on the performance of projects under implementation, but the dimensions they cover vary considerably.** All five institutions require that their project teams periodically report on progress and performance of active projects through a well-defined reporting system, which this report refers to as a project monitoring report (PMR).[8] The scope of this reporting and the approach to assess and summarize project performance vary considerably across the five institutions (Table 2.1). The PMRs

---

7 Procurement and mainstreaming are assigned a pass/fail rating.

8 The various institutions use different terminology for their monitoring reports. This paper uses "project monitoring report" (PMR) as a generic term.

of all institutions cover disbursement performance, progress on outputs and compliance with environmental and social safeguards. The PMRs of all institutions except the IDB also assess performance on procurement (for investment operations), financial management and audits, and compliance with loan covenants.[9] The AfDB, IFAD, and World Bank also rate project management[10] and M&E performance. The IDB tracks progress on specific evaluations if such have been planned and incorporated into the progress monitoring report system. All except the ADB assess and rate progress toward or likelihood of achieving development objectives.[11] The AfDB, IDB, and World Bank PMRs also include a risk assessment.[12] In addition, IFAD assesses progress toward various institution-specific mandates (e.g., targeting and outreach, gender equity, agricultural productivity, nutrition and adaptation to climate change, potential for scaling up, human and social capital and empowerment). All institutions include a summary measure of implementation progress. The IDB assesses and reports on eight criteria. IFAD assesses and rates 27 criteria grouped into five dimensions, the ADB 16 criteria grouped into five dimensions, and the AfDB 12 criteria grouped into four dimensions. The World Bank assesses and rates seven performance criteria and eight risk dimensions. It also assesses the performance under each applicable environment and social performance standard and monitors and reports on compliance with legal covenants.

---

9 While the exact definitions across institutions vary slightly, the assessment of procurement generally covers adherence to procurement guidelines and adequacy, timeliness, efficiency and transparency of procurement processes. Financial management assessments look at the adequacy of financial management arrangements, including capacity to properly budget, execute and account for flow of funds and of internal controls. The audit dimension covers compliance with audit requirements.

10 Project management focuses on how well the implementing agency manages all aspects of project implementation and whether appropriate institutional arrangements and adequately qualified staff with sufficient authority are in place for effective and efficient project execution.

11 There is variation across the five institutions in the terminology for the development objectives progress dimension: ADB uses the term development outcomes; AfDB uses the term results progress; the IDB and IFAD use the term development objectives, and the World Bank uses the term project development objectives. This paper uses the term development objectives as a generic term when discussing the dimension in general and each institution's specific terminology when discussing how a particular institution approaches the assessment of the development objectives dimension.

12 The risk assessments of AfDB, IDB and the World Bank differ in nature. The AfDB's risk assessment is based on the risks identified in the loan documents, the World Bank assesses pre-defined risk criteria including political and governance risks, macroeconomic risks, sector strategy and policy risks, technical design of project, institutional capacity, fiduciary, environmental and social, stakeholder and "other," it then derives an overall risk rating based on the rating of each individual risk assessment. The IDB's progress monitoring report identifies specific risks associated with a particular output or outcome at the time of the report and rates them individually.

Table 2.1. Key dimensions of project progress monitoring

Key monitoring criteria	ADB	AfDB	IDB	IFAD	WB
Disbursements	Rated[a]	Rated	Rated	Rated	Reported
Procurement progress[b]	Rated	Rated	X	Rated	Rated
E&S safeguards	Rated	Rated	Rated[g]	Rated	Rated
Financial management & audit	Rated	Rated	X	Rated	Rated
Project management	X	Rated	X	Rated	Rated
Loan covenants	Reported	Rated	X[f]	Rated	Reported
M&E	X	Rated	Tracked if specific evaluations planned	Rated	Rated
Outputs	Rated	Rated	Rated (via cost performance and schedule performance indices)	Rated (as part of effectiveness)	Reported
Outcomes	Reported if feasible, after Mid-term Review	Rated	Reported	Rated (as part of effectiveness)	Reported
Institutional mandates & priorities	X	X	X	Rated	X
Development objectives progress[c,e]	X	Rated Rating derived from ratings on output progress and outcome progress	Rated Rating not derived from other criteria ratings	Rated Rating derived as weighted average of 14 criteria ratings	Rated Rating not arithmetically derived from other criteria ratings
Overall implementation progress/performance[d,e]	Rated Rating derived as average of 5 criteria ratings	Rated Rating derived as average of 10 criteria ratings	Rated Rating is system generated weighted average of disbursement, cost performance and schedule performance indices	Rated Rating derived as average of 12 criteria ratings	Rated Rating not arithmetically derived from criteria ratings, but taking into consideration performance of 10 factors
Risk assessments	X	Rated	Rated	X	Rated
Discussion of issues and actions	Yes	Yes	Yes	Yes	Yes
Mandatory updates per year	4	2	2	1 2 for problem projects	2
Public disclosure	No	Partial	Partial	yes	Partial

Source: OVE based on review of institution specific documents and interviews.

Notes: ADB = Asian Development Bank; AfDB = African Development Bank. E&S = environmental and social; IDB = Inter-American Development Bank; IFAD = International Fund for Agricultural Development; M&E = monitoring and evaluation; WB = World Bank. [a.] X means the dimension is neither reported on nor rated. The extent to which individual monitoring reports comply with reporting non-system-generated data, particularly on outcomes, varies considerably. [b.] Procurement progress not rated for Policy-Based Operation (PBOs). [c.] The terminology varies across the five institutions. The IDB, IFAD, and World Bank refer to this dimension as "development objectives," with IFAD and IDB rating "likelihood of achieving development objectives" and the World Bank rating "progress toward project development objectives." The AfDB uses the term "results progress," and the ADB uses "development outcome." [d.] Overall implementation progress/performance not rated for PBOs by the IDB. ADB does not apply project performance reports to PBOs. [e.] For further details on the derivation of the ratings on progress toward/likelihood of achieving development objectives and on overall implementation performance/progress see annex tables 1 and 2. [f.] For IDB-approved projects that have not yet reached eligibility, the progress monitoring report tracks the percentage of conditions met prior to eligibility. [g.] At IDB, the E&S safeguards rating is included in the PMR as a tracking indicator. It is not an input to the overall implementation progress indicator.

2.10 **In principle, the PMRs of all institutions include a progress update on the achievement of outputs and, in most cases, outcome indicators captured in the results framework.** All institutions expect project teams to provide in each PMR an update on progress toward achieving outputs. In principle, all institutions (to varying degrees) also expect reporting on outcome progress. A review of a small random sample of PMRs from comparator institutions and interviews with staff from those institutions suggest that both data availability and quality constraints make regular monitoring of outcomes challenging. A results focus during implementation requires clear incentives and commitment from the highest level of the institution for candid and regular results reporting. These priorities need to be clearly communicated from the top down and borrowing countries' project implementation agencies may need to be provided with support to collect and provide outcome data as needed. Interviews with World Bank operational and evaluation staff suggest that the mandatory introduction of a TOC linked to the results framework as part of project design, together with the monitoring of disconnects between ratings in PMRs and final validated ratings of completed projects, have helped make some progress in this respect. To foster regular data collection and reporting, IFAD has introduced a requirement for all projects to include a budget allocation for M&E and launched a program to strengthen in-country rural data collection and management capacity.

2.11 **All five institutions draw on performance indicators from PMRs to identify underperforming projects, but the factors taken into consideration for this purpose vary across the institutions.** All five institutions draw on PMR performance indicators to classify projects into three categories: (i) performing projects, (ii) potential problem projects and (iii) actual problem projects.[13] The factors considered to identify potential and actual problem projects vary significantly across the five institutions (Table 2.2). The IDB is the only institution that relies on a system-generated indicator that does not require any judgment by the project team (see Paragraph 2.12). The ADB, AfDB, and IFAD use a broader range of criteria and ratings from the PMR for project classification. The World Bank uses a combination of PMR ratings and other factors.

13 The classification terminology varies across the five institutions. This report uses a generic classification of performing, potential and actual problem projects.

Table 2.2. Project classification criteria

ADB	AfDB	IDB	IFAD	WB
Rating for Overall Project Performance derived as average of 5 criteria ratings (i) disbursement (ii) contract award for IPF (iii) Financial mgmt. (iv) E&S compliance (v) outputs Overall performance rating is average of 5-criteria ratings *On track*: >0.9 (out of 1) *Requires attention*: 0.9-0.65 *At risk: <0.65*	Ratings for (i) Implementation progress, which is an average of 10 criteria ratings[a] (ii) Results progress based on rating on outputs and outcomes *Non-problematic*: Satisfactory rating on both *Potentially problematic* less than Satisfactory on one *Problematic*: less than Satisfactory on both	Rating based on SI. *Satisfactory:* SI>=2.5 (out of 3) *Alert:* 2<=SI<2.5 *Problem:* 0<SI<2	Ratings for: (1) Likelihood of achieving Development Objectives (SIS 1) which is weighted average of 14 criteria ratings[b] (2) Implementation progress (SIS 2) which is average of 12 criteria ratings[a] *Performing*: SIS 1 and SIS 2 > 3.75 (out of 6) *Potential Problem*: One or both >=3.5 and <= 3.75 *Problem:* One or both indicators < 3.5	Ratings for: • Development Objectives progress • Implementation progress • E&S Standards • Procurement • Financial Management • Project management • M&E Other criteria: -Legal covenants -Effectiveness delays -Disbursement delays -Country environment -Country record • *Potential problem status:* 3 risk flags or IPF or 2 risk flags for PBO • *Problem status:* Moderately Unsatisfactory or lower rating on implementation progress or on Development Objectives Progress

Source: OVE based on review of institution specific documents and interviews.
Note: ADB = Asian Development Bank; AfDB = African Development Bank; E&S = environmental and social; IDB = Inter-American Development Bank; IFAD = International Fund for Agricultural Development; IPF = investment project financing; M&E = monitoring and evaluation; SI = Synthetic Indicator; WB = World Bank.
a. For details on criteria entering Implementation Progress Ratings of the AfDB and IFAD see Annex Table 1.
b. For details on criteria entering Likelihood of Achieving Development Objectives Progress Rating for IFAD see Annex Table 2.

2.12 **The IDB uses a system-generated synthetic performance indicator for project classification that focuses on project delays and cost overruns and does not consider progress on development objectives.** Contrary to the classification systems of the AfDB, IFAD, and World Bank, the IDB's indicator for project classification, called the Synthetic Indicator (SI), focuses on whether outputs are delivered within the expected budget and timeframe and on how the project's disbursement ratio compares to the country's historic disbursement profile, but does not consider progress on achieving development objectives. Although the IDB's progress monitoring report includes a rating on the likelihood of achieving development objectives, the rating is not driven by an explicit assessment of progress on outcomes, as it is in most comparator institutions. The IDB's SI also does not reflect other aspects taken into account by comparator institutions, such as project management, procurement, financial management, environmental and social safeguards or legal covenants. OVE evaluations have found that the IDB's project performance classification during implementation is not a good predictor of a project' rating at completion, as reflected in substantial disconnects between ratings during implementation (from the progress monitoring report) and at closure (in project completion reports—PCRs—and in Validation Reports) (IDB 2024, IDB 2025b).

2.13 **The ADB classifies projects based on an overall project performance rating that draws on the PMR ratings of five criteria.** The ADB's overall project performance rating is derived as the arithmetic mean of the PMR ratings for disbursements, procurement, safeguards and financial management and outputs. Numerical values for each project status category are then defined (Table 2.2). Like the IDB's classification, the ADB's classification does not consider progress toward development objectives, referred to as development outcomes in the ADB.[14]

2.14 **The AfDB's project classification system draws on the PMR ratings for implementation and results progress.** Both the rating for implementation progress and that for results progress (akin to the progress on development objectives in other institutions) are derived from a range of other PMR criteria ratings. The implementation progress rating is the arithmetic mean of ten criteria ratings grouped into 3 dimensions, including compliance with covenants, systems and procedures and project execution and financing. The rating for results progress is derived from the rating on outputs and outcomes. A project requires a satisfactory rating on both, implementation and results progress, to be rated

[14] ADB's PMRs are expected to provide updates and rate progress on outputs. Project Administration Guidelines also instruct teams to report progress on outcome indicators as feasible during the second half of project implementation, but there is no overall assessment of progress toward development objectives (institutionally referred to as development outcomes by ADB) akin to the assessment of progress on or likelihood of achieving development objectives in other institutions.

satisfactory. A rating of less than satisfactory on one dimension results in a project classification of “potentially problematic”, whereas a rating of less than satisfactory on both dimensions classifies the project as “problematic” (Table 2.2).

2.15 **Similar to the AfDB, IFAD’s project classification system draws on the summary ratings for the likelihood of achieving development objectives and for implementation progress.** IFAD’s implementation progress rating is derived as the average of 12 PMR criteria ratings grouped into two dimensions: (i) project management and (ii) financial management and project execution. Its rating for the likelihood of achieving development objectives, in turn, is a weighted average of 14 PMR criteria, including progress on outputs and outcomes and institutional priorities. (See Annex Table 2.) IFAD then uses a combination of the implementation progress and the likelihood of achieving development objectives ratings and defines numeric cut-off points for each rating to classify projects as performing, potential problem or problem projects (Table 2.2).

2.16 **The World Bank uses a more encompassing system to classify projects in potential or actual problem status, taking into consideration an extensive list of ratings from PMRs and other criteria**. The World Bank uses a flag system to identify projects in potential or actual problem status. The flags are defined based on: (i) PMR ratings on progress toward development objectives, implementation progress, environmental and social standards, procurement and financial management performance, and project management and M&E; and (ii) other factors, including the status of key legal covenants, effectiveness and disbursement delays, country environment and country record.[15] Ratings on implementation progress and progress toward development objectives are given a higher weight than other aspects. These two ratings are not arithmetically derived from ratings of other criteria, though project teams are instructed to take those into consideration.[16] A project which has a rating of moderately unsatisfactory or less on either one of these aspects is automatically classified as being in problem status. An investment operation that has raised three flags and a development policy loan that has raised two flags is classified as being in potential problem status. A moderately satisfactory or lower rating on environmental and social standards, procurement, financial management, project management or M&E automatically triggers a flag, as do non-compliance with key legal covenants and effectiveness and disbursement delays. A flag is also triggered in case of a disconnect above 20 percent between the self-evaluation ratings and the ratings of the Independent Evaluation Group (IEG) in the (overall) country portfolio, a high share of country portfolio commitments rated unsatisfactory, certain other country

15 Country record is measured by the percentage of net disconnects between Independent Evaluation Group (IEG) ratings and self-ratings and the share of net commitments associated with unsatisfactory projects in a given country portfolio. A disconnect rate above a certain threshold results in an automatic flag for every project in the portfolio.

16 The World Bank’s implementation progress rating is a summary assessment that takes into consideration performance of 10 factors that are assessed separately. (See Annex Table 1.) The rating is not derived arithmetically from the 10 criteria ratings, rather it is a summary rating assigned by the project team taking into consideration the performance under the 10 criteria. When assigning the rating for progress on development objectives, World Bank project teams need not only take into consideration whether the objectives are likely to be achieved based on progress on outcomes and associated risks, but also whether there are shortcomings in the relevance of the objectives and the efficiency with which they are likely to be achieved.

risk criteria, or in case a project has been rated marginally unsatisfactory or lower on implementation progress or development objective progress for any 24 months cumulative over the life of the project.

2.17 **Institutional incentives are key to ensuring that project performance reporting is candid and evidence-based.** Several evaluations (World Bank 2016, IFAD 2023a, ADB 2020, AfDB 2021) have found that project performance reporting tends to suffer from a lack of credible data, a reluctance by team leaders to report problems, or staff over-optimism. Institutions need to ensure that incentives are right for candid reporting and resources are allocated in project budgets to ensure proper data collection. To provide such incentives, IFAD is allocating more budgetary resources for the supervision of problem projects, while the World Bank monitors disconnects between the rating of progress toward development objectives in the last PMR and IEG's final project ratings,[17] as well as for disconnects of self-ratings in PCRs and IEG final ratings. Disconnects are monitored, and if they exceed the defined standard for a country portfolio, all active projects in that portfolio automatically receive a flag, which in combination with other flags, leads to potential problem project status.

2. Project restructuring[18]

2.18 **All five institutions have policies and procedures in place that allow restructuring a project at the borrower's request**. Except for the IDB, all comparator organizations have updated their restructuring policy or procedures within the last several years.[19] These updates have generally been undertaken with dual objectives: (i) to introduce more flexibility and streamline procedures to encourage proactive restructuring when an operation is off track or affected by changes (e.g., new government priorities, extreme natural events); and (ii) to clarify criteria for restructuring and associated approval authority. The IDB's procedures related to sovereign guaranteed (SG) project reformulation and modification date back to 2011, with some clarifying guidance issued in 2024.

2.19 **All five institutions differentiate between two levels of restructuring, depending on the changes involved**. The first level of restructuring involves changes in project designs which require approval by the respective institution's Board of Directors, while the second level involves changes for which the approval

17 The comparison between project development objectives ratings in the monitoring reports and final IEG project outcome ratings is facilitated by including a qualitative judgment of the continued relevance of and efficiency of progress toward these objectives in addition to outcome progress.

18 For ease of comparison, this section uses the term restructuring, although the terms used by each comparator organization vary. Similarly, to facilitate comparison this paper differentiates between level 1 and level 2 restructuring, though the terms used across institutions vary.

19 Updates by comparator institutions were undertaken in the following years: ADB 2018, AfDB 2022, IFAD 2018, World Bank 2017. By comparison, IDB's procedures date back to 2011.

authority has been delegated to various levels of the institution's management. The extent of delegation for approval of restructuring from the Board to management varies across institutions.

2.20 **The World Bank's Board has delegated the most wide-ranging approval authority to management, including approval for modifications of the project development objectives.** The only project modifications requiring approval by the World Bank's Board involve those that result in a change to the highest level of safeguards risk classification, extension of the expiration date of a Guarantee and alternative procurement arrangements. In addition, World Bank Management can at its own discretion decide to submit project changes to the Board which it considers involve substantial institutional or operational risks (including substantial changes to project development objectives). Substantial changes to the scope, development objectives, or expected outcomes in all other comparator institutions require Board approval. The ADB includes in this category any downward revisions of expected outcomes, the AfDB includes changes that affect at least 10 percent of project outputs and outcomes. IFAD and the IDB are less specific about what changes to expected project results require Board approval. IFAD specifies changes in results and targets that significantly affect the project's TOC overall. Similarly, the IDB stipulates that substantial changes in objectives or underachievement of expected results that can be expected to lead to less than satisfactory effectiveness ratings at project completion require Board approval. Like at the World Bank, project restructuring that results in a higher environmental and social safeguards risk classification requires Board approval at the AfDB and IFAD, while such changes are not explicitly singled out for Board approval at the ADB or IDB (Table 2.3).

Table 2.3. Project changes requiring Board approval

Institution	Required changes
ADB	• Restructuring that fundamentally affects the scope of approved project outcomes, including among others downward revision of expected outcomes.
AfDB	• Significant change to project design, scope, or development objectives. As a rule of thumb, a restructuring that affects at least 10 percent of the project's results (outputs and outcomes) is considered significant. • Modification that results in a reclassification of the project environmental and social categorization to Category 1, or in a change in the climate risk category.
IDB	Substantial and Fundamental Changes to operations including: • change in borrower, guarantor, guaranteed party • changes in financial terms or scope of guarantees • changes in lending category or instrument • significant exceptions to procurement policy • changes in purpose of loan or guarantee • substantial change of objectives or substantial shortfall in outcomes affecting achievement of objectives and end of project effectiveness rating • any other change deemed substantial and fundamental by management
IFAD	• Significant and substantial change in scope or characteristics of project. • Change in environmental and social safeguards classification.
World Bank	• Change from a lesser safeguard category to a Category A. • Alternative procurement arrangements. • Extension of Bank guarantee expiration date. • Any level 2 restructuring that Management decides to submit to the Board due to it involving significant operational or institutional risk (including substantial change in project development objectives).

Source: OVE based on review of institution specific documents and interviews.
Note: ADB = Asian Development Bank; AfDB = African Development Bank; IDB = Inter-American Development Bank; IFAD = International Fund for Agricultural Development.

2.21 **For restructuring with delegated approval (level 2), the approval authority within management varies across institutions**. At the ADB, the project team leader approves minor changes that do not affect the DMF, while changes affecting the DMF require approval by the Sector Director. At the AfDB, approval authority rests with the Vice President for operations who can delegate this authority to lower levels. At the IDB, the specific level of management authorized to approve a project change that is not considered substantial depends on the type of change. At IFAD, all modifications that do not require Board approval are approved at the Vice-Presidential level. At the World Bank, changes to project development objectives, and changes that trigger new environmental and social safeguards policies or involve a closing date extension of more than two years for investment lending operations require Vice-Presidential approval, while all other level 2 changes are approved by the Country Director. Changes to project development objectives for Program-for-Results Financing also require Vice-Presential approval.

2.22 **The ADB, AfDB, and World Bank all require that management regularly report to the Board on projects that were restructured**. The ADB reports in its quarterly portfolio updates to the Board on Board-approved project restructuring and on management-approved restructuring that involves a change to the DMF. The AfDB is expected to prepare bi-annual reports on all (Board and management approved) project restructurings. It also includes a section on project restructuring with a focus on strategic and operational issues in its Annual Portfolio Performance Report to the Board. The World Bank reports on level two restructurings that involve a change to project development objectives or trigger a new safeguards policy in its quarterly operations update to the Board as well as in the annual report to the Board on Operational Policy Waivers. Information on project restructurings in the World Bank is also available to Board members through the Executive Directors' Portal. The World Bank and IFAD monitor how promptly action is taken on problem projects via a proactivity index that is monitored and reported on in the CRF of IFAD and the Organizational Effectiveness and Efficiency Dashboard of the World Bank. The IDB reports on reformulations to the Board in the Operational Management Reports.

2.23 **Striking a balance between providing incentives to restructure projects when problems arise and ensuring that such restructuring is not done for the sole purpose of achieving a positive project rating is important**. While streamlining of restructuring procedures in comparator institutions has been motivated by a desire to address project underperformance proactively, easy restructuring procedures can also harbor the risk of project teams watering down the results framework toward the end of a project for the purpose of a project achieving a positive overall rating at completion. To counter this risk, the World Bank is applying a split rating method when evaluating restructured projects. Such projects are evaluated against the originally approved development objectives and associated expected results and then also against the revised results frameworks (and objectives if these were changed during project restructuring) and two outcome ratings are determined. The overall project outcome rating for the restructured project is then derived as the weighted average of both outcome ratings with the weights being the share of resources disbursed before and after restructuring. Using the share of disbursements as weights provides incentives for restructuring non-performing projects early, when relatively few resources have been disbursed. The IEG's 2023 Report on Results and Performance of the World Bank Group finds that timely restructuring has had a positive effect on project performance by limiting the impact of implementation challenges. It also finds that projects restructured early in the project cycle were more likely to achieve their intended outcomes than those that adapted later (World Bank, 2023b). None of the other comparator institutions apply this method. Instead, formally restructured projects in all other reviewed institutions are evaluated against revised results frameworks, although the AfDB defines limits above which changes in targeted outcomes need to be approved by the Board, and the ADB only allows for upward revision of targets without Board approval.

C. Evaluation of completed projects

2.24 **All five institutions require the preparation of a self-evaluation at project completion, and several have recently revised their guidelines for such evaluations**. The AfDB and IFAD have recently updated their guidelines for project evaluation. The World Bank has done so in 2017 for investment project financing and in 2022 for development policy financing. The ADB is in the process of updating its guidelines. The IDB's PCR guidelines were updated in 2020. Recent revisions by the AfDB, IFAD, and World Bank for policy-based operations have put more emphasis on the need to support evaluations with evidence and on better defining rating categories, with clearly defined quantitative bands of achievement rates for effectiveness ratings.

2.25 **All institutions draw on the Organisation for Economic Co-operation and Development–Development Assistance Committee (OECD-DAC) criteria to evaluate their projects, but the number of core criteria used and how they are applied varies across the five institutions.** The ADB, AfDB, and IDB assess four core criteria: relevance, effectiveness, efficiency,[20] and sustainability, and then derive an overall project performance rating (Table 2.4). The World Bank uses relevance, effectiveness and efficiency as core criteria to derive an overall performance rating for investment projects and relevance and effectiveness to derive an overall performance rating for PBOs. It discusses risks to the continuation of development results (akin to sustainability) as an unrated (non-core) criterion. IFAD supplements the four OECD-DAC core criteria (relevance, effectiveness, efficiency, sustainability) with specific criteria related to its institutional mandate: innovation, gender equality and women's empowerment, scaling up, environment and natural resource management and adaptation to climate change.[21] IFAD's overall project performance rating is derived from the assessment and ratings of each of these dimensions. The following paragraphs describe in more detail how each institution evaluates each core criterion and how it derives the overall project performance ratings.

2.26 **The aspects covered under the relevance criterion vary somewhat across the five institutions**. The AfDB, IDB, and IFAD use a similar approach to

---

[20] IDB and the World Bank do not assess efficiency for policy-based operations.

[21] IFAD-specific criteria are defined as follows: *innovation* measures the extent to which interventions introduced a novel solution to improve performance or address challenges related to rural poverty reduction; *scaling up* assesses to what extent solutions introduced by IFAD have been adopted or diffused by other development partners, the private sector or the government beyond IFAD's project; *gender equality and women's empowerment* looks at the extent to which IFAD interventions have contributed to greater gender equality and women's empowerment; and *environment and natural resource management and climate change adaptation* looks at the extent to which an IFAD project has contributed to enhancing environmental sustainability and resilience to climate change in small-scale agriculture.

assessing relevance, looking at the relevance of project objectives in view of country development needs, institutional and country strategy priorities, as well as the relevance of project design given aspired development objectives (vertical logic). They also look at whether project development objectives and design remained relevant throughout implementation or were adjusted to remain relevant to changing circumstances. The ADB supplements the aspects applied by the AfDB, IDB, and IFAD with an assessment of whether projects had any innovative effects and showed potential for scaling up. The World Bank limits the relevance assessment to an examination of the extent to which project development objectives are aligned with the institution's country strategy priorities for investment loan operations. For policy-based operations (PBOs), the World Bank undertakes an assessment of the relevance of prior actions (policy conditions).

2.27 **Although all institutions include an assessment of the extent to which aspired outcomes and objectives were achieved when evaluating effectiveness, how the criterion is assessed varies across the five institutions**. All institutions use the project results framework adopted at approval and subsequent formal modifications as a basis to evaluate effectiveness. The IDB, IFAD, and World Bank focus on the achievement of intended outcomes and objectives as measured by the indicators in project results frameworks. In addition to assessing the achievement of outcomes, the ADB and AfDB also include an assessment of outputs against their targets in the overall effectiveness rating. The ADB allows scoring effectiveness solely based on outputs when outcome data are missing, but outputs were achieved, and the evaluation can provide a credible justification of why the outputs can be expected to have led or will lead to the aspired outcomes. The ADB furthermore includes an assessment of the results of implementing its environmental and social safeguards measures and gender action plan in its effectiveness assessment. The IDB, IFAD, ADB, and AfDB have clearly defined numeric cut-offs (percent of targets achieved) for each effectiveness rating score. The World Bank provides general guidance on effectiveness ratings, allowing for evaluator judgment when assigning effectiveness ratings for investment loan operations but not for PBOs. To achieve a positive rating for the achievement of a development objective, the ADB, IDB, and IFAD require that, on average, 80 percent of targets were achieved, while the AfDB sets the cut-off at 75 percent. The World Bank defines no specific numeric achievement targets for investment loan projects, but it does so for PBOs, setting the cut-off for a positive (top half of the rating scale) indicator rating at 66 percent, provided that the relevance of the indicator has been rated positively.[22]

2.28 **The approach to measuring efficiency also varies across the five institutions, while assessment of sustainability is fairly uniform across the institutions that use this criterion**. The IDB and the World Bank do not assess the efficiency of PBOs, given that PBO funding goes to the general budget. The ADB and AfDB assess the efficiency of PBOs and investment loans but apply the efficiency criterion differently to the two operation types. All institutions strive to include

22 When assessing effectiveness of policy-based operations, the World Bank first rates the relevance of each indicator with respect to the objective to which it is assigned. Then it assesses the achievement rate of the indicator and rates the latter, if the indicator relevance has been rated negatively, the achievement rating is downgraded. Once all indicator achievements under an objective have been rated, the objective rating is derived as the numeric average of all indicator ratings. The overall effectiveness rating is then derived as the numeric average of all objective ratings.

results of an economic analysis (cost-benefit or cost-effectiveness analysis) as a measure of efficiency for investment operations. The IDB downgrades efficiency when no credible economic analysis is presented, but such an analysis was deemed feasible. The World Bank downgrades efficiency if such an analysis is lacking in the PCR but was included in the project appraisal document. The other institutions include additional criteria on efficiency. The ADB supplements economic analysis with an assessment of process efficiency (time and cost overruns) and allows for unit cost assessments when an economic rate of return is not presented. The AfDB supplements economic analysis with an assessment of implementation timeliness and the implementation progress rating assigned during supervision. IFAD supplements economic analysis with an assessment of program costs, unit cost comparisons and implementation timeliness. The assessment of sustainability is fairly uniform across the four institutions that include this aspect among their core project evaluation criteria (ADB, AfDB, IDB, IFAD).

2.29 **All five institutions derive the overall project performance (or outcome) rating from ratings assigned to the core criteria, but there is variation in how the core criteria ratings are aggregated to derive the overall outcome rating**. The AfDB and the World Bank use a decision tree to derive the overall project performance rating from the core criteria ratings. The ADB and IFAD[23] calculate the simple average of the core criteria to determine the overall project performance rating. The IDB calculates a weighted average of core criteria and supplements that with two additional decision rules: (i) for a project to be rated (fully) Successful, three core criteria (relevance, effectiveness, and sustainability) need to be rated Satisfactory or higher; and (ii) the maximum rating a project can achieve if it has a (fully) Unsatisfactory rating in any of these three criteria is Partly Unsuccessful. The AfDB, IDB, and World Bank[24] accord higher weight to the effectiveness rating than to the other core criteria when deriving the overall project performance rating. Contrary to the IDB, the AfDB and the World Bank do not allow for a positive overall project performance rating with a negative effectiveness rating.

Table 2.4. Evaluation criteria to calculate overall project ratings at completion

	Core criteria	Non-core criteria	Rating scale for the criteria	Overall project rating
ADB[a]	• Relevance • Effectiveness • Efficiency • Sustainability	• Development impact • Bank and cofinancier performance • Borrower performance	Four-point scale for core and non-core criteria	Four-point scale: Highly Successful, Successful, Less than Successful, Unsuccessful. Derived as arithmetic mean of the 4 core criteria. For a positive[c] (Successful or higher) rating, the average score needs to be 2.5 or higher (out of 4).

23 IFAD's recently updated project completion report (PCR) guidelines state that the overall performance rating has to be consistent with core criteria ratings but does not need to be a strict arithmetic mean. However, the Independent Evaluation Office (IEO) applies an arithmetic mean when deriving the overall project performance rating in its validations of PCR ratings.

24 The World Bank gives equal weight to the relevance and effectiveness ratings when deriving the overall rating for policy-based operations.

	Core criteria	Non-core criteria	Rating scale for the criteria	Overall project rating
AfDB (2023)[b]	• Relevance • Effectiveness • Efficiency • Sustainability • Coherence is assessed qualitatively but not rated	• Bank Performance • Borrower Performance	Four-point scale for core and non-core criteria	Six-point rating scale: Highly Successful, Successful, Mostly Successful, Mostly Unsuccessful, Unsuccessful, Highly Unsuccessful. Derived from 4 core criteria ratings based on a decision tree, with more weight given to effectiveness than to the other criteria, as shown by the rules below: - For a positive rating (Mostly Successful or higher) at least 3 criteria ratings need to be Satisfactory or higher, including relevance and effectiveness. - With an Unsatisfactory effectiveness rating, the overall project rating is Highly Unsuccessful.
IDB	*Investment Loans*: • Relevance • Effectiveness • Efficiency • Sustainability *Policy-Based Loans*: • Relevance • Effectiveness • Sustainability	• Bank Performance • Borrower Performance	Four-point scale for core and non-core criteria	Six-point scale: Highly Successful, Successful, Partly Successful, Partly Unsuccessful, Unsuccessful, Highly Unsuccessful. *Investment Loans*: Overall project rating is calculated as weighted average of the 4 core ratings. Weight of effectiveness rating is 40 percent, weight of other 3 ratings is 20 percent each. *Policy-Based Loans*: Overall project rating is calculated as weighted average of the 3 core ratings. Weight of effectiveness rating is 60 percent, weight of other 2 ratings is 20 percent each. For a positive rating (Partly Successful or higher) the average score needs to be 4 or higher, but additional rules apply: (i) for a Successful overall project rating, relevance, effectiveness and

	Core criteria	Non-core criteria	Rating scale for the criteria	Overall project rating
				sustainability need to be Satisfactory or higher. (ii) with an Unsatisfactory rating for relevance, effectiveness or sustainability, the highest possible overall project rating is Partly Unsuccessful.
IFAD	• Relevance • Effectiveness • Efficiency • Sustainability • Innovation • Gender Equality & women's empower-ment • Scaling Up • Environment and natural resource management • Adaptation to climate change	• Borrower performance • IFAD performance	Six- point scale for core and non-core ratings	Six-point scale: Highly Satisfactory, Satisfactory, Moderately Satisfactory, Moderately Unsatisfactory, Unsatisfactory, Highly Unsatisfactory. Per PCR guidelines, the overall rating should be consistent with individual criteria ratings, but it does not need to be strictly an arithmetic mean. The independent evaluation office applies an arithmetic mean in PCR validations.
World Bank	*Investment Loans:* • Relevance of objectives with respect to country strategy • Effectiveness • Efficiency *Policy-Based Loans:* • Relevance of prior actions (policy conditions) • Effectiveness	• Risk to development outcome (discussed/ not rated) • Bank performance • Relevance of results indicators (PBOs), M&E quality (IL)	*Investment Loans*: • Core criteria: four-point scale • Bank performance: six-point scale • Quality of M&E: four-point scale *Policy-Based Loans*: • Core criteria: six-point scale. • Relevance of indicators: six-point scale Bank Performance: six-point scale	Six-point scale: Highly Satisfactory, Satisfactory, Moderately Satisfactory, Moderately Unsatisfactory, Unsatisfactory, Highly Unsatisfactory. *Investment Loans*: Overall project outcome rating derived from 3 criteria ratings based on a decision tree. Higher weight accorded to effectiveness rating. For a positive rating (Moderately Satisfactory or higher) the effectiveness rating has to be positive (top half of the rating scale). *Policy Based Loans*: Overall project rating derived from the 2 core criteria based on a decision tree. For a positive rating (Moderately Satisfactory or higher), effectiveness and relevance need to have a positive rating (top half of rating scale).

Source: OVE based on review of institution specific documents and interviews

Note: ADB = Asian Development Bank; AfDB = African Development Bank; IDB = Inter-American Development Bank; IFAD = International Fund for Agricultural Development; M&E = monitoring and evaluation; PBO = policy-based operation; PCR = project completion report.
a. The ADB is revising PCR guidelines. The above presentation is based on current guidelines. Separate guidelines for PBOs will be developed, with PBO relevance assessment likely to focus on the relevance of prior actions and the relevance of vertical logic from prior actions to outcomes.
b. The AfDB revised the approach and criteria for project evaluation in 2023. The criteria and ratings depicted above reflect the new guidelines, which are, however, not yet in use.
c. "Positive rating" is defined here as a rating in the top half of the rating scale.

2.30 **Although all five institutions have overall clearly defined approaches to evaluating project performance, evaluations have found that self-assessments are often affected by a lack of credible data and candor.** IFAD's 2023 Annual Review of Independent Evaluations raises concerns about the validity of output and outcome data presented in PCRs given limited capacity of project implementation units to collect such data, particularly for surveys (IFAD 2023b). Similarly, the AfDB's evaluation of the Self-Evaluation System (AfDB 2020) found that the main weakness of the system was in the application of established procedures, standards and norms. It pointed to three factors that affected the proper functioning of the AfDB's self-evaluation system, namely (i) low compliance with established procedures; (ii) limited resources for M&E during supervision; and (iii) a deficiency in candor and a positive bias in assessing performance. In response to these findings, the AfDB has updated its PCR guidelines to introduce more clarity on the supporting evidence needed and more clearly define ratings categories. The ADB's Annual Evaluation Review 2020, which also focused on the institution's self-evaluation system, concluded that the system is affected by weaknesses in the M&E systems during implementation (ADB 2020). It further found that reliability of the data and information generated by the ADB project self-evaluation system was affected by the lack of incentives to produce realistic and candid completion reports. It ascribed this to the ADB's system tending to penalize failure rather than recognize causes and encourage learning. To address these issues, the ADB has recently updated its project administration instructions related to project monitoring and restructuring and is revisiting its PCR guidelines. The IEG's evaluation of the World Bank's self-evaluation system (World Bank 2016) also pointed to M&E and data weaknesses. More recent evaluation work suggests that improvements in M&E since then have had a positive effect on project adaptation and the demonstrated results of World Bank operations, though further challenges on outcome orientation persist, particularly related to measuring results of capacity building efforts (World Bank 2023).

2.31 **All five institutions consider that project self-evaluations should fulfill a dual function of accountability and learning.** To this effect, each institution's PCR is expected to include lessons. However, evaluations of the self-evaluation system at the World Bank and AfDB have found that self-evaluations are generally seen by staff as accountability rather than learning tools (World Bank 2016, AfDB 2020), pointing out that insufficient efforts are made to consistently use the information generated for institutional learning. The ADB's, AfDB's, and IFAD's independent evaluation offices maintain databases with lessons that can be searched by project team leaders when designing new operations. Interviews suggest, however, that the quality of lessons is often low, limiting their usefulness.

D. Independent validation

2.32 **To enhance the credibility of the project evaluation system, the independent evaluation offices of all five institutions validate project self-evaluations.** All evaluation offices except that of the AfDB validate 100 percent of project self-evaluations. The AfDB is obliged to validate a sample of at least 65 percent of projects, but depending on the number of self-evaluations, it validates a higher share. All evaluation offices, other than the IDB's, also undertake a limited number of in-depth independent project evaluations each year (Table 2.5). The independent evaluation offices of all institutions report validated project ratings to the Board annually. The ADB, AfDB,[25] and IDB also report validated ratings in their management's annual development effectiveness reports and the World Bank does so in its Organizational Effectiveness and Efficiency Dashboard.

Table 2.5. Coverage and use of validations and independent in-depth project evaluations

	Validation coverage of self-assessed projects	Number of independent project evaluations	Use of validated ratings	Reporting of ratings disconnects
ADB	100 percent	4-5 per year	• Annual Development Effectiveness Review • CRF • IEO Annual Evaluation Review	• IEO variance memo
AfDB	65 percent-100 percent	No specific number, but are conducted as part of project cluster evaluations	• Annual Development Effectiveness Review • CRF • IDEV validation synthesis (sent to the Board for information only)	• IDEV validation synthesis
IDB	100 percent	None	• DEO • OVE's Report on project validations • CRF	• OVE report on project validations
IFAD	100 percent	4-5 per year	• Limited reporting in CRF • IEO Annual Report on Independent Evaluations	• IEO Annual Report on Independent Evaluations
World Bank	100 percent	About 20 percent of projects	• Organizational Effectiveness and Efficiency dashboard • Public project performance database • IEG RAP of the World Bank Group	• IEG RAP • Portfolio dashboards.

Source: OVE based on review of institution specific documents and interviews
Note: ADB = Asian Development Bank; AfDB = African Development Bank; CRF = corporate results framework; DEO = Development Effectiveness Overview; IDB = Inter-American Development Bank; IDEV = Independent Development Evaluation; IEG = Independent Evaluation Group; IEO = independent evaluation office; IFAD = International Fund for Agricultural Development; OVE = Office of Evaluation and Oversight; RAP = Results and Performance.

2.33 **Validating project self-evaluations provides incentives for candid, evidence-based reporting and evaluation**. One tool to provide incentives for honest

25 The AfDB's Annual Development Effectiveness Report reporting on project performance has not consistently been done based on validated ratings, but the 2023 ADER (AfDB 2023b) states that it is reporting validated ratings.

reporting is measuring and monitoring the disconnect between self-ratings of project performance and validated ratings. Although the independent evaluation offices of all five institutions measure the disconnect between self-ratings and validated ratings, how this information is reported and used varies. The AfDB, IDB, IFAD, and World Bank evaluation offices report on disconnects in their annual reports on the validation exercise to the Board. The validation report of the ADB's evaluation office does not include this information. However, management's Development Effectiveness Report and its Annual Portfolio Performance Review are discussed almost concurrently (within a few days) with the evaluation office's Annual Evaluation Review, providing the Board with an opportunity to ask questions on performance differences reported in the Annual Evaluation Review and management's own reports (particularly the Annual Portfolio Review, which focuses on the performance of the active portfolio). The ADB's evaluation office also prepares a variance memo that outlines the disconnects, which it discusses with management. Reporting on disconnects alone provides limited incentives for candid and credible reporting in self-evaluations if there is no follow-up on disconnects and if self-evaluation is not supplemented by independent evaluation. While self-evaluation constitutes an essential component of multilateral development banks' (MDBs') accountability and learning ecosystems, independent evaluation is warranted to mitigate its inherent limitations.

2.34 **The World Bank and IFAD put more emphasis on ratings disconnects than the other institutions reviewed**. Although IFAD does not include disconnect rates in its corporate reporting, the IFAD results framework reports on both the share of projects with an overall positive rating, both based on PCRs and based on independent evaluation office (IEO) ratings. The World Bank actively monitors disconnects on an ongoing basis, displays them in portfolio dashboards and raises a risk flag for a country's entire portfolio when the disconnect rate exceeds a certain target (Table 2.2). This has provided managers and staff with incentives to avoid disconnects and pay more attention to the quality of completion reports. However, the strong focus on disconnects and the aim to minimize these can also influence the candor of self-evaluations. The IEG's evaluation of the self-evaluation system concluded that the validation process has significant influence over behaviors and incentives, and affects the content, candor and usefulness of the self-evaluation. Project teams are careful about what they put in PCRs lest it be used by IEG to downgrade project performance and generate a disconnect (World Bank 2016). The independent evaluation of the AfDB's self-evaluation system concluded that lack of attention to the disconnect rates is an important factor behind ineffective management oversight of PCR quality and lacking candor in PCRs (AfDB 2020). ADB management found that monitoring disconnects did not incentivize credible reporting in self-evaluations and has therefore started to monitor the quality of self-evaluations.

III. MONITORING PROJECT RESULTS AT THE CORPORATE LEVEL

3.1 **Project results monitoring in all five institutions is part of a larger system of results monitoring**. All five institutions monitor results at the corporate, portfolio and project levels. Corporate, sector, and country portfolio performance monitoring draw significantly on project performance information.

3.2 **The Asian Development Bank (ADB), African Development Bank (AfDB), International Fund for Agricultural Development (IFAD), and Inter-American Development Bank (IDB) have similar corporate results frameworks (CRFs) that are structured into three or four tiers.** They all track outputs and outcomes, as well as operational and organizational effectiveness. The top tier provides the development context and evolution of key development performance indicators in the areas of the institution's intervention. The second tier monitors the institution's contribution to key development results, aggregating results from individual projects. The third tier covers operational performance (e.g., quality at entry, portfolio and project performance, portfolio alignment with institutional priorities), while the fourth tier monitors organizational performance (human resources, efficiency of resources use, process efficiency). IFAD and the IDB have combined operational and organizational performance reporting into a single tier. Over time, all institutions have undertaken efforts to refocus and streamline their results frameworks. The number of indicators monitored varies across the four institutions, ranging from 105 at the AfDB to 65 at the IDB.

3.3 **The World Bank Group has recently revised its corporate results monitoring system and separated reporting on results from reporting on organizational effectiveness and efficiency.** The new World Bank Group scorecard adopts a thematic structure and is focused on results. It contains 8 vision indicators, 22 client context indicators and 22 associated World Bank Group results indicators. The scorecard is complemented by the Organizational Effectiveness and Efficiency Dashboard, which tracks operational performance based on 35 indicators similar to those in tier three/four of the other institutions' results frameworks.

3.4 **Corporate results monitoring is expected to fulfill multiple functions**. These include serving as a tool for accountability and communication of corporate performance on one hand and as a tool for learning and corporate decision-making. A recent benchmarking of multilateral development banks' (MDBs') corporate results monitoring concluded, however, that the information generated by these results measurement systems was mostly used for accountability and communication purposes, rather than for decision-making and learning. (AfDB 2021). Thus, corporate results monitoring appears to be afflicted by a similar challenge as project self-evaluations, which have been found to be more of an accountability and reporting instrument than a learning tool (World Bank 2016, AfDB 2020, IFAD 2023b, ADB 2020). Recent evaluations of the AfDB and IFAD CRFs (AfDB 2021, IFAD 2023a) furthermore point to issues of questionable data reliability at the project level, which then affect the quality of data reported in the CRFs. To help address this issue, IFAD is carrying out impact evaluations on a sample of projects each year, which are then used to extrapolate portfolio-wide contributions to tier two indicators.

3.5 **All five institutions report on the rate of completed projects that have achieved a positive overall performance rating.** The AfDB, ADB, IDB, and IFAD include in their corporate result frameworks information on the share of completed projects with a positive (top half of the rating scale) overall performance rating assigned by the independent evaluation office. The World Bank does so in its Organizational Effectiveness and Efficiency Dashboard, which also reports on the Independent Evaluation Group's (IEG's) rating of the monitoring and evaluation (M&E) quality of operations. The AfDB and ADB also include select validated core criteria ratings.[26] IFAD includes additional project evaluation criteria ratings, but the ratings are those reported in the project completion reports (PCRs) rather than those validated by the independent evaluation office.

3.6 **All institutions other than the World Bank include indicators on portfolio quality in their CRFs. At the World Bank, this performance is monitored in the Operational Effectiveness and Efficiency Dashboard.** At the ADB, the CRF includes the share of projects rated satisfactory or higher on the five key dimensions assessed in the project performance rating system, as well as the share of projects that have been at risk for four or more quarters (as a proxy for proactivity on underperforming projects). Similarly, at IFAD, the share of projects rated moderately satisfactory or higher on implementation progress and a proactivity index (which measures what share of problem projects in the previous period have undergone a formal process to address the problems) are reported in the CRF. At the AfDB, the share of projects flagged at risk and those facing implementation challenges figure in the CRF (in addition to disbursement and procurement indicators). The IDB's CRF includes the share of active projects with satisfactory performance. At the World Bank, the share of projects in problem status and a related proactivity index (which monitors actions taken to deal with flagged projects in the preceding 12 months) are monitored and reported on in the Organizational Effectiveness and Efficiency Dashboard. IFAD and AfDB also report on the share of projects with a satisfactory quality at entry score,[27] while the ADB reports on the share of design- and procurement-ready infrastructure projects in its 2019−24 CRF. The IDB's CRF does not report on the Development Effectiveness Matrix (DEM) scores.

3.7 **Including project and portfolio performance indicators in CRFs can help focus managerial attention on such performance.** Interviews with staff from comparator organizations carried out for this report suggest that when CRFs contain targets for project and portfolio performance and results delivery that are filtered down to the level of unit managers, such indicators are taken seriously. However, interviews also suggest that such targets can lead to overly optimistic reporting and ratings in an effort to meet performance targets.

---

26 AfDB core criteria ratings reported in its corporate results framework (CRF) include percent of projects rated satisfactory on effectiveness and sustainability. ADB core criteria ratings reported in its CRF include percent of projects rated successful on relevance, effectiveness, efficiency, and sustainability.

27 AfDB will discontinue the indicator which measures the share of projects with a satisfactory readiness score and replace it with more granular indicators that reflect the share of projects with low scores in each of the four quality at entry assessment dimensions.

IV. PROJECT DEVELOPMENT EFFECTIVENESS GOVERNANCE

4.1 **Although the organizational structure differs across the five institutions, their governance structure for development effectiveness management is similar**. In all five institutions, the development of policies and guidelines governing quality at entry, project monitoring, and, in most cases, self-evaluation is the responsibility of a central department with input and feedback from operational departments. At the Asian Development Bank (ADB), developing guidelines for project completion reports (PCRs) is the responsibility of the independent evaluation office. In other institutions, these guidelines are developed by management with inputs or feedback from the independent evaluation department.

4.2 **In all five reviewed institutions, developing project results frameworks is the responsibility of the operational team preparing the project, whereas the responsibility for assessing the results framework during project design varies.** Quality at entry assessments of the African Development Bank (AfDB) and the International Fund for Agricultural Development (IFAD), which include evaluability assessments, are coordinated by a central department that seeks inputs from various specialists such as financial management, safeguards, procurement, and sector specialists. At the Inter-American Development Bank (IDB), the Development Effectiveness Matrix (DEM) is prepared by project teams and reviewed by the central department (Office of Strategic Planning and Development Effectiveness). At the World Bank, project results frameworks and the associated theory of change (TOC) are reviewed as part of the overall project quality assurance process at regional level, with inputs from regional monitoring and evaluation (M&E) specialists. Similarly, at the ADB, the project design and monitoring framework (DMF) is reviewed as part of the regular project design quality assurance process.

4.3 **In all five institutions, preparing project supervision or monitoring reports is the responsibility of project teams**. They are approved by sector managers/division chiefs at all institutions except IFAD, where the regional director approves them. At the IDB and the World Bank, sector management/division chief approval is contingent on inputs and validation by country management units.

4.4 **Preparing PCRs is also the responsibility of project teams in all organizations.** It is, however, often contracted out to consultants under the supervision of the project team leader. Signing off on these self-evaluations is the responsibility of sector managers or directors. The independent evaluation offices validate PCRs (including their ratings) in all institutions. In most institutions, project teams are also responsible for providing project-specific inputs for second-tier corporate results monitoring indicators, where such indicators are relevant and included in the project results framework. At IFAD, such information is collected by a central department based on impact evaluations carried out on a sample of projects and then extrapolated to the portfolio.

4.5 **Reporting on aggregate project performance and results is similar across the five institutions**. Data from individual project monitoring reports (PMRs) feeds into operational dashboards accessible to operational managers in all institutions (Table 4.1). PMRs are made public, at least in part, by all institutions except the AfDB. Performance of the active portfolio is reported to the Board by all institutions. Project self-evaluations are shared with the Board for information and then made public by all institutions except IFAD, where they are not shared with the Board before being made public. A report summarizing the results of validations of project self-evaluations is produced by the independent evaluation office of all institutions and sent to the Board annually. The validation reports are discussed by either a committee of the Board or the full Board at all institutions except the AfDB. In the latter, the evaluation office's report is sent to the Board for information. All institutions produce annual updates of their corporate results frameworks (CRFs), generally in the context of a development effectiveness review, which is also sent to and discussed by the Board. The Board of the ADB discusses the Annual Portfolio Performance Review and the Development Effectiveness Review simultaneously and the independent evaluation office's Annual Evaluation Review, which summarizes among other topics the results of that year's validations within a few days of management's two reports, therefore allowing for some checks and balances on management's reporting.

Table 4.1. Development effectiveness reporting

	Project monitoring	Performance of completed projects	CRF
ADB	Project performance rating system, not public	PCR shared with the Board and made public. IEO Annual Evaluation Review reports on validated ratings discussed with the Board within a few days of Annual Portfolio Performance Review and Development Effectiveness Review.	Development Effectiveness Review, discussed with the Board simultaneously with Annual Portfolio Performance Review.
AfDB	Implementation progress report, public	PCR shared with the Board and made public. IDEV validation synthesis sent to the Board for information.	Annual Development Effectiveness Review, discussed with the Board.
IDB	Progress monitoring report, partly public	PCR shared with the Board for information and made public. OVE validation report, discussed with the Board.	DEO discussed with the Board.
IFAD	Supervision Report, public	PCR made public IEO Annual Evaluation Review discussed with the Board	Report on IFAD's Development Effectiveness discussed with the Board.
World Bank	Implementation and Supervision Report, partly public	PCR sent to the Board and made public. IEG's Report on Results and Performance of the World Bank Group discussed with the Board.	Corporate Scorecard Update and Organizational Effectiveness and Efficiency Dashboard discussed with the Board.

Source: OVE based on review of institution specific documents and interviews
Note: ADB = Asian Development Bank. AfDB = African Development Bank; CRF = corporate results framework; DEO = Development Effectiveness Overview; IDB = Inter-American Development Bank; IDEV = Independent Development Evaluation; IEG = Independent Evaluation Group; IEO = independent evaluation office; IFAD = International Fund for Agricultural Development; OVE = Office of Evaluation and Oversight; PCR = project completion report.

4.6 **The role of the Board with respect to development effectiveness is similar across all five institutions.** The Board approves individual projects, including their results frameworks, and it approves level one project restructurings. It receives PCRs and regular updates on the status of the portfolio. It also receives and, in most cases, discusses the report summarizing the results of validations by the independent evaluation office, except at the AfDB where the report is sent to the Board for information only. The Board also approves and subsequently receives and discusses updates on the corporate results monitoring framework.

V. CONCLUSION

5.1 **The complement of development effectiveness management instruments for sovereign guaranteed (SG) operations is similar across the five institutions reviewed, but the design and application of individual instruments varies.** All reviewed institutions undertake an assessment of project design, have a project monitoring system that includes identification of potential or actual problem projects, require a self-evaluation system with project achievement ratings being validated by the institution's independent evaluation office, and report aggregate results in the institution's corporate results framework (CRF). There is, however, variation across the institutions on the design and application of these development effectiveness management system building blocks.

5.2 **The African Development Bank (AfDB), Inter-American Development Bank (IDB), and International Fund for Agricultural Development (IFAD) use formal scoring systems to assess projects at entry, with the IDB's scoring being more focused on evaluability than those of the other two institutions**. The AfDB, IDB, and IFAD all require a formally scored assessment of projects at entry. The IDB's Development Effectiveness Matrix (DEM) scoring is more focused on evaluability than the quality at entry assessments by the AfDB and IFAD. The Asian Development Bank (ADB) requires a project design and monitoring framework (DMF) as part of project design but only scores certain implementation readiness aspects of infrastructure projects. Most institutions require that project design include a theory of change (TOC) of varying specificity linked to the results framework. Interviews with the Independent Evaluation Group (IEG) and operational staff at the World Bank suggest that this requirement has helped sharpen projects' results focus and evidence base.

5.3 **The monitoring systems for projects under implementation cover similar core dimensions across the five institutions but they vary in the number of areas they assess.** IDB's progress monitoring report system assesses and rates fewer dimensions than that of the comparator institutions and puts less emphasis on the rating of the likelihood of achieving project development objectives than most other institutions.

5.4 **All five institutions draw on the ratings in the PMRs to identify potential and actual problem projects, but IDB and ADB do not consider progress toward development objectives for this purpose**. IDB's identification of problem projects is limited to the Synthetic Indicator produced in the PMRs and does not consider progress on development objectives nor other dimensions considered by the comparator institutions, such as project management, procurement, and environmental and social safeguards. The AfDB, ADB, and IFAD draw on a broader range of PMR ratings, though only the AfDB and IFAD include progress on development objectives or results. The World Bank's system considers an extensive list of PMR ratings (according substantial weight to implementation and progress on development objectives ratings) and additional factors, including, among others, the share of disconnects between ratings of the Independent Evaluation Group (IEG) and self-ratings in the country portfolio.

5.5 **The World Bank's Board has delegated the most wide-ranging authority for project restructuring to management among the five institutions.** All institutions except the IDB have revised the project restructuring policies in the last several years. The objective of these revisions was generally twofold: (i) to introduce more flexibility and streamline restructuring procedures to encourage managers and staff to proactively restructure operations at risk or off track, and (ii) to clarify criteria for restructuring and associated approval authority. All five institutions have a two-tiered restructuring policy, with changes in one tier (level 1) requiring Board approval and changes in the other tier (level 2) being approved by various levels of managers. The World Bank's Board has delegated the most wide-ranging authority to management for approval of project restructuring, including approval of changes to project development objectives. In the other four institutions, changes to development objectives require Board approval, with IFAD and the IDB stipulating that substantial changes require Board approval.

5.6 **To incentivize early action on projects experiencing problems and ensure that projects are not simply restructured to reduce results targets in line with achievements toward the end of the project, the World Bank uses a split rating system when evaluating the performance of projects at completion**. The system evaluates projects against the original results framework as well as against the revised framework and derives a weighted average rating with the weights equivalent to the resources spent before and after restructuring. The ADB, AfDB, and IFAD evaluate against revised frameworks if a project has gone through a formal level 1 or level 2 restructuring, while the IDB evaluates against the results framework in effect within 60 days of project eligibility unless a project was formally restructured with Board approval.

5.7 **All five institutions require the preparation of a self-evaluation at project completion and draw on Organisation for Economic Co-operation and Development–Development Assistance Committee (OECD-DAC) criteria**. All institutions assess completed projects on relevance and effectiveness. All institutions also evaluate efficiency for investment operations, but the World Bank and IDB do not apply the efficiency criteria to policy-based operations (PBOs), given that loan funds go to the general budget. All institutions except the World Bank also include sustainability among its core criteria. The World Bank assesses but does not rate risks to development outcomes at completion (akin to sustainability). IFAD complements the OECD-DAC core criteria with additional criteria aligned with its mandate, such as innovation, gender equality, scaling up, environment and natural resource management, and adaptation to climate change. When assessing effectiveness, the IDB, IFAD, ADB, and AfDB have clearly defined achievement bands for each point on the rating scale, while the World Bank allows for more evaluator judgment on investment loans but not on PBOs.

5.8 **All five institutions derive an overall project performance rating from the core criteria ratings, but the way in which the overall rating is derived varies.** The AfDB and World Bank use a decision tree to derive the overall project performance rating from core criteria ratings. The ADB and IFAD calculate the simple average of the core criteria ratings to determine the overall performance rating. The IDB calculates a weighted average of the core criteria ratings. The AfDB and IDB give higher weight to the effectiveness rating than to other core criteria when deriving the overall outcome rating. The World Bank does so for investment projects.

5.9 **The evaluation offices of all institutions validate project self-evaluations**. All evaluation offices except that of the AfDB validate 100 percent of self-evaluations, while the latter validates at least 65 percent. All evaluation offices report on the results of validations to their Boards. The Boards of all institutions except the AfDB discuss these reports. The World Bank and IFAD emphasize disconnects between self-ratings and validated ratings to provide incentives for evidence-based reporting and assessment.

5.10 **The corporate results reporting systems of all institutions incorporate some elements of the project development effectiveness assessments**. All institutions report on the share of completed projects that received positive overall outcome ratings by the independent evaluation office. The AfDB, ADB, IFAD, and World Bank also report on selected other dimensions of completed projects, while the IDB does not. All institutions other than the World Bank include indicators on portfolio quality in their corporate results reporting, while the World Bank includes them its Operational Effectiveness and Efficiency Dashboard. Interviews suggest that when targets on project and portfolio performance from corporate results reporting are filtered down to unit managers, managers pay attention to them. However, such a system needs to be complemented with an incentive structure that encourages candid and realistic reporting on project performance and regular updating on the status of results indicators.

References

African Development Bank (2020), Independent Development Evaluation, Evaluation of AfDB's Self-Evaluation System and Processes, Abidjan.

African Development Bank (2021), Independent Development Evaluation, Mid-Term Evaluation of AfDB's Results Measurement Framework 2016-2025, Abidjan.

Asian Development Bank (2020), Independent Evaluation Department, 2020 Annual Evaluation Review: ADB's Project Level Self-Evaluation System. Manila.

Inter-American Development Bank (2008), Development Effectiveness Framework (DEF), document GN-2489. Washington DC.

Inter-American Development Bank (2024), Office of Evaluation and Oversight, Evaluation of the Development Effectiveness Framework, document RE-583-2. Washington DC.

Inter-American Development Bank (2025a), Office of Evaluation and Oversight, Proposed Work Program and Budget of the Office of Evaluation and Oversight (OVE), 2025–2026, document RE-600-3. Washington DC.

Inter-American Development Bank (2025b), Office of Evaluation and Oversight, IDB Group Project Performance: The 2024 Validation Cycle, document RE-606. Washington DC.

International Fund for Agricultural Development (2023a), Independent Evaluation Office, Review of the IFAD12 Results Management Framework, Rome.

International Fund for Agricultural Development (2023b), Independent Evaluation Office, Annual Report on the Independent Evaluation of IFAD 2023, Rome.

World Bank (2016), Independent Evaluation Group, Behind the Mirror: A Report on the Self-Evaluation Systems of the World Bank Group, Washington DC.

World Bank (2023), Independent Evaluation Group, Results and Performance of the World Bank Group 2023, Washington DC, 2023.

Office of Evaluation and Oversight - OVE

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

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

iadb.org/evaluation

linkedin.com/showcase/idb-ove

@BID_evaluacion