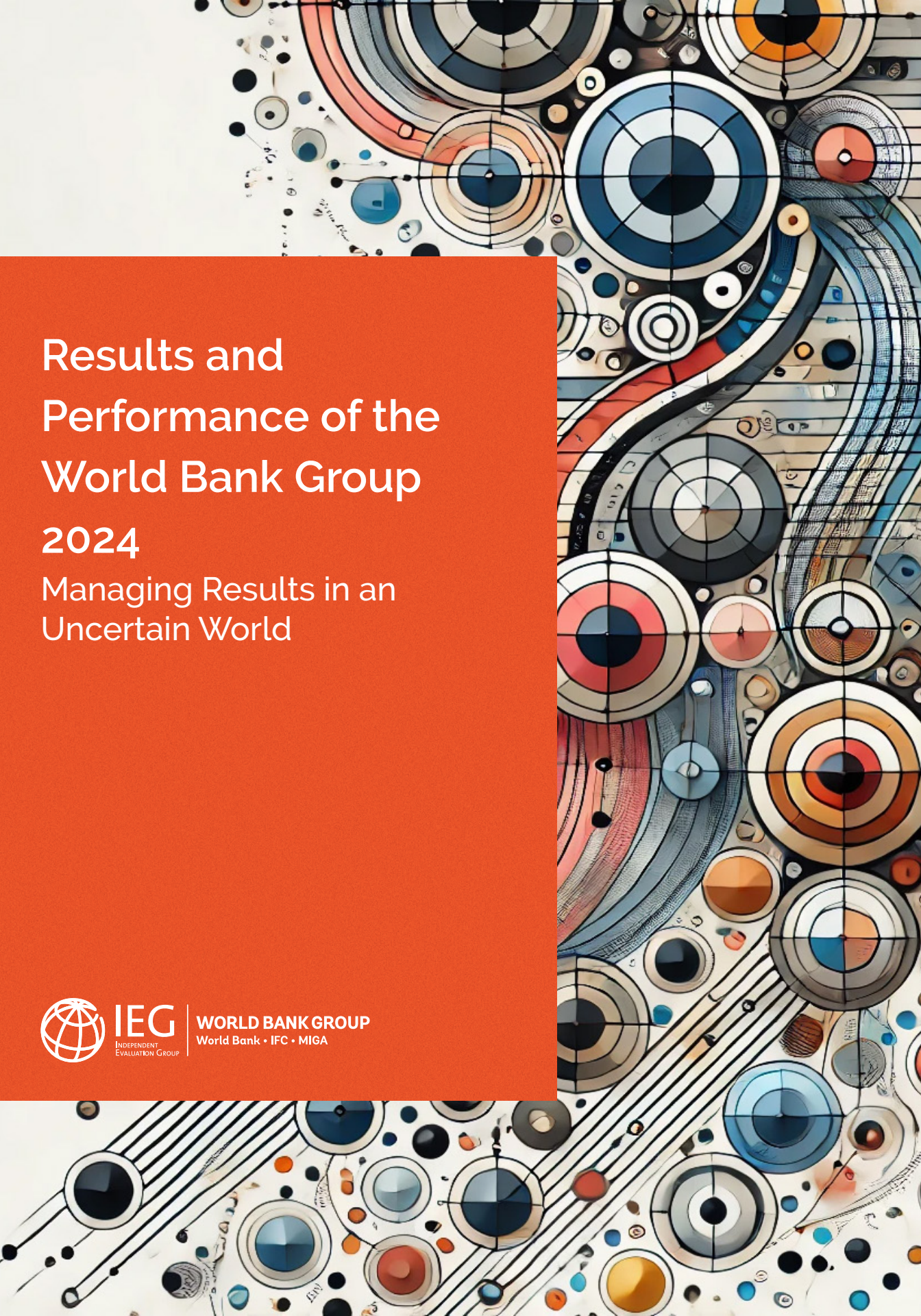


Results and Performance of the World Bank Group 2024

Managing Results in an Uncertain World

IEG
Independent Evaluation Group

WORLD BANK GROUP
World Bank • IFC • MIGA

1818 H Street NW
Washington, DC 20433
Telephone: 202-473-1000
Internet: www.worldbank.org

ATTRIBUTION
Please cite the report as: World Bank. 2025. *Results and Performance of the World Bank Group 2024: Managing Results in an Uncertain World*. Independent Evaluation Group. World Bank.

COVER PHOTO
Created using GPT-3 using the prompt, "The concept is that things work differently for different groups. The image should be abstract." November 15, 2024.

EDITING AND PRODUCTION
Amanda O'Brien

GRAPHIC DESIGN
Luísa Ulhoa

Results and Performance of the World Bank Group 2024

Managing Results in an Uncertain World

March 7, 2025

Contents

Abbreviations vi

Acknowledgments vii

Overview viii

Management Comments xxv

1. Introduction 1

Background and Audience 1

Objectives, Question, Scope, and Use 3

Methods 5

Organization of the Report 6

2. World Bank 8

Trends 11

Challenges 20

Levers 27

3. International Finance Corporation 36

Investment Projects 38

Advisory Services 57

4. Multilateral Investment Guarantee Agency 68

Trends 70

Challenges 73

Levers 76

5. Country Program 79

Trends 82

Challenges 87

Levers 95

6. Conclusions 99

Fragile and Conflict-Affected Situations 99

Operational Efficiency and Effectiveness: Preparation and Processing Time 100

World Bank Group Scorecard 101
Joint Country Representation 102

Bibliography 105

Boxes

Box O.1. Validation Frameworks and Methodologies of the Independent Evaluation Group viii
Box 1.1. Relevance of *Results and Performance of the World Bank Group 2024* to the Better Bank Initiatives 2
Box 1.2. Key Terms and Concepts in *Results and Performance of the World Bank Group 2024* 5
Box 2.1. Main Performance Ratings in the World Bank 10
Box 2.2. Delivery Challenges in Operations Taxonomy 20
Box 5.1. Ratings and Perceptions at the Country Level Analyzed by *Results and Performance of the World Bank Group 2024* 81
Box 5.2. Important Conditions for World Bank Group Collaboration 93

Figures

Figure 2.1. World Bank Project Rating Trend and Coverage 12
Figure 2.2. Portfolio Composition Shift from FY20 to FY23 14
Figure 2.3. Gap in Average Project Volume Between Operations in Fragile and Conflict-Affected Situations and in Non–Fragile and Conflict-Affected Situations 15
Figure 2.4. Share of All Lending Projects in Operation 15
Figure 2.5. Average COVID-19 Exposure Among Projects Closed During FY20–23 16
Figure 2.6. Distribution of Outcome Ratings by COVID-19 Exposure 18
Figure 2.7. World Bank Performance Ratings 19
Figure 2.8. Challenges in Institutional Capacity of Stakeholders and Project Factors 22
Figure 2.9. Density Distribution of Project Preparation Days Between Initiation and Approval 25

Figure 2.10. Correlations Between Project End Systematic Operations Risk-Rating Tool Rating and Challenges 29
Figure 2.11. Comparison of Average Systematic Operations Risk-Rating Tool Ratings 30
Figure 2.12. Share of Projects with an Objective on Expanding Access to Services 33
Figure 3.1. Performance Ratings in International Finance Corporation Investment Projects 38
Figure 3.2. Trends in Development Outcomes for International Finance Corporation Investment Projects, Calendar Years 2013–23 40
Figure 3.3. Development Outcomes by Region, Calendar Years 2013–15 Versus 2021–23 41
Figure 3.4. Share of Projects Evaluated and Validated by the Independent Evaluation Group by Region, Calendar Years 2013–15 Versus 2021–23 42
Figure 3.5. Most Prevalent Factors Linked to Development Outcomes, Throughout the International Finance Corporation and by Subgroup 45
Figure 3.6. Top 10 Work Quality Factors, Net Effect 51
Figure 3.7. Association Between Front-End Work Quality and Development Outcome, Calendar Years 2021–23 52
Figure 3.8. Average Elapsed Time from Mandate to First Disbursement Versus Development Outcome 53
Figure 3.9. Association Between Additionality and Development Outcome, Calendar Years 2021–23 54
Figure 3.10. Performance Ratings in International Finance Corporation Advisory Services Projects 57
Figure 3.11. Trends in Development Effectiveness for International Finance Corporation Advisory Services Projects, FY13–23 59
Figure 3.12. Performance on Development Effectiveness by International Development Association Status, FY13–15, FY16–18, and FY21–23 60
Figure 3.13. International Finance Corporation Advisory Services Development Effectiveness, Work Quality, and Project Preparation and Design 62
Figure 3.14. Associations Between Work Quality, Project Preparation and Design, and Development Effectiveness, FY21–23 64
Figure 4.1. Performance Ratings in Multilateral Investment Guarantee Agency Guarantee Projects 69
Figure 4.2. Development Outcomes for Multilateral Investment Guarantee Agency Guarantee Projects, FY13–23 71

Figure 4.3. Development Outcomes for Multilateral Investment Guarantee Agency Guarantee Projects by Region, FY13–18 Versus FY18–23 72
Figure 4.4. Most Prevalent Factors Linked to Development Outcomes 74
Figure 4.5. Project-Level Versus Foreign Investment–Level Outcome Achievement Rates 76
Figure 5.1. Country Program Ratings, FY13–23 83
Figure 5.2. Country Program Ratings by Region, FY13–23 84
Figure 5.3. Country Program Ratings in the Country's Latest Completion and Learning Review Validation 86
Figure 5.4. World Bank Group Performance at the Country Program Level 88
Figure 5.5. World Bank Group Collaboration at the Country Program Level 93

Tables

Table 2.1. Mann–Whitney *U* Test and Ordinal Logistic Regression on Project Ratings and Long Preparation Time (more than 1,393 days) 26
Table 5.1. Client Country Perceptions and World Bank Group Performance 97

Appendixes

Appendix A. Methodology 114
Appendix B. World Bank Dashboards 167
Appendix C. Decomposition Analysis of International Finance Corporation Project Rating Trends 169
Appendix D. International Finance Corporation Investment Outcome Types 178
Appendix E. Association Among Development Outcomes, Work Quality, and Additionality in International Finance Corporation Investment Projects 184
Appendix F. Project Examples of Factors Linked to International Finance Corporation Investment Project Performance 192
Appendix G. Decomposition Analysis of Multilateral Investment Guarantee Agency Project Rating Trends 197
Appendix H. Multilateral Investment Guarantee Agency Guarantee Outcome Types 202
Appendix I. Project Examples of Factors Linked to Multilateral Investment Guarantee Agency Guarantee Project Performance 205
Appendix J. One World Bank Group Collaboration 210

Abbreviations

AIMM	Anticipated Impact Measurement and Monitoring
ASA	advisory services and analytics
CLRV	Completion and Learning Review Validation
COS	Country Opinion Survey
CPF	Country Partnership Framework
CY	calendar year
DOTS	Development Outcome Tracking System
FCS	fragile and conflict-affected situations
ICR	Implementation Completion and Results Report
IDA	International Development Association
IEG	Independent Evaluation Group
IFC	International Finance Corporation
IPF	investment project financing
M&E	monitoring and evaluation
MIGA	Multilateral Investment Guarantee Agency
RAP	*Results and Performance of the World Bank Group*
SORT	Systematic Operations Risk-Rating Tool

All dollar amounts are US dollars unless otherwise indicated.

Acknowledgments

This report was prepared by an Independent Evaluation Group team co-led by Stephen Porter (senior evaluation officer) and Melvin Vaz (senior evaluation officer), under the guidance of Theo Thomas (director, Human Development and Economic Management), Timothy Johnston (manager, Corporate and Human Development), Beata Lenard (manager, Financial and Private Sector Micro Unit), and Galina Sotirova (former manager, Corporate and Human Development), and the overall guidance of Sabine Bernabè (Director General, Evaluation).

The evaluation team included Rocio Manchado Garabito, Diana Goldemberg, William Haworth, Aarre Laakso, Gaby Loibl, Melissa Metz, Chikako Miwa, Daniel Palazov, Konstantin Panov, Xiaoxiao Peng, Ichiro Toda, and Stefan Triendl. Other Independent Evaluation Group colleagues who made valuable contributions are Leonardo Bravo, Corky De Asis, Susan Caceres, Soniya Carvalho, Unurjargal Demberel, Jenny Gold, Ananda Ghose, Izlem Hajdenberg, Ramachandra Jammi, Donna Jeffrey, Priyanka Jetwani, Giuseppe Larossi, Eduardo Maldonado, Christopher David Nelson, Avjeet Singh, Min Song, Carlos Otobed Stagliano, Mercedes Vellez, Viktoriya Yevsyeyeva, and Virginia Ziulu.

The external advisory panel was composed of senior international development experts and evaluation professionals: Gabriela Yarahuan Perez (Inter-American Development Bank and a professor at Centro de Investigación y Docencia Económicas, Mexico), Ede Jorge Ijjasz-Vasques (nonresident senior fellow at the Brookings Institution), and Sabine Schlorke (former International Finance Corporation manager, Global Manufacturing, and manager, Global Trade Supply Chain Financing).

The team is grateful to Lily Chu and Estelle Raimondo for their comments on an early version of this report and to other Independent Evaluation Group colleagues who shared their insights and feedback throughout the processes. The team also thanks Operations Policy and Country Services, International Finance Corporation, and Multilateral Investment Guarantee Agency colleagues who engaged with us during the preparation of the report.

Overview

The *Results and Performance of the World Bank Group (RAP) 2024* report by the Independent Evaluation Group (IEG) is the 14th annual report in the series. The *RAP* series aggregates and interprets evidence on World Bank Group performance, mainly using IEG's validations of World Bank, International Finance Corporation (IFC), and Multilateral Investment Guarantee Agency (MIGA) self-evaluations of projects and country programs. These validations summarize the extent to which the Bank Group institutions have achieved key aspects of project and country program design, implementation, and results.

The *RAP* series follows the principles of continuity, symmetry, and innovation. *RAP 2024* maintains continuity with earlier reports by reviewing performance over a 10-year period based on ratings from IEG validations using standardized methodologies (box O.1). It maintains the principle of symmetry across institutions by analyzing common or similar factors linked to performance while also acknowledging their differences. Hence, *RAP* includes separate chapters for World Bank, IFC, MIGA, and country programs, which allows more in-depth discussion of performance trends and relevant topics for each institution. *RAP 2024* includes innovative analysis of (i) processing time for IFC investment projects, (ii) risk identification and mitigation, (iii) IFC work quality and additionality, (iv) the Country Opinion Surveys, and (v) results measurement in IFC investment projects.

RAP 2024 can inform the learning and accountability processes of the Bank Group's Board of Executive Directors. The evidence presented in this report focuses on significant changes and substantive patterns in performance relevant to the institutions of the Bank Group, but it does not identify causal relationships. Drawing on this evidence, the Board can use *RAP 2024* for learning because it helps the Board further examine the challenges facing Bank Group institutions and highlights the levers that can be pulled to enhance their performance. Furthermore, *RAP 2024* aids accountability because it helps the Board understand trends and changes in portfolio performance.

RAP 2024 is produced at a time when the work of the Bank Group is increasingly exposed to risky contexts and periodic shocks. The Bank Group has committed to giving increased priority and resources to lower-income countries and countries classified as fragile and conflict-affected situations (FCS), where most of the world's poor people reside (World Bank 2024c, 2024e). The COVID-19 pandemic and subsequent shocks to global food and fuel prices have highlighted the importance of building resilience and responding proactively to domestic and external shocks. Climate change is likely to make such shocks more frequent and more severe. Previous IEG evidence has described areas where the Bank Group has successfully responded to increased risks and maintained performance, for example, through rapidly adapting its operational and country portfolios, adjusting lending volumes, and using new digital technologies during the COVID-19 pandemic (World Bank 2022h, 2024a, 2024b). Building on these findings, this report highlights challenges confronted in risky contexts at a portfolio level and suggests responses that can help enhance performance.

Key Findings

RAP 2024 highlights several cross-cutting findings. Although the World Bank, IFC, and MIGA have different business models and different evaluation methodologies for operations and projects (box O.1), some common themes and findings emerge that are outlined in this section. While the ratings themselves cannot be compared across institutions because different methodologies are used, rating trends can be compared.

Box O.1. Validation Frameworks and Methodologies of the Independent Evaluation Group

World Bank Group institutions employ different frameworks and methodologies when rating operations. World Bank operations use an objective-based methodology to derive project performance ratings. As such, the outcome rating is based on the extent to which the objectives stated at design (or formally modified) were achieved. The World Bank's self-evaluation and the Independent Evaluation Group's validation

(continued)

Box O.1. Validation Frameworks and Methodologies of the Independent Evaluation Group (cont.)

ratings are aggregated across operations (for example, outcome ratings are based on a six-point scale, ranging from highly satisfactory to highly unsatisfactory). Similarly, the International Finance Corporation (IFC) advisory services project performance ratings are derived from an objective-based methodology, which establishes minimum thresholds for rating and assessing projects' effectiveness. Country programs also follow an objective-based methodology. By contrast, evaluation systems and performance ratings for IFC's investment projects and the Multilateral Investment Guarantee Agency's guarantee projects both are objective based and include market or industry benchmark-based methodologies, particularly for measuring financial performance to ensure the sustainability of IFC investments and Multilateral Investment Guarantee Agency guarantees. Bank performance ratings in the World Bank and work quality in IFC are rated separately from outcomes; although the terminology is different, they both measure how well implementation and design issues were addressed, and these are largely within the control of the institutions.

All rating methodologies align with good practice standards for evaluating public and private sector projects, as established by the Evaluation Cooperation Group of multilateral development banks, except for technical assistance or advisory services, which do not yet have established good practice standards (ECG 2012). Further explanation of rating methodologies is provided for each Bank Group institution at the beginning of the relevant chapters.

While project and country program ratings are important proxies for the Bank Group's development effectiveness, they are not direct measures of contributions to key outcomes (for example, poverty reduction, increased learning, or reduced gender inequality). As highlighted in recent thematic evaluations, projects with less ambitious objectives (for example, increasing access to services or other outputs) may be rated successful even in the absence of evidence on improvement in higher-level outcomes. Given the fundamental differences in methodologies across the institutions, ratings cannot be compared among the World Bank, IFC, and the Multilateral Investment Guarantee Agency. However, the trends in ratings can be compared.

Sources: Independent Evaluation Group; World Bank 2024a, 2024b, 2024f.

The Bank Group's performance ratings have plateaued or declined, which is linked to exposure in riskier country contexts. For the World Bank, ratings for operations have recently plateaued because of increased portfolio exposure in FCS. For IFC and MIGA, recent declines have been driven primarily by declining outcome ratings in International Development Association (IDA) and blend countries (without significant portfolio shifts toward IDA and blend or FCS). Bank Group country program development outcome ratings are 19 percentage points lower for both IDA-eligible countries, compared with those eligible under the International Bank for Reconstruction and Development, and 21 percentage points lower for FCS countries than for non-FCS countries. In contrast, Bank Group performance in country program ratings has remained at the same levels in FCS, IDA, and blend countries.

COVID-19 has not yet affected the overall ratings of the operations of the World Bank, although the performance of IFC advisory services and IFC investment projects with financial institutions has been negatively affected. The Bank Group's response to COVID-19 appears to have mitigated negative impacts on the ratings of World Bank projects. For IFC investment projects with financial institutions, some clients were able to adapt to COVID-19, but the pandemic detracted from the success of other clients. For IFC advisory services, performance suffered despite proactive support and restructurings.

Bank Group performance can be improved even with risks increasing. First, the quality of the design of operations and projects is strongly associated with better results. Quality of design can include realism and relevance of objectives with respect to country context and appropriate sequencing of interventions. Second, across the Bank Group portfolio, effective risk identification and proactive risk management—particularly for risks under direct or indirect Bank Group control—are linked with stronger performance. This finding is particularly important in FCS, where the analysis of World Bank Systematic Operations Risk-Rating Tool ratings shows that larger risk reductions during implementation were associated with better outcomes, whereas failure to mitigate risks contributed to lower outcomes. Third, client capacity challenges are linked to lower performance across institutions, particularly in FCS or in IDA-eligible countries. Developing public and private clients' organizational capacity to adapt to changing contexts and implement fit-for-purpose solutions (for example, in procurement), implement coordination

mechanisms, and establish performance reporting systems can contribute to better outcomes.

The Bank Group's implementation of results monitoring could enhance outcome ratings and contributions to development effectiveness. Gaps in the design and implementation of results monitoring are persistent concerns in project and Country Program Evaluations and in IEG thematic evaluations. Monitoring and evaluation practices are subject to World Bank influence and are correlated with outcome ratings. Despite improvements over the past decade, more than one-third of World Bank operations rated by IEG have inadequate monitoring and evaluation practices. MIGA is behind on submitting self-evaluations of guarantee projects, which hinders both accountability and learning. IFC could also improve the measurement of outcomes, particularly market outcomes (that is, changes in the market beyond those narrowly linked to the project). IEG validations identified that 83 percent of country programs over the past 10 years had major inadequacies in their results frameworks.

Trends in Performance

Performance ratings across the Bank Group's institutions have not consistently increased over the long term.[1] For the World Bank, project ratings have plateaued after a steady increase over the past decade. IFC project ratings recovered in the past five years but remain lower than they were a decade ago. MIGA project ratings declined slightly over the past decade.

- World Bank project outcome ratings increased from an average rating of 3.8 (out of a maximum rating of 6) in 2013 to 4.3 in 2020, then plateaued at this level between FY 2020 and FY23. World Bank performance has trended gradually upward since FY13, reaching 4.3 in FY23.
- The development outcome ratings of IFC investment projects improved over the medium term, from 41 percent mostly successful or better (calendar year [CY]16–18) to 51 percent mostly successful or better (CY21–23), but they remain lower than they were a decade ago (53 percent in CY13–15). IFC work quality ratings have declined over the past decade, from 62 percent satisfactory or better (CY13–15) to 55 percent (CY21–23).

- The development effectiveness ratings of IFC advisory services improved over the medium term, from 41 percent mostly successful or better (FY16–18) to 50 percent mostly successful or better (FY21–23), but they remained below 61 percent in FY13–15. Work quality ratings improved from 48 percent satisfactory or better (FY16–18) to 59 percent (FY21–23).
- The development outcome ratings of MIGA guarantee projects declined slightly, from 69 percent satisfactory or better in FY13–18 to 68 percent satisfactory or better in FY18–23. In the 15 MIGA guarantee projects that IEG validated during FY21–23, MIGA fully or partially achieved 81 percent of the intended project-level outcomes, but only 66 percent of the intended foreign investment–level outcomes. (Project-level outcomes refer, for example, to benefits to stakeholders, society, the environment, and the economy. Foreign investment–level outcomes refer, for example, to creating positive demonstration effects that signal opportunities to other market participants or potential capital providers.)
- The development outcomes improved in Bank Group country programs from 68 percent moderately satisfactory or above in FY13 to 78 percent moderately satisfactory or above in FY20. Bank Group performance in country programs was stagnant through FY20, with low ratings observed in approximately 40 percent of country programs.

World Bank operations and Bank Group country programs in FCS are exerting downward pressure on overall outcome ratings. The plateau in World Bank outcome ratings is mostly due to an increase in the share of FCS operations in the portfolio, which have lower ratings on average for most of the FY13–23 period. The proportion of projects with full or partial exposure to FCS increased from 31 percent in FY20 to 37 percent in FY23, magnifying the effect of their lower average ratings on the overall World Bank average outcome rating. The share of country programs in FCS with development outcomes rated moderately satisfactory or above is 55 percent, compared with 76 percent in non-FCS countries between FY13 and FY23.

For IFC and MIGA, declining project ratings in IDA and blend countries were the major contributor to overall declines in development outcome ratings. For IFC investment projects, the development outcome ratings of IDA and blend projects declined from 54 percent mostly successful or better

(CY13–15) to 46 percent (CY21–23), and the outcome ratings of FCS in IDA and blend countries declined sharply, from 50 percent mostly successful or better to 18 percent over the same period. For IFC advisory services, ratings of projects in IDA and blend countries declined from 59 percent mostly successful or better in FY13–15 to 43 percent in FY21–23, even though there were fewer projects in these countries over the latter period. The share of FCS projects in the overall active IFC investment portfolio has remained stable (at 11 percent); however, the number of projects in FCS has been increasing since FY21. Development outcome ratings of MIGA projects in IDA and blend countries declined from 74 percent satisfactory or better to 50 percent between FY13–18 and FY18–23, and ratings for MIGA projects in Sub-Saharan Africa declined from 72 percent satisfactory or better to 50 percent over the same period.

Challenges Influencing Performance

A wide variety of challenges affected the performance of Bank Group operations and projects. *RAP* undertook an analysis of the key challenges that affected project outcomes, including those linked to country context (over which the Bank Group has limited or no influence); institutional capacity of clients and stakeholders (over which the Bank Group has indirect influence); and factors related to project design, finance, and monitoring and evaluation (over which the Bank Group has direct influence).

Country context challenges were common throughout the portfolio but were more acute in FCS:

- The World Bank identified institutional capacity and financial management challenges in 75 and 76 percent of operations, respectively. In FCS, political interference and electoral cycles were intense challenges, identified in 74 and 80 percent of operations, respectively.
- The most prevalent challenges for IFC investment projects overall—and in several subgroups (IDA and blend, IDA and blend in FCS, and Latin America and the Caribbean)—were heightened business and economic risks. Risks related to business models, cyclicality, or the operating environment occurred in 25 percent of projects reviewed, and risks related to economic issues occurred in 24 percent of projects. In IDA and blend projects in FCS, civil unrest

(which occurred in 21 percent of these projects) and asset quality (which occurred in 25 percent of projects and included factors such as nonperforming loans in a financial institution's portfolio) were prevalent.

» The most prevalent challenges for MIGA projects were cost overruns and construction delays (which occurred in 46 percent of projects reviewed, mostly in real sector projects), quality of the company that owns and implements the project (which occurred in 31 percent of projects and included the company's ability, technical expertise, and track record), and legal or regulatory risk (which occurred in 27 percent of projects).

COVID-19 has affected the implementation of World Bank operations but not its performance. For operations that have closed since FY20, an increasing proportion of their life span occurred since the beginning of the COVID-19 pandemic, yet the outcome ratings of these operations are similar. Recent IEG evaluations and *RAP 2023* suggest that the World Bank responded to the pandemic by increasing restructuring and applying lessons from past crises. These actions led to 60 percent of World Bank country programs substantially realigning their portfolios to address the evolving needs arising from the pandemic (World Bank 2022h).

Although some IFC clients were able to adapt to the challenges created by the COVID-19 pandemic, the consequences of the pandemic detracted from the success of other clients. COVID-19 was identified as a factor in 19 percent of IFC investment projects in CY20–23. However, it was among the top factors only in the Financial Institutions Group. The consequences included drops in business volumes, banks forced to curtail loans to small and medium enterprises, and clients being downsized. Financial institutions with strong risk management, sound credit underwriting, and flexible business strategies were better able to withstand crises such as COVID-19. Conversely, COVID-19 was the most important factor in explaining why the development effectiveness of evaluated IFC advisory services projects declined during FY21–23, despite proactive actions during project design and supervision. The time frame for adaptation may also have been a factor: IFC advisory services projects are typically implemented in three years, whereas IFC investments are typically evaluated between five and seven years after financing is provided.

Operational and Country Program Design

The World Bank has opportunities to reinforce performance by improving aspects of operational design, especially in FCS. More successful World Bank operations, for example, grounded their design in practical operational lessons and insights, aligned objectives with on-the-ground realities, and sequenced tasks responsively. Since FY20, a growing share of World Bank operations in FCS improved alignment with their contexts by having more objectives focus on expanding access to services in addition to continuing efforts to preserve institutional strengths. Moreover, IEG evaluations have indicated the importance of World Bank efforts in, for example, the Geo-Enabling Initiative for Monitoring and Supervision in reinforcing elements of performance that are relevant for the design of operations (World Bank 2021d). The operations designed or restructured to include a focus on access to services in FCS received significantly higher average outcome ratings, although increased access does not necessarily contribute to improved outcomes—for example, improved learning or gender equality (World Bank 2024a, 2024b). Nonetheless, in some FCS contexts, sustaining service access may be both appropriate and more feasible.

IFC work quality (which includes front-end work, particularly project preparation and design, and project supervision and administration) is strongly associated with the performance of IFC investment and advisory services projects. A review of 19 IFC investment projects whose work quality was rated unsatisfactory or whose development outcomes were rated highly unsuccessful during CY21–23 found three front-end work quality factors that contributed to weak development outcomes: (i) market assessment (15 projects); (ii) client quality (10 projects); and (iii) assumptions, financial models, and project costs (9 projects). In addition, in FY21–23, the work quality of IFC advisory services projects was 59 percent satisfactory or better, whereas only 9 percent of projects with low work quality ratings achieved positive development effectiveness ratings.

While project preparation times must be sufficient to ensure strong design and implementation readiness, World Bank operations with very long preparation times (above the 90th percentile) have lower outcomes. The World Bank has recently made efforts to shorten project preparation times (World

Bank 2024c), which currently average two years from initiation to approval, including 10 percent of projects being prepared in less than seven months (including emergency operations and much of the COVID-19 response), and 10 percent of projects requiring nearly four years or more. Operations with the longest preparation times (more than 1,393 days) are more likely to encounter significant challenges with elements of their design and the development of institutional capacity. These operations also have significantly lower outcome, Bank performance, and monitoring and evaluation ratings on average. Shorter preparation time was not statistically associated with lower outcomes, however, which likely reflects the complex relationships between preparation time, design quality, and implementation readiness.

For IFC investment projects, spending sufficient time on front-end work may be particularly important in challenging contexts and for complex projects. Average IFC investment processing time is just over 12 months, with little difference between mostly successful or better and mostly unsuccessful or worse projects. However, in challenging contexts (specifically, in Africa, Middle East and North Africa, FCS, and IDA and blend), processing mostly successful or better IFC investment projects takes a few months longer than processing mostly unsuccessful or worse projects and projects that are especially complex (such as those in the Infrastructure industry group). Reducing key front-end work quality and preparation factors (such as market assessment; client quality; or assumptions, financial models, and project costs) may contribute to weaker development outcomes. For example, assumptions not based on feasibility studies or market assessments could contribute negatively to development outcomes. Finally, IFC advisory services cannot accurately measure preimplementation scoping time because not all advisory services projects go through the Concept Note stage (for example, some subprojects of approved programmatic umbrellas or "fast-track" projects that were follow-ons from previous engagements). In addition, out of the 411 standard advisory services projects evaluated and validated by IEG (FY13–23), 67 projects (16 percent) did not record a Concept Note date.

Collaboration among Bank Group institutions has improved over the past decade but could be further strengthened. The Bank Group has sought to improve synergies on country programs between institutions for almost three decades. Bank Group collaboration has increased: before FY16,

less than half of country programs demonstrated collaboration in at least one program, compared with 85 percent since then. However, only half of Country Partnership Frameworks since FY16 have demonstrated collaboration in more than one sector (with three-quarters of all collaboration in energy, agriculture, infrastructure, financial, or investment). The institutions have yet to consistently leverage each other's capacities across Country Partnership Framework periods and in multiple sectors. IEG evidence suggests that successful Bank Group collaboration on country programs rests on its institutions working toward a shared view of the priorities within a given sector. This includes a shared understanding of actors, opportunities and constraints to grow the sector, what needs to be done to address constraints, and the alignment of objectives across both public and private sectors.

Risk Identification, Mitigation, and Adaptation

Risk identification and mitigation measures enhance responses to challenges and are associated with improved outcome ratings. The Systematic Operations Risk-Rating Tool helps address key challenges the World Bank faces. Identifying and then mitigating risk through the Systematic Operations Risk-Rating Tool is useful because project risk ratings assigned by management and outcome ratings are correlated. For example, increases in risk ratings during implementation correlate with decreases in outcome ratings. Larger risk reductions during implementation were associated with better outcomes, whereas failure to respond to risks contributed to lower outcomes. Operations that identified and mitigated risks undertook, for example, extensive political economy analysis and drew lessons from the World Bank's previous operations and economic and sector work. A similar finding arises from a desk-based review of IEG evaluations and validations in CY20–23 for key factors linked to the performance of IFC investment projects. The review showed that, when assessing the risk of investment projects, IFC must focus on client quality and on broader factors related to market developments in the sector and the country's macroeconomic challenges.

Adaptive management can help respond to client challenges in country programs and World Bank operations. Adaptive management is an iterative approach to decision-making whereby interventions and portfolios are adjusted based on evidence and evolving context (World Bank 2020d). As such,

it encompasses managing risk effectively, seizing new opportunities, and developing fit-for-purpose solutions. An IEG analysis of Country Opinion Surveys found that when respondents had less favorable perceptions of adaptive practices, such as responsiveness, flexibility, and accessibility, the country programs had lower Bank Group performance and development outcome ratings. Furthermore, analysis of text discussing adaptive management in operations found a statistically significant correlation between positive sentiments on adaptive management and higher outcome ratings. Specific examples of risk-responsive adaptations included changes to scope, timelines, activities, results frameworks, budgets, and procurement. Examples of adaptive management outside of risk management included scaling up projects based on lessons learned or supporting new activities through cost savings during implementation.

In IFC advisory services projects, however, adaptive management did not always result in improved development effectiveness ratings, particularly with respect to restructuring during COVID-19. To better understand the reasons for declining IFC advisory services ratings despite consistent IFC work quality, IEG undertook a desk review of 31 projects validated by IEG for which development effectiveness was rated mostly unsuccessful or worse, but work quality was rated satisfactory or excellent. For 25 of these projects, COVID-19 strongly contributed to weak development effectiveness. Nineteen of these projects were restructured, 11 of them explicitly due to COVID-19. Seven of these projects were restructured early (between March 2020 and July 2020). Regardless, the restructuring was insufficient: all 11 projects received weak development effectiveness ratings, demonstrating that adaptive management was inadequate in these projects. Other prevalent factors in the weak development effectiveness ratings included the client's commitment or motivation (11 projects), change in scope or premature termination of advisory services (11 projects, 5 of which were at IFC's initiative, while 6 were not), and project design (8 projects).

The consolidation of all Bank Group guarantees under MIGA presents both a risk and an opportunity. Until recently, MIGA provided only political risk guarantees to the private sector and was responsible for the supervision of only environmental and social safeguards and country risk assessment. The consolidation of all current World Bank and IFC guarantees under MIGA may

require MIGA to develop the capacity to supervise both public and private sector guarantees.

Client Quality and Capacity Building

Client selection and complementary capacity building are powerful levers that IFC can use to positively influence development outcomes. Client quality includes the ability, technical expertise, or track record of the IFC client, including the quality of the client management team and their skills, contractor competency, familiarity, and acumen. This factor was highlighted in 41 percent of CY20–23 IFC investment projects that IEG reviewed. It had a positive influence on IFC investment development outcomes 75 percent of the time overall but less often for Africa (41 percent) and FCS projects (50 percent). Selecting clients with proven or promising business models, good financials, strong risk management frameworks, and flexible business strategies can mitigate business risk and help clients, particularly financial institutions, withstand shocks like COVID-19. This review of IEG evaluations and validations showed that in challenging contexts where selecting high-quality clients may not be feasible, IFC can influence client quality by providing support for capacity building.

The World Bank can also undertake institutional capacity building to address challenges associated with lower ratings. Various World Bank and external studies have identified institutional capacity as a critical issue for improving development effectiveness (OECD 2008; Otoo et al. 2009; World Bank 2005a, 2005b, 2017a, 2018a, 2018b, 2022b; World Bank Group 2017). In the World Bank, institutional capacity building means improving the effectiveness of country development by changing the formal and informal rules that structure interactions across multiple organizations (World Bank 2005b, 2018b; World Bank Group 2017). *RAP 2024* finds that there is a nonlinear relationship between the number of institutional capacity challenges and the outcome rating. Specifically, the negative effect on outcome ratings becomes more pronounced when there is more than one challenge, with a more noticeable downward shift in outcome ratings as the number of challenges identified increases from one to three. Substantial improvements in outcome ratings can be achieved through mitigating institutional capacity risks. Successful operations often included extensive efforts to build institutional

capacity, such as developing formal coordination mechanisms and shared reporting systems.

Results Monitoring

Gaps in the design and implementation of results monitoring are a persistent concern identified across IEG evaluations. These evaluations include project and Country Program Evaluations covered by *RAP* and recent IEG thematic evaluations (World Bank 2024a, 2024b).

Results frameworks in Country Partnership Frameworks have persistent shortcomings that affect their ability to support implementation. The evidence indicates that a weak results framework is "a key determinant of unsatisfactory outcome performance at the country program level" (World Bank 2015d, 1). Nevertheless, 83 percent of Completion and Learning Review Validations reported major inadequacies in results frameworks.[2] Moreover, there has been no substantial improvement in countries with two Completion and Learning Review Validations: out of 81 countries, 16 improved their results frameworks to a good rating, while 9 declined to fair or below in their most recent Completion and Learning Review Validation. Frequent shortcomings are noted with the intervention logic and the chosen indicators. For example, indicators may focus on inputs or outputs rather than outcomes, or they may be overly reliant on indicators of operations that fail to capture the full extent of the country program objectives and do not properly account for the contributions of advisory services and analytics, IFC, MIGA, policy dialogue, development policy financing, or the Bank Group's convening role. Although Performance and Learning Reviews often adjust the results framework, many weaknesses remain unresolved. Previous IEG reports raised these concerns and also found that these practices generate incentives not aligned with an outcome orientation at the country level (World Bank 2020d, 2022c).

MIGA is behind on submitting self-evaluations of its guarantee projects, which prevents IEG from having an accurate picture of its overall development outcome ratings. A total of 19 MIGA projects are pending self-evaluations for the FY21–23 period (45 percent of planned self-evaluations during this period). Of these 19 projects, 11 involved engagements from

MIGA's legal team, including renegotiation of possible investment term modifications. As a result, self-evaluations of these MIGA projects were postponed. The pending self-evaluations from MIGA mean that it is difficult to comment too definitively on trends in development outcome ratings.

IFC could improve the measurement of outcomes, particularly market outcomes, by recording more complete information about projects. IEG conducted a desk-based review of 173 IFC investment projects evaluated and validated by IEG during CY21–23. The review identified 842 individual outcomes (676 project-level outcomes and 166 market-level outcomes). IEG could not verify 96 outcomes, of which 76 were project-level outcomes and 20 were market-level outcomes. This represents 11 percent of the total project-level outcomes and 12 percent of the total market-level outcomes. During the *RAP 2023* Board discussions, IFC management noted that the introduction of the Anticipated Impact Measurement and Monitoring system—an ex ante monitoring tool—in 2017 overcomes the issue of verifying outcomes. To assess this assertion, IEG conducted an analysis of 21 projects evaluated and validated by IEG with "live" Anticipated Impact Measurement and Monitoring scores (projects that IFC assigned ex ante Anticipated Impact Measurement and Monitoring scores at Board approval). The analysis found that 22 percent of outcomes did not have an indicator in the tracking system (17 percent of project-level outcomes and 43 percent of market-level outcomes). Moreover, IFC could not track most market-level outcomes because these outcomes did not have indicators or indicators were never tracked. Therefore, the identification and tracking of outcome indicators, particularly for market outcomes, remains a challenge.

Despite improvements in monitoring and evaluation quality over time, one-third of World Bank operations received ratings of modest or negligible. The average monitoring and evaluation quality rating of investment project financing and Program-for-Results increased from 2.1 in FY13 to 2.6 in FY20. Since FY20, the average rating has plateaued between 2.6 and 2.7. This indicates that although 64 percent of operations were rated as substantial and above in FY23, there are few examples of high-quality practice, and more than one-third of operations remain with inadequate monitoring and evaluation practices. World Bank projects with challenges in project data and monitoring have lower outcome ratings. Challenges included poorly

designed or misaligned metrics that failed to capture intended outcomes, difficulty setting appropriate targets because of missing baselines, and low quality of progress reporting. Recent evaluations, including *RAP 2021* and *RAP 2023*, have identified opportunities for enhancing indicators, data availability and baselines, and reporting and supervision.

[1] A World Bank project's development outcome rating reflects the extent to which major relevant objectives were achieved, or are expected to be achieved, efficiently. For IFC investment projects, the development outcome rating reflects the extent to which the project achieved its intended development objectives and delivered sustainable results. For IFC advisory services, a project's development effectiveness reflects the extent to which its intended development results are achieved, strategically relevant, and efficient. For MIGA guarantee projects, a project's development outcome reflects the extent to which the project achieved its intended development objectives and delivered sustainable results.

[2] The Completion and Learning Review Validation (CLRV) was called the Completion and Learning Review Review (CLRR) before May 1, 2023.

Management Comments

Management of the World Bank Group welcomes the Independent Evaluation Group (IEG) report *Results and Performance of the World Bank Group 2024: Managing Results in an Uncertain World* and thanks the team for addressing the comments provided. This marks the 14th annual *Results and Performance of the World Bank Group (RAP)* report and covers a period marked by high levels of uncertainty, external shocks, and increasingly risky contexts for the Bank Group's work. Management welcomes the report's recognition of the Bank Group's commitment to prioritizing and allocating more resources to low-income countries and countries classified as fragile and conflict-affected situations (FCS) and responding proactively to domestic and external shocks, particularly during the COVID-19 pandemic and ongoing global crises. The insights provided by the report are valuable for improving project preparation and supporting adaptive management during implementation.

World Bank Group Overall Comments

Management underscores the Bank Group's progress in strengthening operational effectiveness, noting that further efforts are underway to enhance development outcomes. Management welcomes the report's findings, particularly the improvement in development outcome ratings, which rose from 68 percent of country programs rated moderately satisfactory or above in FY 2013 to 78 percent in FY20, surpassing the 70 percent corporate target set in the Corporate Scorecard for FY19–23. Management notes that reforms are underway in the Country Engagement Framework, along with the launch of the new Bank Group Scorecard to drive further improvement and selectivity in the Country Partnership Framework (CPF). Greater selectivity in CPF objectives is expected through greater alignment with the new Bank Group Scorecard outcome areas, Bank Group programming will be further integrated, proactive risk mitigation will be strengthened through annual business planning, and the quality of CPF results frameworks will be improved through alignment with the new Bank Group Scorecard. Management's efforts to revamp country advisory services and analytics products, such as the Country Growth and Jobs Reports, Country Private Sector Diagnostics,

and Country Climate and Development Reports, will help inform Bank Group programming, using all Bank Group instruments, including the International Finance Corporation (IFC) and the Multilateral Investment Guarantee Agency (MIGA), and upstream and advisory services.

Management acknowledges the report's finding that challenging operational environments may impede improvements in the Bank Group's performance and outcome ratings and is committed to continuing to prioritize efforts to mitigate these risks and strengthen development outcomes. The report highlights an emerging gap between outcome ratings and Bank Group performance ratings for FCS contexts, notably for projects with extended exposure times to the pandemic and other compounding shocks in the challenging global environment. Management notes that these observed results likely reflect an interaction effect of both FCS and support for crisis response since managing for results was more complex in FCS countries during the COVID-19 period. While appreciative of the report's insights and conclusions, management notes that some of them may be based on small sample sizes that can be influenced by minor percentage changes within the margin of error. The Bank Group is committed to improving monitoring and evaluation ratings, and management is pleased to report that proactive measures have been taken, which include refining monitoring and evaluation methodologies, revising indicators, adjusting targets through restructuring, and collecting additional evidence on project achievements.

Management welcomes the notable improvement in collaboration among the World Bank, IFC, and MIGA over the past decade and is committed to further enhancing synergy and impact. Stronger collaboration across the Bank Group will be driven by a shared view of strategic priorities in client countries in alignment with Bank Group corporate priorities and systematic efforts to leverage each other's capacities across CPF periods and in multiple sectors. The One World Bank Group approach is also central to country programming, diagnostics, and operations to strategically determine when to tap into public or private sector solutions—or both.

World Bank Management Comments

Management appreciates the report's finding that despite the challenging global environment of 2020–23, the World Bank successfully maintained stable overall ratings for its operations and monitoring and evaluation (M&E) with the share of operations validated moderately satisfactory or above increasing to 84 percent. The report finds a sustained upward trend in Bank performance ratings, rising from an average of 3.8 (out of 6) in FY13 to 4.3 in FY23, plateauing at 4.3 during the COVID-19 pandemic and global crises (FY20–23). The report primarily attributes the stability in the ratings to a 6 percent increase in the proportion of projects with full or partial exposure to FCS—historically associated with lower average ratings for much of the period covered—amplifying their impact on the overall World Bank average outcome rating. Management notes that the pandemic, restricted movements, and limited monitoring capabilities and evaluation efforts potentially affected Bank Group performance ratings. Management also wishes to emphasize the importance of the World Bank's adaptive risk management approach and the significant efforts made to maintain portfolio performance despite these complex challenges, which demand more targeted, flexible, and responsive approaches, as evidenced by a realignment of 60 percent of its portfolio. Management notes that IEG's identification strategy does not allow for empirically isolating the effects of COVID-19 and disentangling the negative effects from the mitigating effects, inferring that the negative shock from the pandemic and the mitigating effects of the World Bank's operations would have had a net zero effect on World Bank performance. In a counterfactual scenario without the pandemic and global shocks, the positive pre-pandemic trend in year-on-year improvements in overall performance and M&E ratings would likely have continued.

Management acknowledges IEG's insightful observation of the association among proactive risk identification, mitigation measures, and improved outcome ratings. The report's finding underscores the validity of management's commitment to strengthen adaptive management practices through several promising avenues, including proactively restructuring operations, rolling out the crisis preparedness and response tool kit, scaling up programmatic approaches under the Global Challenge Programs and multiphase

approaches, and accelerating the implementation of effective operations. The report's findings support management's plans to continue strengthening adaptative risk management practices, including the calibration of the Systematic Operations Risk-Rating Tool risk ratings based on updated risk assessments and implementation of risk mitigation measures during project implementation. Management is also planning on adopting agile and innovative approaches using new technologies to collect and analyze emerging risk data to inform timely course correction.

Management recognizes the report's finding that client capacity challenges are linked to lower performance, particularly in FCS, International Development Association (IDA), and blend countries. Simplified design can improve performance and should feature interventions aiming at building client capacity and adequate identification and mitigation of institutional capacity risks. Management reaffirms that several initiatives have been taken to address the client's institutional capacity gaps. The World Bank's strategic decision to expand in FCS contexts has been coupled with specific initiatives under the fragility, conflict, and violence strategy that are designed to build capacity among government agencies and partners and improve their effectiveness. The suite of tools emphasized in FCS contexts includes the Geo-Enabling Initiative for Monitoring and Supervision, hands-on expanded implementation support, World Bank–facilitated procurement, and others. The Geo-Enabling Initiative for Monitoring and Supervision builds capacity in digital data collection and analysis among government agencies and partners. New approaches such as hands-on expanded implementation support and World Bank–facilitated procurement directly assist borrowers in their procurement efforts when they are faced with capacity, market, or supply obstacles (both hands-on expanded implementation support and World Bank–facilitated procurement proved especially effective as part of the World Bank's COVID-19 pandemic response). Third-party monitoring and implementation arrangements, partnerships including with United Nations agencies and regional organizations, and increased field presence and staff facetime will further support client capacity in FCS. Management has recently launched a new Client Capacity Program to address project implementation units' operational learning needs during portfolio

implementation. The Change Management Program for Procurement also includes a strong focus on client capacity and training.

Management welcomes IEG's recognition of the doubling in the share of projects with M&E quality ratings of substantial or above, rising from 29 percent in FY13 to 64 percent in FY23. While recognizing the need for further work to improve the remaining one-third of projects that fall below the "substantial" threshold, management believes that the promising trend observed over the last 10 years, including at the CPF level, indicates that such gains can be maintained. Management agrees that further achievements would be facilitated by the continued provision of enhanced staff training and resources to improve M&E capacity, improvements in results frameworks and indicator selection, and enhancements in data analysis. Management concurs that continued progress in improving M&E systems, including moving results frameworks from inputs and outputs to outcomes, will be critical for the achievement of World Bank objectives and for providing adequate data for the new Bank Group Scorecard.

Management acknowledges the report finding that projects with exceptionally long preparation times (exceeding the 90th percentile) tend to underperform, noting that long lag times are often due to a lack of client buy-in or overly complex project designs. Substantial efforts are underway to shorten average project preparation time, using risk-based approaches to target more resources on more complex and high-risk operations. The report indicates that both undermine the ability of projects to achieve their objectives. Management welcomes further consultation as it deepens its analysis of the complex relationships among preparation time, design quality, and implementation readiness and refines its nuanced approach to preparation time, considering factors such as project complexity, country and sector context, and political stability.

International Finance Corporation Management Comments

Management notes positively the improving trends in the medium term, particularly given the challenging external context of recent years, while acknowledging a slight decline over the long term. Development outcomes for IFC investment projects improved by 10 percentage points over the medium

term from 41 percent to 51 percent, while declining 2 percentage points over the long term. The improvement over the past five years is encouraging, particularly in Africa, which also saw the largest increase in the share of active portfolio and where operating conditions are often challenging. Furthermore, management believes that learning from positive trends is as important as learning from negative trends and found IEG's identification of IFC investment subgroups whose performance improved over the medium term, including Disruptive Technologies and Funds, which saw a 34 percentage point increase in the development outcome rating from 14 percent to 48 percent, helpful. While in-depth analysis to identify drivers of performance over the medium term is beyond the scope of this year's *RAP,* IFC management would welcome for future *RAPs* to examine the drivers of medium-term performance.

Management acknowledges the weak outcome ratings for IFC investments in FCS, IDA, and blend countries, where factors outside of IFC's control continue to be key drivers of performance. Achieving high development impact in projects in FCS and IDA-eligible FCS countries remains challenging, including due to economic issues, increased civil unrest and fragility, and associated business risk. However, management's proactive risk management and capacity-building efforts have allowed it to maintain a critical presence in these regions and play a countercyclical role at a time when foreign direct investment flows to these markets saw substantial decreases. Management notes that IFC's upstream and advisory services have been instrumental in improving project design, increasing private sector capacity, and fostering sustainable market outcomes, particularly in IDA and FCS countries. IFC has refined its approach in FCS, with upstream and advisory work tailored to highly fragile markets. Through initiatives such as FCS Africa's Local Champions Initiative, IFC identifies and supports potential clients in areas such as financial management and environmental and social compliance to help build a pipeline of bankable projects in FCS. For Middle East and Pakistan (the region after Africa with most FCS countries), an IDA-FCS Investment Accelerator program is being developed. The program will coordinate all IFC's relevant tools and advisory products (environmental and social advisory, corporate governance advisory, feasibility support, among others), under a single platform to address the key hurdles, helping develop a sustainable pipeline of investable projects

and sponsors. IFC is also systematically scaling up its conflict sensitivity interventions to better work in these markets.

Management acknowledges the improvements in development outcomes in Europe and South Asia and the decline in investment project ratings in the remaining regions over the long term. Latin America and the Caribbean contributed to the decrease in IFC-wide development outcomes more than other regions due to its large share of reviewed projects (25 percent) and the decline in its ratings over the long term (from 60 percent of projects rated mostly successful or better to 47 percent). This is tied to a strategic shift in upper-middle-income countries in the Latin America and the Caribbean region, where IFC is taking more risks to innovate and test new business models by developing new products and markets and working with new clients. Management thanks IEG for bringing to our attention the gaps in the Latin America and the Caribbean region's work quality that need improvement. Lessons highlighted will serve to instruct teams to continue testing similar business models and to develop stronger risk mitigants for future projects.

Management welcomes IEG's plan to develop a taxonomy of factors linked to the development effectiveness of IFC advisory services for future studies. Unlike for IFC investment projects, there is no typology in IEG for factors linked to performance for IFC advisory services projects at this point. This limits the analysis of challenges that can be conducted and thus management's understanding of levers we can employ to drive development effectiveness, after controlling for external factors. Management would thus find the taxonomy helpful in promoting learning on the effectiveness of future advisory services engagements and understanding what makes for high-quality preparation in IFC advisory services projects.

Management finds the analysis on challenges and levers insightful and relevant and welcomes the additional preliminary analysis that IEG has conducted on work quality, too. These analyses offer a helpful framework to identify areas that can influence development results, including those within and outside management's control. This is particularly relevant as IFC continues to focus on more challenging markets where issues around client quality, risks, and lack of market information make employing available levers more challenging. Overall, IFC is providing capacity building in these

operating environments to expand the pool of high-quality clients. As we continue to focus on more challenging markets, it would be helpful for IEG and Bank Group management to reflect more generally on how the outcomes and performance should be measured across time when the risk environment has fundamentally changed, and there is an expectation for the institution to take more risks. It is important to note that not all challenges and risks—even if noted on the front end—can be mitigated.

Management continues to strengthen our capabilities for increased tracking of outcomes through enhancements to the Anticipated Impact Measurement and Monitoring (AIMM) system, and we find that IEG's gap assessment is misaligned with the AIMM framework and approach. In FY24, management launched the AIMM Navigator, a technology platform that allows IFC to track development impact more systematically and consistently within the AIMM framework. IFC tracks outcome indicators related specifically to impact claims and those that are needed for corporate reporting purposes. IFC intends to track only the most significant indicators that underpin the AIMM assessments and are reflected in the Development Impact Indicator table, which is included in the Board paper. In contrast, IEG's typology, developed for *RAP 2021*, tries to cover all potentially relevant outcomes included in the AIMM narratives, driving the reported gap in outcome tracking.

Multilateral Investment Guarantee Agency Management Comments

MIGA notes *RAP 2024's* confirmation that MIGA has maintained a consistent trend in development outcome ratings, between 65 percent and 70 percent. Following a constructive and evidence-based dialogue with IEG colleagues, both parties agree that the one percent decline in development outcome ratings over the long term is not statistically significant at the MIGA-wide level. However, MIGA also recognizes the challenges in interpreting the ratings due to the delay in completing the current self-evaluation program.

MIGA acknowledges IEG's findings on the lower outcome ratings among projects in IDA and blend countries. These ratings indicate challenging environments in these countries in achieving expected development outcomes. MIGA noted IEG's additional analysis, which highlighted that key factors

affecting outcomes, such as cost overruns, construction delays, and legal and regulatory risks, are beyond MIGA's control.

MIGA emphasizes its strong record in monitoring and managing project risks throughout the life of the guarantee contracts. This involves monitoring project and country risks, development indicators, and environmental and social risks. MIGA plays a proactive role as an honest broker in helping resolve emerging contractual or country risks.

MIGA highlights that development outcome ratings for fragility, conflict, and violence–status projects over the past six-year cycle are at 75 percent, surpassing MIGA's overall average. For the last decade, fragility, conflict, and violence–status projects have consistently achieved strong development outcome ratings, reflecting MIGA's effectiveness in delivering development impact in challenging environments. MIGA continues to prioritize this area, leveraging lessons learned to enhance development outcomes.

MIGA acknowledges IEG's observation regarding the realization of foreign investment effects, which extend beyond direct project benefits. It is important to note that these evaluations took place before the introduction of the ex ante development impact assessment tool. MIGA had already been extensively focused on specifying and assessing foreign investment effects. These experiences led to improvements that culminated in the Impact Measurement and Project Assessment Comparison Tool framework, MIGA's ex ante development outcome assessment tool, piloted in FY19 and fully launched in FY20. The recent introduction of the Bank Group Scorecard, which features a stronger outcome focus, has also helped us sharpen our attention in development impact assessment and the tracking of such effects, especially those aligned with higher levels of outcomes. Furthermore, the launch of the Bank Group guarantee platform enhances the collaboration across Bank Group institutions further, with opportunities for furthering country-level collective engagement that can support stronger project supervision and monitoring.

In this regard, MIGA welcomes IEG's ongoing commitment to various lenses of outcome analysis across the Bank Group and believes that IEG can enrich data quality assurance and methodological fine-tuning for sharpening the outcome assessment.

1 | Introduction

Background and Audience

Results and Performance of the World Bank Group (RAP) 2024 presents an annual review of the Bank Group's operational and country effectiveness that draws principally on evidence from the Independent Evaluation Group (IEG). This is the 14th annual report on the results and performance of the Bank Group. The report aggregates and interprets evidence related to the results and performance of the World Bank, which includes the International Bank for Reconstruction and Development and the International Development Association (IDA); the International Finance Corporation (IFC), including IFC investment services and IFC advisory services; the Multilateral Investment Guarantee Agency (MIGA); and Bank Group country programs.[1]

The main audience for *RAP 2024* is the Bank Group's Board of Executive Directors. *RAP 2024* seeks to help readers understand performance ratings and factors linked to them as the Bank Group undertakes its Better Bank initiatives (box 1.1). The report also contains in-depth analyses that can assist management and staff with the ongoing design and support to the implementation of operations and country programs.

Throughout this *RAP*, we refer to three aspects of the ratings:

» Trends. We describe overall performance trends by key ratings and important subgroups. Like a Scorecard report, *RAP 2024* presents trends, describes changes in trends, and provides a deeper analysis of specific issues based on the IEG validations of Bank Group self-evaluations.

» Challenges. Having described trends, we then identify challenges associated with performance, often in relation to subgroups. Challenges are persistent factors negatively linked to ratings. They are often drawn from institution-specific taxonomies because each institution responds to different kinds of challenges.

» Levers. Finally, we identify levers, which are actions within the Bank Group's influence that management can take to address challenges or improve performance. Levers are also institution specific. Hence, a factor that appears as a lever for one institution may appear as a challenge for another.

Box 1.1. Relevance of *Results and Performance of the World Bank Group 2024* to the Better Bank Initiatives

Results and Performance of the World Bank Group 2024 provides evidence of the Bank Group's evolution toward becoming a Better Bank. The Better Bank initiatives implement changes in how the Bank Group operates, building on processes initiated by governors at the annual meetings in 2022 as part of the evolution. The initiatives for becoming a Better Bank seek to assist countries with navigating intertwined crises, tackling global challenges, and achieving the vision of a world free of poverty on a livable planet.

The findings presented in *Results and Performance of the World Bank Group 2024* relate to a cross-cutting issue and three of the main initiatives being undertaken within the Better Bank initiatives:

- A cross-cutting issue is the Bank Group reinforcing its commitment to support people and countries affected by fragility, conflict, and violence.
- Joint country representation aims to enable truly integrated solutions that span both the public and private sectors, leveraging the entirety of our knowledge and experience and amplifying our collective impact.
- The operational efficiency and effectiveness initiative aims to quicken the pace and simplify Bank Group policies, processes, and systems.
- The World Bank Group Scorecard is a strategic management tool that is used to drive action for results.

Sources: Independent Evaluation Group; World Bank 2024c.

In this *RAP*, we systematically document trends, challenges, and levers for each institution as follows:

- Trends. Outcome ratings across the Bank Group's institutions have not consistently increased over the long term, mostly because of changes in portfolio shares and shifts in ratings in challenging contexts.
- Challenges. The top challenges linked to performance include the following:
  - World Bank: institutional capacity of implementing agencies and key stakeholders, operational design, project data and indicators, and financial management
  - IFC: business risk, asset quality, economic issues and civil unrest, and declining development effectiveness ratings
  - MIGA: cost overruns or construction delays, project company quality, and legal or regulatory risk
  - Country programs: relevance, risk identification and mitigation, and support to implementation
- Levers. The Bank Group institutions have opportunities to reinforce performance by the following:
  - Improving operational and project design (World Bank and IFC), which entails the One World Bank approach, ensuring appropriate preparation time and processing time (World Bank and IFC) and identifying and mitigating risks (World Bank and IFC)
  - Ensuring client quality (IFC) and undertaking capacity building (World Bank and IFC), which entails developing adaptive management (World Bank and country program)
  - Improving results monitoring, which entails recording more complete information (IFC and MIGA)

Objectives, Question, Scope, and Use

The objective of *RAP 2024* is to identify trends in the Bank Group's performance ratings, challenges that may constrain performance, and levers that can be used to improve performance. To meet this objective, the main question to be addressed by *RAP 2024* is, What do IEG's validations tell us about how the Bank Group's performance changed over time and across subgroups?

The scope of *RAP 2024* is based on the principles of continuity, symmetry, and innovation. Continuity is provided through standardization between *RAPs*, which allows for comparison of the Bank Group's performance ratings across key breakdowns between 2020 and 2023. *RAP 2024* continues to review performance over a 10-year period based on ratings from IEG validations and standardized methodologies undertaken previously for analyzing factors linked to performance. Symmetry is maintained within this report among Bank Group institutions through the analysis of trends, challenges, and levers. However, we acknowledge that the institutions are different in nature; therefore, perfect symmetry is not possible. Hence, we discuss certain topics in more depth for some institutions than for others. To innovate, *RAP 2024* includes analysis of (i) project preparation time for the World Bank, processing time for IFC investment projects, and preimplementation scoping time for IFC advisory services projects; (ii) data from the Systematic Operations Risk-Rating Tool (SORT) for the World Bank; (iii) factors specific to work quality and additionality that are associated with development outcomes in IFC investment projects; (iv) the Country Opinion Surveys (COSs); and (v) levers related to outcome indicators and measuring results for IFC.[2] Each innovation is defined, and we elaborate on findings in the chapters for the institutions.

RAP 2024 can inform learning and accountability. The evidence presented in *RAP 2024* focuses on statistically significant changes and substantive patterns in performance relevant to the portfolio level, although it does not identify causal relationships. Drawing on this evidence, the Bank Group can use *RAP 2024* for learning because it helps the Bank Group further examine the challenges linked to performance and the levers that can be pulled to enhance its portfolio. *RAP 2024* aids accountability because it helps the Board understand trends and changes in portfolio performance. Being informed about past performance and the outlook for the portfolio helps the Bank Group make informed choices on investment. As with previous *RAPs*, the 2024 report does not provide formal recommendations based on its findings. Nevertheless, the identification of levers provides information regarding what the Bank Group can do to improve performance. In addition, to aid learning and accountability, a set of online charts for the World Bank accompanies this report, providing further breakdowns of data.

Methods

RAP 2024 applies a structured methodological design with the two main steps of data set construction followed by analysis. The three main data sets constructed across institutions relate to ratings, factors linked to performance, and outcome types. These data sets are the essential ingredients for the analysis of the *RAP*. In the second step, structured analysis is undertaken by drawing on a single or multiple data sets. The statistical analysis of the data sets includes both descriptive and inferential techniques. A text analysis of IEG evaluations and validations identified factors linked to performance, sentiments associated with their descriptions, and examples of challenges and levers. Supervised machine learning models were applied to automatically categorize text into factors linked to performance through logistic regression, transformer, and Naïve Bayes models. Appendix A provides full details on the sampling, data sources, and analytic methods used for *RAP 2024* and their strengths and limitations.

Bank Group institutions use different frameworks and methodologies when rating operations. Outcome and performance ratings mean different things for the different institutions (box 1.2). Each chapter defines key ratings for the respective institutions. All rating methodologies align with good practice standards for evaluating public and private sector projects, as established by the Evaluation Cooperation Group of multilateral development banks (ECG 2012).[3, 4]

Box 1.2. Key Terms and Concepts in *Results and Performance of the World Bank Group 2024*

Performance. The extent of achievement of key aspects of design, support to implementation, or results, based on ratings defined by the World Bank, the International Finance Corporation, and the Multilateral Investment Guarantee Agency, and for World Bank Group country programs.

Self-evaluation. A formal assessment of a project, program, or policy conducted by or for those in charge of the activity. In the Bank Group, self-evaluation takes the form of a systematic written account of the performance of a project or operation,

(continued)

Box 1.2. Key Terms and Concepts in *Results and Performance of the World Bank Group 2024* (cont.)

with ratings assigned based on the criteria defined in guidelines to ensure comparability among reports.

Ratings. Ratings are quantitative summaries for assessing performance relative to an operation, a country program, or advisory services objectives. Ratings summarize the self-evaluation narrative into categories or values that enable aggregation.

Validation. The Independent Evaluation Group's independent, critical review of the evidence, results, assessments, and ratings from self-evaluation.

Portfolio composition. The *Results and Performance of the World Bank Group* portfolio dynamics mirror those of a stock market portfolio. The portfolio consists of various subgroups, similar to a stock market portfolio. Changes in the allocation of operations among these subgroups (akin to stock positions) or their ratings (like stock prices) affect the weighted average of the overall portfolio. Consequently, the overall average rating shifts are determined by both shifts in subgroup ratings and changes in the portfolio's composition.

Significance. A measure that indicates whether the results of a statistical analysis are unlikely to have occurred by chance, typically determined by a p value threshold (such as $p < .05$).

Sources: Independent Evaluation Group; World Bank 2021c, 2023e.

Organization of the Report

The report is organized into six chapters. Following this introduction, chapter 2 focuses on the results and performance of the World Bank, chapter 3 on the results and performance of IFC, chapter 4 on the results and performance of MIGA, and chapter 5 on the results and performance of Bank Group country programs. Each chapter examines performance trends, analyzes challenges, and describes levers. Chapter 6 provides concluding remarks on the relevance of the findings to the Better Bank initiatives. These chapters are complemented by appendixes that provide additional information and supporting methodologies.

[1] The institutions define performance differently, measure it differently, and call it by different names. We use the term *performance* as a shorthand to capture in a single word all the ratings relevant to these institutions for this report. Box 2.1 lists and defines the specific ratings.

[2] This year's *RAP* notes challenges linked to results measurement of the World Bank and in country programs. However, we do not develop the challenges into levers in this *RAP*. Previous *RAPs* and IEG evaluations focused on results measurement systems (for example, *The World Bank Group Outcome Orientation at the Country Level*, *Behind the Mirror: A Report on the Self-Evaluation Systems of the World Bank Group*, and *Results and Performance of the World Bank Group 2023*). This year's *RAP* does not prioritize new analysis in these areas for the World Bank, unlike IFC. A planned early-stage evaluation by IEG will review developments linked to the Bank Group's new Corporate Scorecard.

[3] The exception to this is technical assistance or IFC advisory services because the Evaluation Cooperation Group has yet to define these rating standards.

[4] There is no framework in IEG to consolidate and report the various performance ratings of the three Bank Group institutions on a single, uniform scale. Currently, the underlying criteria are incommensurable. The Bank Group has started to implement a new Corporate Scorecard, and changes in the Bank Group self-evaluation and IEG validation practices will be discussed as part of this and other Better Bank initiatives. Updates on the implementation of these reforms will inform future *RAPs*.

2 | World Bank

Highlights

World Bank outcome ratings have plateaued during FY 2020–23, along with a shift in the share of projects toward countries classified as fragile and conflict-affected situations. In this period, average ratings remained stable at about 4.3, and the percentage of moderately satisfactory or above ratings showed minimal change, inching up slightly from 83 percent to 84 percent.

Among operations closed since FY20, an increasing proportion of their life span occurred during the COVID-19 pandemic. Yet there are no statistically significant differences in outcome ratings among operations that closed before COVID-19 (no exposure), those that operated for part of their life during the pandemic (partial exposure), and those that were approved and closed within the pandemic (full exposure).

Despite improvements in monitoring and evaluation quality over time, one-third of World Bank operations received ratings of modest or negligible. Certain Global Practices may require particular attention given the introduction of the new Scorecard and indicators.

The portfolio continues to experience challenges with institutional capacity, operational design, project data and monitoring, and financial management that the World Bank can influence. Among operations that addressed each issue, challenges were encountered by 75 percent of operations in institutional capacity, 45 percent in operational design, 46 percent in project data and monitoring, and 76 percent in fiduciary compliance.

Effective risk identification and mitigation and adaptive management emerged as crucial strategies for the World Bank in addressing challenges, especially in countries classified as fragile

and conflict-affected situations. Larger risk reductions during implementation are significantly associated with higher outcome ratings, while those that fail to mitigate high risk throughout their life cycle tend to receive lower outcome ratings.

This chapter examines the performance trends of World Bank operations closed between FY13 and FY23, as evaluated by IEG by June 30, 2024 (see box 2.1 for a description of the main performance ratings). It delves into the challenges encountered during project design and implementation, exploring their impact on outcomes. The analysis also identifies key areas within the World Bank's control that could mitigate these challenges and enhance project performance. We focus on statistically significant changes and substantive patterns evident over four data points in the analysis of the portfolio. Online dashboards can be accessed through appendix B and enable interested readers to undertake their own breakdowns of the data.

Box 2.1. Main Performance Ratings in the World Bank

In assessing the World Bank's performance (figure B2.1.1), the Independent Evaluation Group validates the Implementation Completion and Results Reports through Implementation Completion and Results Report Reviews. The Implementation Completion and Results Reports are completed by operations based on guidance issued by the World Bank's Operations Policy and Country Services. The Independent Evaluation Group also conducts evaluations of operations through Project Performance Assessment Reports, which can adjust the ratings of the Implementation Completion and Results Report Reviews.

Ratings

Outcome. The extent to which a project efficiently achieved, or was expected to achieve, its relevant objectives. The outcome rating brings together three underlying dimensions: relevance, efficacy (objectives achievement), and efficiency. The outcome is rated on a six-point scale: highly satisfactory, satisfactory, moderately satisfactory, moderately unsatisfactory, unsatisfactory, and highly unsatisfactory.

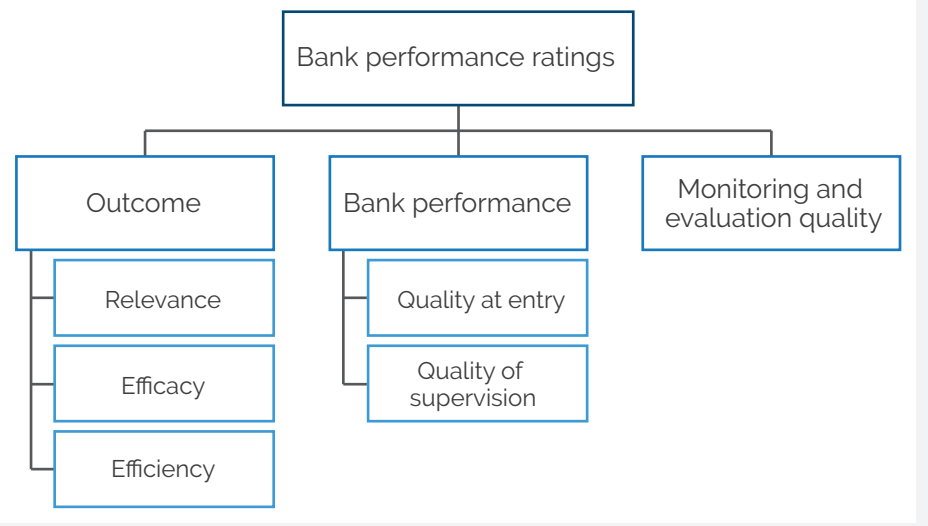
Bank performance. The extent to which services provided by the World Bank ensured quality at entry of the project and supported effective implementation through appropriate supervision. Bank performance and its two constituent elements—quality at entry and quality of supervision—are rated on a six-point scale: highly satisfactory, satisfactory, moderately satisfactory, moderately unsatisfactory, unsatisfactory, and highly unsatisfactory.

(continued)

Box 2.1. Main Performance Ratings in the World Bank (cont.)

Monitoring and evaluation quality. The quality of the design and implementation of monitoring and evaluation arrangements of the operation and the extent to which the results are used to improve performance. It is rated on a four-point scale: high, substantial, modest, and negligible. The monitoring and evaluation rating applies only to investment project financing and Program-for-Results operations.

Figure B2.1.1. Performance Ratings in World Bank Investment Projects



Sources: Independent Evaluation Group; World Bank 2023e, 2024d.

Note: This is the ratings structure for investment project financing and Program-for-Results; development policy financing has a slightly modified ratings structure (see appendix A).

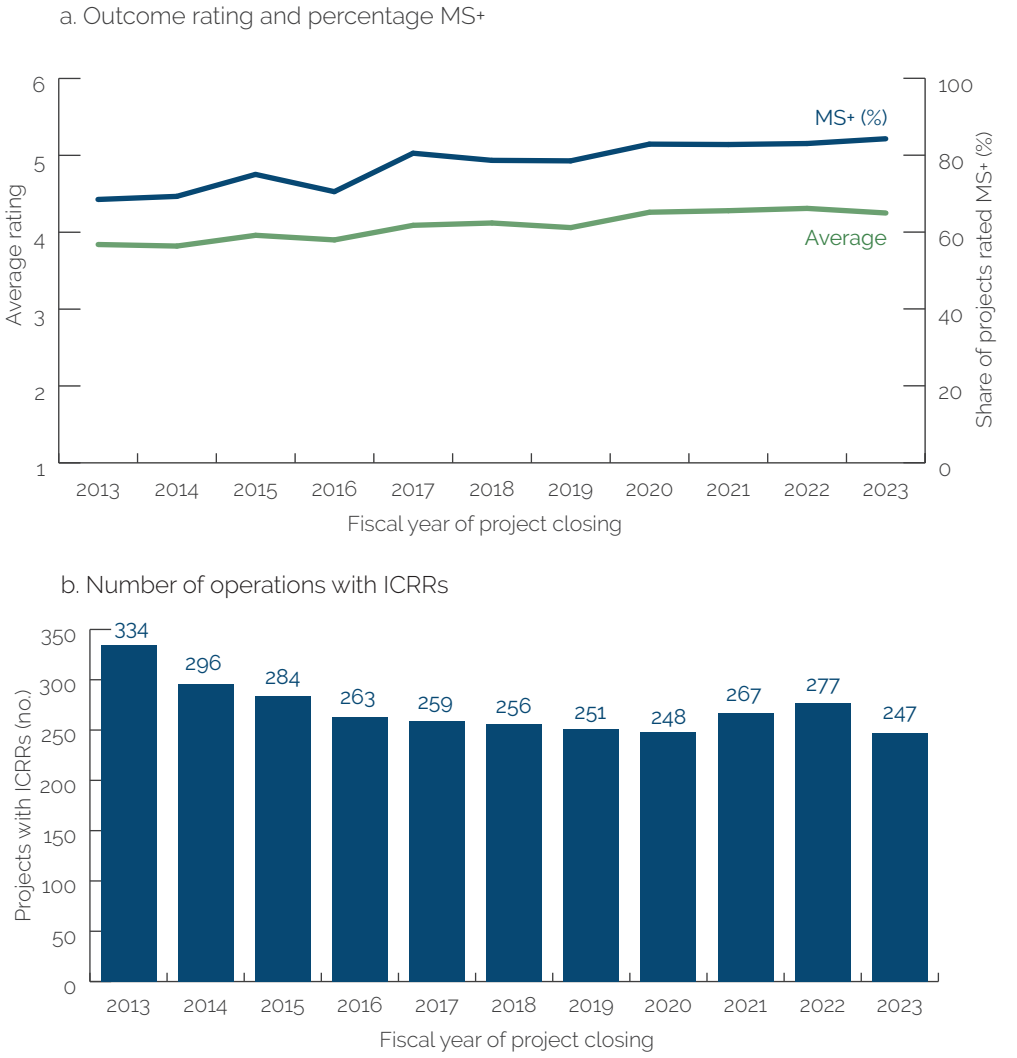
Source: Independent Evaluation Group.

Trends

World Bank outcome ratings have plateaued recently after a long period of increase (figure 2.1). Between FY13 and FY20, portfolio-level outcome ratings saw significant improvement. The average rating increased from 3.8 in 2013 to 4.3 in 2020, while the percentage rated moderately satisfactory or above rose from 68 percent to over 83 percent. In contrast, the FY20–23 period was characterized by plateauing performance. Average ratings remained

stable at about 4.3, and the percentage of moderately satisfactory or above ratings showed only a slight increase, from 83 percent to 84 percent.

Figure 2.1. World Bank Project Rating Trend and Coverage



Source: Independent Evaluation Group.

Note: The *RAP* data have an inherent selection bias in their coverage. Not all projects closed during recent fiscal years, especially FY23, have been evaluated by the Independent Evaluation Group yet. More ICRs of recently closed projects could arrive later. Projects with ICRs and ICRRs completed relatively quickly after closure tend to have higher ratings than those with delayed evaluations. This pattern was analyzed in depth in *RAP 2023*. *N* = 2,982 operations included in *RAP 2024* analysis (*N* = 2,191 for FY13–20 and *N* = 1,039 for FY20–23). The dashboard for further review of rating breakdowns is available (see also appendix B for more details). ICR = Implementation Completion and Results Report; ICRR = Implementation Completion and Results Report Review; MS+ = moderately satisfactory or above; *RAP* = *Results and Performance of the World Bank Group*.

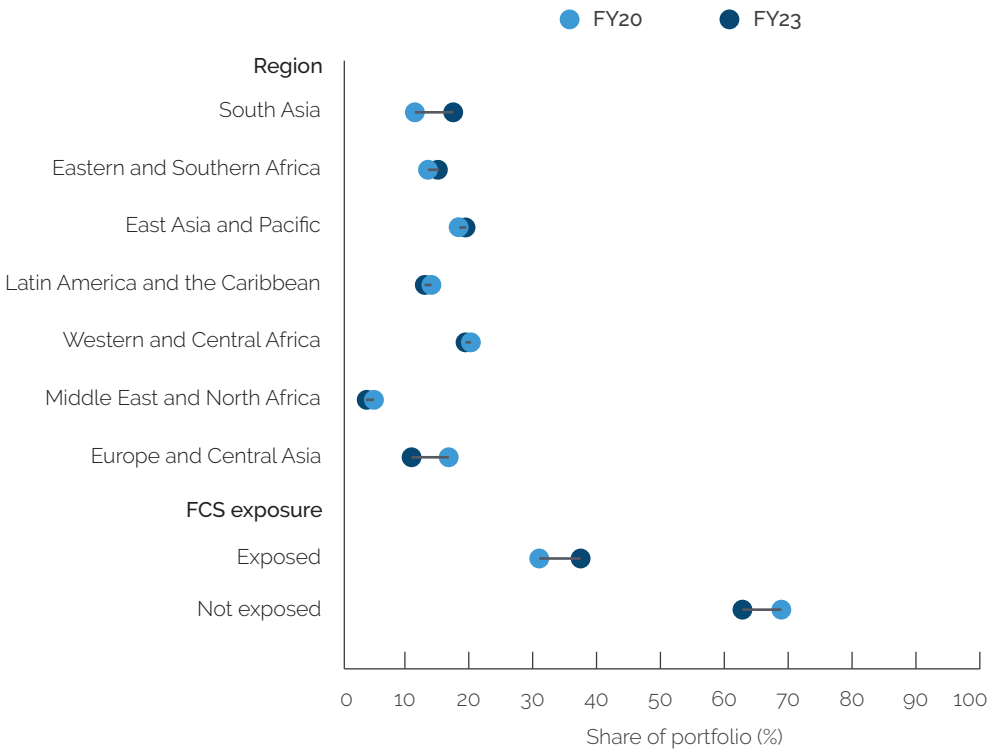
In the period between FY13 and FY20, the increase in the average outcome rating was driven by improvements across a wide range of subgroups within the portfolio. This upward trend is substantive in a wide spectrum of the subgroups, such as Regions, Practice Groups, those exposed to fragile and conflict-affected situations (FCS), agreement types (for example, the International Bank for Reconstruction and Development and IDA), and net commitment sizes. Nearly all subgroups experienced increases in their outcome ratings, with most of these upward shifts being statistically significant.

The recent plateau in portfolio performance occurred amid a notable shift in portfolio composition toward FCS. Between FY20 and FY23, there was no clear trend of increasing or decreasing in ratings across subgroups, and no statistically significant shift was observed for any subgroup. For example, operations with FCS exposure (meaning the country of operation was classified as FCS for at least one year during the operation's lifetime)[1] had their average outcome ratings and those rated moderately satisfactory or above fluctuate, generally at a level below non–FCS-exposed operations, and there are no statistically significant differences in average outcome ratings for FCS operations between FY20 and FY23 or between consecutive years. However, there are substantial changes in the portfolio composition, with shifts in the portfolio share among different subgroups (figure 2.2). For example, the share of operations in FCS contexts, historically with lower outcome ratings, has risen. The proportion of closed operations with full or partial exposure to FCS contexts increased from 31 percent in FY20 to 37 percent in FY23. Conversely, the share of operations in subgroups with historically high outcome ratings declined. For example, the portfolio for Europe and Central Asia decreased from 17 percent to 11 percent. Moreover, the increasing importance of FCS is demonstrated by the shrinking gap in average project volume between FCS and non-FCS operations since FY21 (figure 2.3).

The World Bank's entire lending portfolio has been shifting toward lower-rated contexts (figure 2.4). The lending portfolio is a measure of the total number of all operations active at any stage of a fiscal year. The analysis found that the proportion of projects operating in FCS contexts increased from 16 percent in FY13 to 25 percent in FY24. A similar pattern emerged at the regional level with a rising share of projects in Africa, where outcome ratings tend to be below average. Simultaneously, there is a shrinking share

of projects in East Asia and Pacific, where outcome ratings have traditionally outperformed other Regions. For example, the portfolio is decreasing in countries with high outcome ratings, such as China and Viet Nam. Overall, shifts in portfolio composition for the current lending portfolio mean that there are fewer operations in high-performing countries and more in contexts where average outcome ratings can be lower. The change in distribution toward risky contexts is likely to continue to flow into the closed portfolio that is rated by IEG.

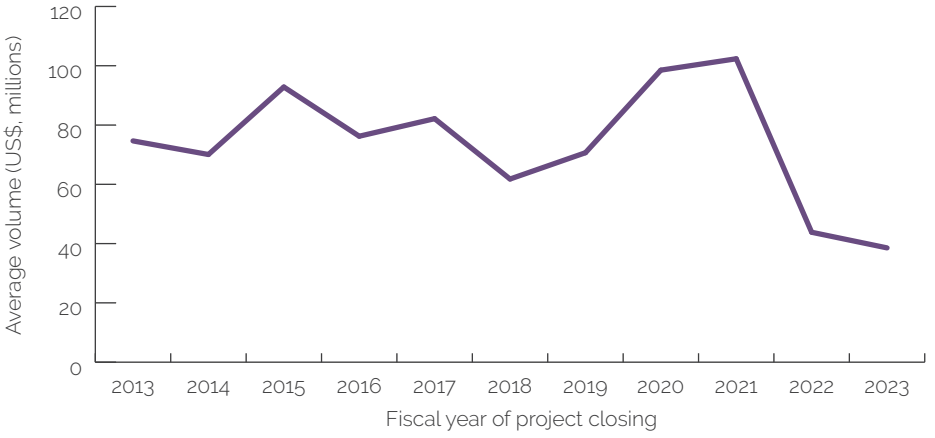
Figure 2.2. Portfolio Composition Shift from FY20 to FY23



Source: Independent Evaluation Group.

Note: Not exposed = the country in which operation was active was never classified as FCS on the annually updated list. Exposed = the country of operation was classified as FCS for at least one year during the operation's lifetime. FCS exposure provides a more accurate estimations of its effect on ratings compared with the country's FCS status in operation's closing fiscal year. A country's FCS status can change over time. For example, a country might be classified as FCS due to an emergency but removed from the list the following year as conditions improve. Using only end-year status would miss the impact of FCS status in other years. FCS = fragile and conflict-affected situations.

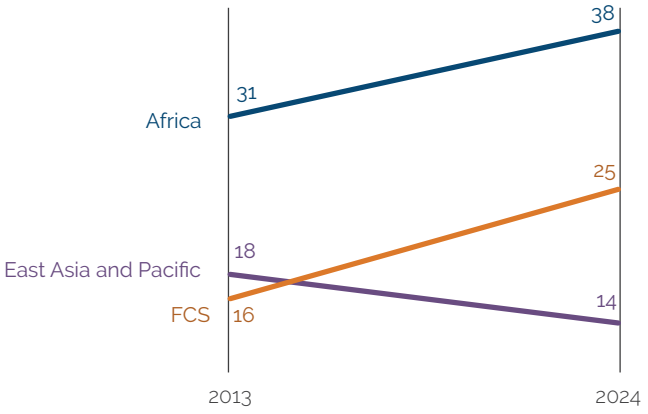
Figure 2.3. Gap in Average Project Volume Between Operations in Fragile and Conflict-Affected Situations and in Non–Fragile and Conflict-Affected Situations



Source: Independent Evaluation Group.

Note: The chart shows the gap in average volume between FCS and non-FCS operations. An upward trend indicates that non-FCS operations have a higher volume. A downward trend shows that the average size of FCS operations is increasing. FCS = fragile and conflict-affected situations.

Figure 2.4. Share of All Lending Projects in Operation (percent)

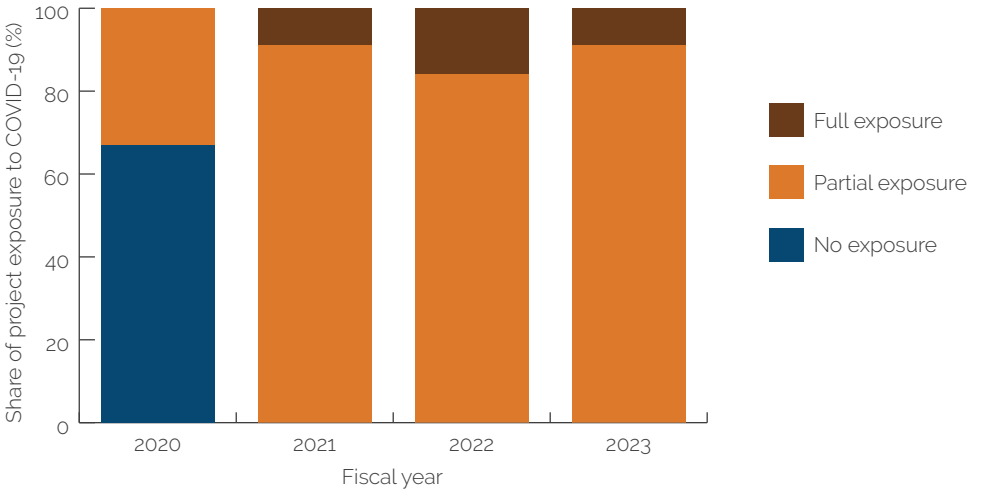


Source: Independent Evaluation Group.

Note: This chart includes all lending projects active at any stage of a fiscal year, including those that will not be rated by the Independent Evaluation Group, as they do not produce Implementation Completion and Results Reports, such as those projects below $5 million. As such, the chart provides an indication of the total efforts being undertaken by the World Bank in different contexts. Africa = Africa West and Africa East combined; FCS = fragile and conflict-affected situations.

Among operations closed since FY20, an increasing proportion of their life span occurred during the COVID-19 period (defined as March 1, 2020, through May 1, 2023). All operations closed since FY20 have spent part of their lifetime during COVID-19. Operations closed in FY23 spent an average of 49 percent of their life span during the COVID-19 period, the highest so far. While most operations have been partially affected by COVID-19, only a small share had their entire life span (full exposure) within the pandemic time frame (figure 2.5). The rising COVID-19 exposure aligns with the growing share of investment project financing (IPF) operations that have cited epidemics as a challenge in their Implementation Completion and Results Reports (ICRs), from 2 percent in FY20 to 86 percent in FY23. (A detailed discussion of the challenges identified in ICRs can be found in the Challenges section in this chapter.)

Figure 2.5. Average COVID-19 Exposure Among Projects Closed During FY20–23



Source: Independent Evaluation Group.

Note: Full exposure refers to a project approved and closed entirely within the COVID-19 period; partial exposure refers to a project operated partially during the period; no exposure refers to a project closed before the period. *N* = 248 for FY20, *N* = 267 for FY21, *N* = 277 for FY22, and *N* = 247 for FY23.

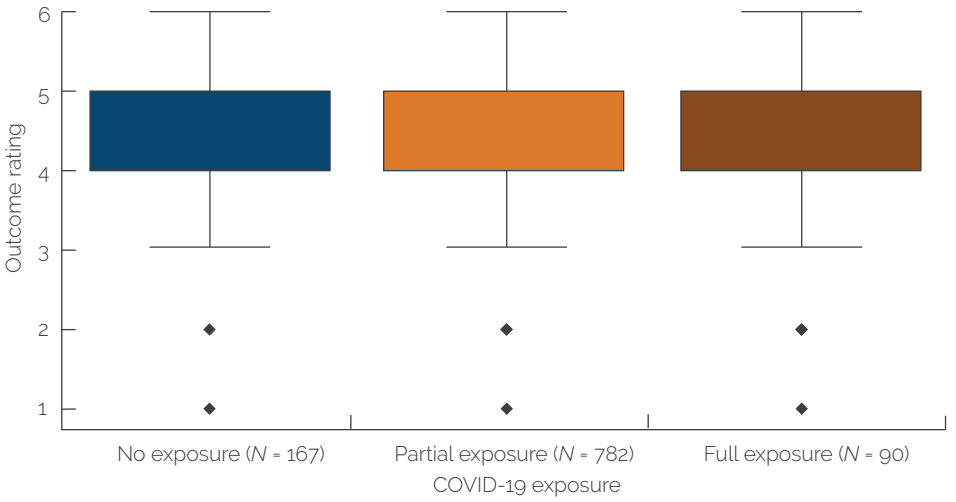
Outcome ratings of operations with different extents of their lifetime spent in COVID-19 are similar, despite expectations to the contrary. No statistically significant correlation exists between COVID-19 exposure and project outcome ratings. As shown in figure 2.6, operations with varying degrees of

lifetime exposure to COVID-19 demonstrate similar average outcome ratings. Although the result contradicts expectations of a decrease in ratings for operations exposed to COVID-19, recent IEG evaluations and *RAP 2023* provide some explanation for the stability of ratings. Unlike previous IEG studies on the global financial crisis, which found that operations already underway when the crisis hit had weaker ratings, recent IEG evaluations of the Bank Group's COVID-19 response reveal that the World Bank applied lessons from the past, leading to significant adjustments in operations and the application of new digital technologies. Furthermore, *RAP 2023* showed extensive restructuring whereby operations adapted by repurposing components, revising indicators, reallocating financing, and extending durations. Extensive restructuring was evident in approximately 60 percent of World Bank country programs as part of the COVID-19 response, which substantially realigned their portfolios to address the evolving needs arising from the pandemic (World Bank 2022h). These adjustments include significant modifications to existing projects, enhanced analytic work in relevant sectors, and the introduction of new support initiatives.

Bank performance ratings have had a consistent upward trend since FY13. The rating reached 90 percent of operations rated moderately satisfactory or above, and an average rating of 4.35 in FY23 (figure 2.7, panel a), which has increased incrementally in each year since FY16. Unlike outcome ratings, the World Bank has greater control over the quality of entry and supervision that underpin the Bank performance rating. The increase in average ratings and a lower percentage of operations being rated moderately unsatisfactory and below suggest that there have been ongoing steady improvements in issues related to operations' designs and support to implementation.

In FCS contexts, a gap has emerged between outcome and Bank performance ratings (figure 2.7, panel b). For much of the past decade, outcome ratings and Bank performance ratings in FCS have tracked each other closely.[2] However, since FY20, a notable divergence has appeared between these two ratings. This disparity increased in FY23, with average outcome ratings at 4.0 and Bank performance ratings at 4.2, a statistically significant difference. The widening gap suggests that improvements in Bank performance may not fully translate into better outcome ratings, especially in challenging environments.

Figure 2.6. Distribution of Outcome Ratings by COVID-19 Exposure



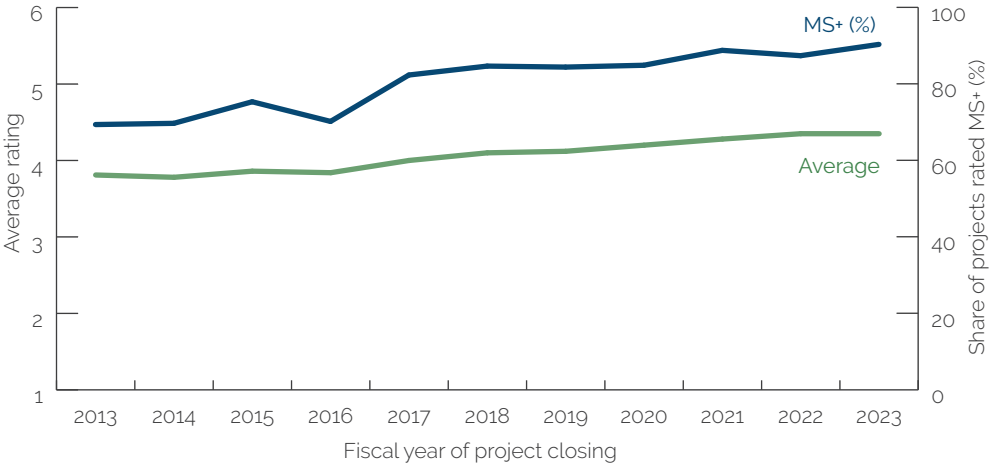
Source: Independent Evaluation Group.

Note: The average outcome ratings of operations with no exposure, partial exposure, and full exposure are 4.28, 4.27, and 4.27, respectively. No exposure = no part of operation lifetime spent in COVID-19 emergency period; partial exposure = some part of operation lifetime spent in COVID-19 exposure, but less than 100 percent; full exposure = entire lifetime of the operation is within COVID-19 emergency period. The COVID-19 emergency period is FY20–23.

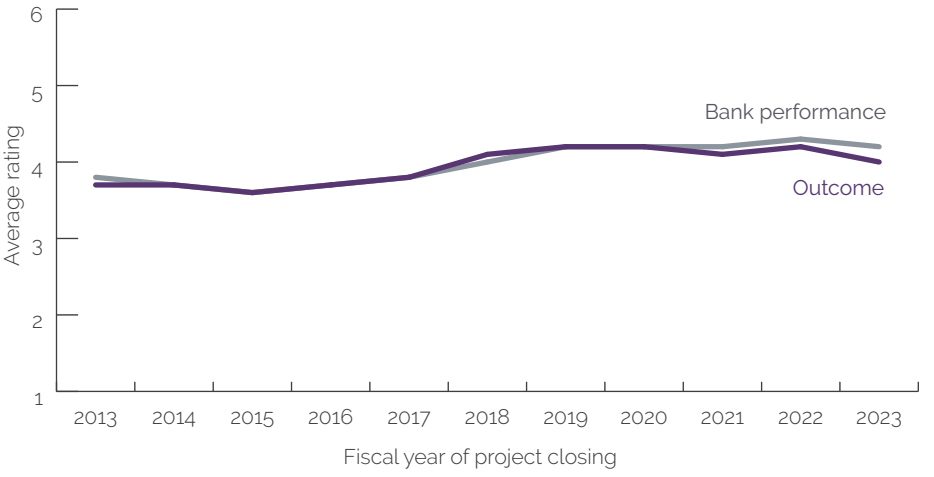
Despite improvement in monitoring and evaluation (M&E) quality ratings in FY20, a considerable share of operations is still rated modest or negligible. The average M&E quality rating of IPF and Program-for-Results increased rapidly from 2.1 in FY13 to 2.6 in FY20. Since FY20, the average rating has plateaued between 2.6 and 2.7. The percentage of IPF and Program-for-Results projects with M&E quality rated substantial or above increased from 29 percent in FY13 to 57 percent in FY20 and more incrementally to 64 percent in FY23. Contributing to this are some improving Global Practices. For example, the Water Global Practice's ratings have improved consistently each year since FY19, rising from 26 percent rated substantial or above to 85 percent rated substantial or above in FY23. Yet more than one-third of operations still have inadequate M&E practices. This gap is concerning, as M&E practices are largely under the World Bank's influence and are associated with overall outcome ratings. Certain Global Practices may require particular attention given the introduction of the new Scorecard and indicators; for example, during FY20–23, 64 percent of Transport operations and 57 percent of Governance operations had M&E quality ratings of modest or negligible.

Figure 2.7. World Bank Performance Ratings

a. Average rating and percentage rated MS+



b. FCS outcome rating versus Bank performance rating



Source: Independent Evaluation Group.

Note: FCS = projects operated in countries classified as fragile and conflict-affected situations during their lifetime; MS+ = moderately satisfactory or above.

Challenges

Challenges are persistent factors often negatively associated with project performances at both the country and project levels. The *RAP* analysis focuses on areas within the World Bank's sphere of influence, such as the institutional capacity of stakeholders, project design, data and monitoring, and finance, because these are frequently cited by ICRs and often linked to outcome ratings. Understanding these challenges is essential for the World Bank to make informed adjustments to operation design and implementation, thereby improving project performance. To analyze challenges, *RAP 2024* created a comprehensive data set by combining existing data from *RAP 2021* and *RAP 2023* with new classifications of factors from the Key Factors That Affected Implementation and Outcome section of ICRs. This data set covers all IPF projects closed during FY18–23 as of December 2023. The 12 factors analyzed in the *RAP* related to three clusters—context, institutional capacity of stakeholders, and project—with sentiment tagged to each factor (box 2.2).

Box 2.2. Delivery Challenges in Operations Taxonomy

Results and Performance of the World Bank Group 2024 builds on previous *Results and Performance of the World Bank Group* reports by continuing to use an adapted version of the DeCODE (Delivery Challenges in Operations for Development Effectiveness) taxonomy (table B2.2.1). Developed by the World Bank's Global Delivery Initiative in 2016, DeCODE identifies typical delivery challenges that could affect operational performance from design to closure. The taxonomy's validity is ensured through an iterative process involving literature reviews, text analytics, and practitioner consultations. In *Results and Performance of the World Bank Group 2024*, the analysis is structured around three main clusters: context, institutional capacity of stakeholders, and project. These clusters are further divided into 12 categories, some of which include subcategories. The World Bank has limited influence on context, indirect influence on institutional capacity of stakeholders, and direct influence on project-related factors. Detailed definitions of these clusters, factors, and subfactors can be found in appendix A.

(continued)

Box 2.2. Delivery Challenges in Operations Taxonomy (cont.)

Table B2.2.1. Delivery Challenges

Context	Institutional Capacity of Stakeholders	Project
» Legislation and regulations » Governance and politics » Political interference » Electoral cycles » Conflict and instability » Disasters and emergency response » Natural disasters » Epidemics » Business environment » Macroeconomic environment	» Coordination and engagement » Commitment and leadership » Human resources and organizational capacity	» Project design » Appropriate objectives or design » Time allocation or task sequencing » Stakeholder selection » Beneficiary targeting » Project finance » Procurement » Financing mechanism » Budgeting » Financial management and reporting » Project data and monitoring » Indicators » Data availability and baselines » Reporting and supervision

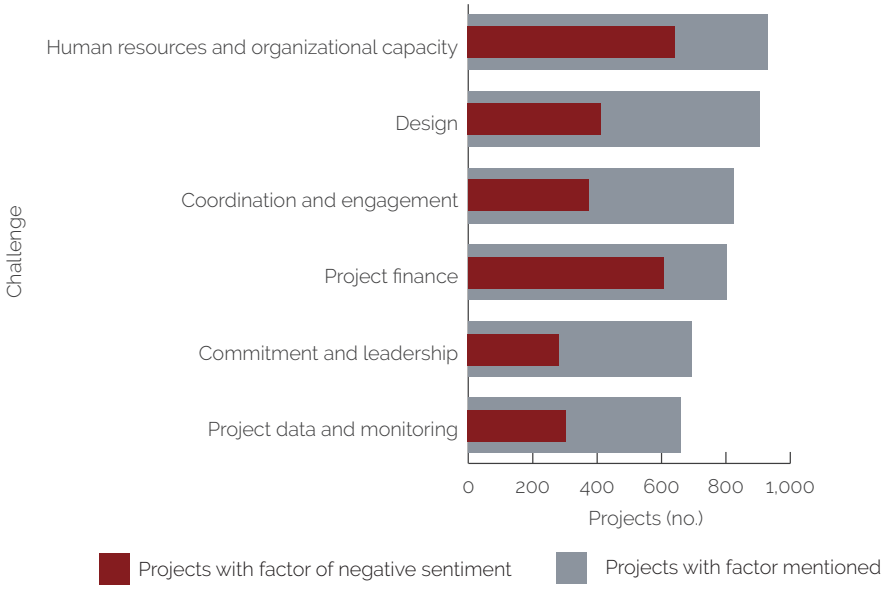
Sources: Independent Evaluation Group; World Bank 2023e.

Note: The original DeCODE taxonomy has 15 categories. Three were dropped (social and cultural, environmental and geography, and basic infrastructure) because too few examples of these were identified in developing the training data for the machine learning model. The cluster name *institutional capacity of stakeholders* is adapted from the original term *stakeholder*. It is also recognized that *stakeholder* is a term that may imply hierarchal narratives, which is not the intention in this box (Reed and Rudman 2023).

IEG found that challenges within the World Bank's sphere of influence could lead to lower project outcome ratings. The analysis revealed that all six factors within the institutional capacity and project clusters were addressed by nearly 60 percent or more projects (figure 2.8), and challenges in all the factors correlate with lower outcome ratings. Notably, these factors all fall within the World Bank's sphere of influence to varying degrees. This finding underscores the significant impact that operational difficulties can have on

project success and highlights areas where the World Bank can intervene to improve outcomes.

Figure 2.8. Challenges in Institutional Capacity of Stakeholders and Project Factors



Source: Independent Evaluation Group.

Note: The dashboard that supports further review of factors linked to performance is available (see also appendix B for more details).

Challenges with factors that link to the institutional capacity that affect outcome ratings—human resources and organizational capacity, coordination and engagement, and commitment and leadership—are within the World Bank's indirect influence.[3] In the World Bank, institutional capacity-building efforts aim to improve the effectiveness of country development by changing the formal and informal rules that structure interactions across multiple organizations (World Bank 2005b, 2018b; World Bank Group 2017).[4] The *RAP* analysis found that 75 percent of projects encountered one or more challenges in human resources and organizational capacity, coordination and engagement, and commitment and leadership factors. Human resources and organizational capacity challenges include a lack of qualified personnel or their difficulty in acquiring necessary skills. Coordination and engagement challenges stem from complex administrative

structures, ambiguous role definitions, or insufficient communication strategies. Commitment and leadership challenges arise from shifts in leadership, evolving priorities, or a lack of shared vision among stakeholders.

Institutional capacity plays a crucial role in the success of World Bank operations. Various World Bank and external studies have identified institutional capacity as a critical issue for improving development effectiveness (OECD 2008; Otoo et al. 2009; World Bank 2005a, 2005b, 2017a, 2018b, 2018c, 2022b; World Bank Group 2017). There is a nonlinear association between the extent of institutional capacity challenges and the outcome rating. Specifically, the negative effect of challenges on outcome ratings becomes more pronounced when there is more than one challenge, with a more noticeable downward shift in outcome ratings as the number of challenges increases from one to three. Moreover, similar findings on the importance of addressing institutional capacity challenges have been highlighted in previous IEG products—for example, IEG's evaluation of the World Bank's early support to addressing COVID-19 health and social response (World Bank 2022h).

Institutional capacity challenges are more prevalent in FCS contexts than in the overall portfolio. The share of projects reporting institutional challenges and the average number of challenges per project are both higher in FCS contexts than in the overall portfolio, underscoring the importance of attending to this subgroup. Overall, among projects that discussed institutional capacity, 78 percent of operations mentioned one or more challenges in FCS, compared with 75 percent overall. Among operations that addressed human resources and organizational capacity, challenges were cited in 75 percent of FCS, compared with 69 percent overall. Similarly, for projects that mentioned commitment and leadership, challenges were more prevalent in FCS contexts (48 percent) than overall (41 percent).

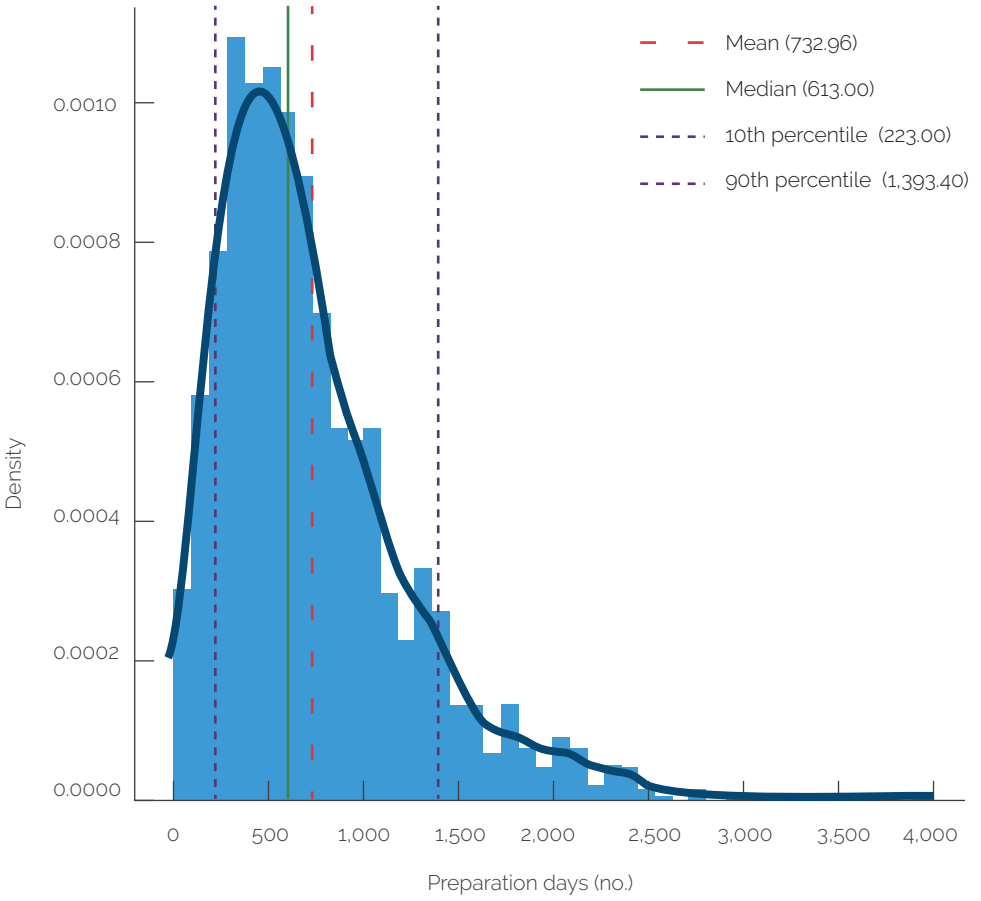
Beyond institutional challenges, operations in Africa and FCS also encounter more country-level contextual obstacles. In Africa, among the projects that discussed political interference and business environment challenges, 71 percent and 75 percent, respectively, reported facing these challenges, compared with 66 percent and 69 percent at the overall level. Out of the FCS projects that discussed political interference and electoral cycles, 74 percent and 80 percent identified these as challenges, compared with 66 percent and

73 percent, respectively, at the overall level. Additionally, in FCS contexts, 85 percent of the projects that addressed the macroeconomic environment reported it as a challenge, compared with 66 percent overall. Along with the ongoing portfolio shift toward Africa and FCS contexts, these findings underscore the need for mitigation efforts that focus on these specific challenges.

The World Bank has opportunities to reinforce performance by improving the design of operations. There are four aspects that define the project design category: appropriate objectives or project design, time allocation or task sequencing, stakeholder selection, and beneficiary targeting. Among projects that discussed challenges in project design, 45 percent reported one or more challenges in these four areas. The most frequently discussed subfactors were appropriate objectives or design and time allocation or task sequencing, which were addressed by 64 percent and 48 percent of projects, respectively. Challenges were identified in 25 percent of the operations addressing objectives or design and in 51 percent of those addressing time allocation or sequencing. The importance of project design has been a subject of concern in various World Bank reports, including those by IEG. Tracing back to the Wapenhans report in 1992, the *RAP* findings echo concerns that have been raised previously, such as recognizing risks from limitations in country capabilities, giving systematic attention to whether the operation can be implemented, keeping complexity minimal, and considering intended outcomes and risks when designing indicators (World Bank Group 1992).

Operations with the longest preparation times are more likely to encounter significant challenges in institutional capacity and project design. The distribution of IPF preparation times shows a notable proportion taking extended periods. As shown in figure 2.9, for IPF projects closed during FY13–23, the 10 percent with the longest preparation times took more than 1,393 days for preparation (measured from initiation to approval). The analysis reveals a significant correlation between projects with extremely long preparation times and the occurrence of challenges related to institutional capacity and project design. These extended preparation periods not only reflect difficulties encountered during the design stage but also serve as an early warning of potential obstacles in institutional capacity that operations may face during implementation.

Figure 2.9. Density Distribution of Project Preparation Days Between Initiation and Approval



Source: Independent Evaluation Group.

Note: This graph displays the density distribution of preparation time of 2,213 investment project financing operations closed during FY13–23. The histogram represents the frequency of projects across different preparation time intervals. The density scale normalizes the data, ensuring that the area under the curve sums to one. Overlaid on the histogram is a kernel density estimate curve, which provides a smooth, continuous representation of the probability density function.

Extremely long preparation periods also correlate negatively with the performance of operations. Operations with preparation times exceeding the 90th percentile tend to underperform compared with those below this threshold, consistent with earlier findings on the association of challenges with ratings. Table 2.1 shows strong negative correlations with outcome, Bank performance, and M&E quality. This negative correlation is particularly pronounced for Bank performance, closely followed by M&E quality. This

finding aligns with *RAP 2015*, which also identified a negative correlation between the months from the Concept Note to project approval and project outcome ratings. Conversely, even though the *RAP 2024* analysis does not show a statistical association between shorter preparation time and project outcomes, other evaluation evidence highlights the importance of adequate preparation for successful outcomes (World Bank 2024f).

Table 2.1. Mann–Whitney *U* Test and Ordinal Logistic Regression on Project Ratings and Long Preparation Time (more than 1,393 days)

	Mann–Whitney *U* Test	Ordinal Logistic Regression Models (independent variable: long preparation time)	
Rating	*p* value	Coefficient	*p* value
Outcome	.000	−0.422491	.0012
Bank performance	.000	−0.638323	.0000
M&E quality	.000	−0.604527	.0000

Source: Independent Evaluation Group.

Note: M&E = monitoring and evaluation.

Challenges in project data and monitoring are more directly within the World Bank's control, with evidence on challenges highlighting opportunities where improvements can be made. Challenges identified in project data and monitoring span indicators, data availability and baselines, and reporting and supervision. Notably, 46 percent reported at least one of these challenges among operations that discussed issues in project data and monitoring. Furthermore, there is a negative correlation between project outcome ratings and challenges in indicators, and reporting and supervision projects with these specific weaknesses tend to have lower outcome ratings. Indicator challenges included poorly designed or misaligned metrics that failed to capture intended outcomes, a lack of clarity on calculation or measurement methods, and overly ambitious targets that were unrealistic for given timelines or country contexts. Data availability and baseline problems encompassed a lack of initial data, difficulty setting appropriate targets due to missing baselines, and challenges in tracking progress without reliable information. Reporting and supervision difficulties involved data collection

and reporting delays, low quality of progress reporting, and obstacles to conducting in-person monitoring visits because of security concerns, travel restrictions, or other external factors.

Financial management challenges are prominent, including those related to procurement, budgeting, financial management and reporting, and financing mechanisms. Overall, 76 percent of operations encountered one or more of these challenges among those mentioned that reported these issues. Procurement was the most frequently cited factor (41 percent of projects, among which 76 percent reported it as a negative factor specifically).

Common challenges in project finance were identifiable across operations. Common procurement challenges included delays due to counterpart fund disbursement, prolonged approval processes, and procurement rules. Budgeting problems involved inadequate or untimely counterpart funding, financing gaps from underestimated costs, and inflexible budget processes. Financial management and reporting challenges encompassed disbursement delays due to various administrative bottlenecks, weak internal controls leading to financial mismanagement, and recurring delays and low quality of financial reports. Financing mechanism challenges included, for example, IPFs' lack of flexibility and complexities in disbursement-linked indicator modality and multidonor trust funds. There is a significant negative correlation between Bank performance and outcome ratings and challenges in budgeting, financial management and reporting, and financing mechanisms. The finding aligns with the IEG evaluation of the World Bank's procurement system, which pointed out that even minor improvements in procurement can substantially enhance project outcomes (World Bank 2024f).

Levers

Risk identification and mitigation and adaptive management are two critical levers that can significantly impact project outcomes. The World Bank's ability to influence project performance is linked to its capacity to activate key "levers"—actions within the organization's control to address challenges or enhance performance. By effectively employing these levers, the World Bank may proactively identify potential obstacles, develop targeted mitigation strategies, and adapt to changing circumstances throughout a project's life

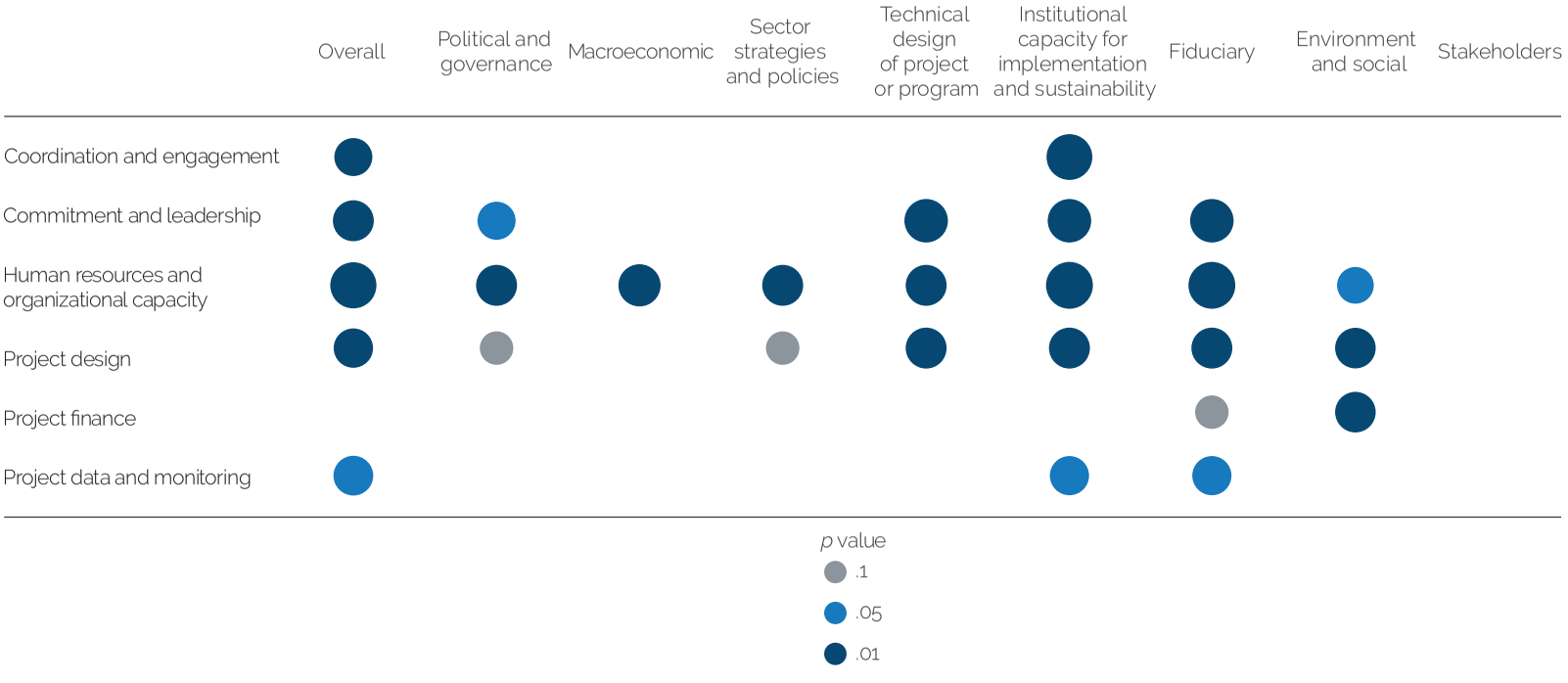
cycle. The following analysis explores how the World Bank uses these levers to address the challenges identified in this chapter and contribute to more favorable outcome ratings.

Risk identification and mitigation can enable better anticipation and response to challenges. The key challenges discussed earlier in this chapter closely align with risks defined by the World Bank's SORT. SORT evaluates residual risk by assessing the likelihood and impact of risks materializing, considering mitigation measures (World Bank 2021b). It identifies specific inherent risks to development outcomes and reviews mitigation strategies. The analysis found statistically significant correlations between the six challenges within the World Bank's sphere of influence and SORT ratings, including overall ratings and ratings of various SORT categories (figure 2.10). Institutional capacity for implementation and sustainability, technical design of project or program, and fiduciary are the three SORT categories that correlate the most with the six challenges.

Risk identification and mitigation look to be important for undertaking more successful operations. The *RAP* analysis revealed statistically significant negative correlations between outcome ratings and end SORT ratings across all categories. Larger risk reductions during operations, measured as the difference between the initial and end SORT rating, are also associated with higher outcome ratings. Operations that fail to mitigate high risk throughout their life cycle are strongly associated with lower outcome ratings. Notably, the most substantial effects of risk reductions were observed in areas where the World Bank has indirect or direct control, particularly in institutional capacity for implementation and the technical design of the operation. Further consideration of actions that support lowering risk ratings would be beneficial, as some shifts may occur as risks do not materialize, whereas others may arise from efforts on the part of the World Bank team and country counterparts.

The analysis of the text of ICRs reinforces the importance of proactively addressing risk. The *RAP* analyzed the sentiment of the text in the Key Factors That Affected Implementation and Outcome section of 1,118 ICRs. This analysis also uncovered a strong association between positive sentiment of risk identification or mitigation actions and higher outcome ratings. This finding is crucial in Africa and FCS contexts, where risk profiles are higher (figure 2.11) and the share of lending projects has grown.

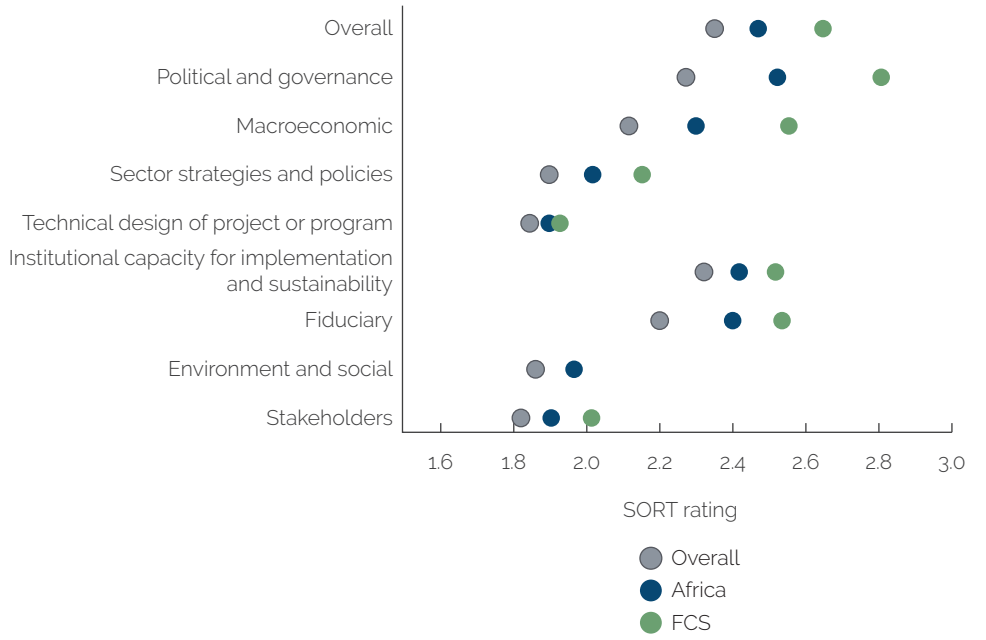
Figure 2.10. Correlations Between Project End Systematic Operations Risk-Rating Tool Rating and Challenges



Source: Independent Evaluation Group.

Note: Spearman correlation is used to calculate correlation coefficients and statistical significance. All correlations are positive. Sample sizes ranged from 652 to 925 projects per pair. Bubble size represents covariance (larger = higher). Bubble color indicates *p* value.

Figure 2.11. Comparison of Average Systematic Operations Risk-Rating Tool Ratings



Source: World Bank operations data.

Note: SORT rating is on four-point scale, with larger number indicating higher risk. Data include 1,613 investment project financing lending projects closed during FY16–23. FCS = fragile and conflict-affected situations; SORT = Systematic Operations Risk-Rating Tool.

The World Bank can effectively identify and mitigate risks in institutional capacity through various strategies. SORT guidance highlights key risk areas in institutional capacity, such as the competence of implementing agencies, implementation arrangements, and M&E systems. The combined manual and artificial intelligence–assisted summarization review of text from 615 ICRs with positive sentiment of risk identification or mitigation actions shows that successful operations often included extensive efforts to influence institutional capacity through, for example, capacity-building initiatives, such as training agency staff on World Bank processes and technical aspects of project management, and establishing well-resourced, dedicated implementing agency units. To manage the complexity of institutional arrangements, successful operations streamlined steering committees by maintaining representation from key stakeholders and oversight functions, along with detailed project operational manuals that clarify roles and responsibilities. For projects involving multiple agencies or levels of

government, formal coordination mechanisms, such as regular interagency meetings and shared reporting systems, ensured consistent communication and alignment. Additionally, enhanced engagement and monitoring—through the involvement of technical experts, regular missions, and community participation—are critical for addressing risks related to capacity and resource challenges in institutional capacity.

The technical aspect of project design, largely controlled by operational staff, is another critical area for risk reduction and mitigation. SORT considers the project design stage to be the primary mitigation measure. The aforementioned text analysis found that operations with realistic and measurable objectives engaged interests that could exert more direct control over implementation. These operations conducted extensive political economy analysis and drew lessons from the World Bank's previous operations and economic and sector work. The design of operations also benefited from engaging key internal and external stakeholders in comprehensive consultation processes involving government entities, local communities, civil society organizations, and international partners. The effective design phase incorporated local stakeholders, which enabled communities to identify needs and plan interventions. Successful approaches also included decentralizing decision-making processes and engaging local institutions to leverage existing structures and knowledge. By grounding designs in practical insights and local needs, operations promoted local ownership and aligned objectives with the on-the-ground realities of their specific locations.

Appropriate design of operations is paramount in FCS contexts, where complex challenges demand tailored, realistic, and adaptable approaches. *IDA21 Policy Package: The "Lenses" Paper* emphasized the need for tailored solutions that account for limited government capacity and resources in FCS contexts (World Bank 2024e). *World Bank Group Strategy for Fragility, Conflict, and Violence 2020–2025* and its Mid-Term Review noted that crisis-related operations could perform well because of their focused, simple, and realistic nature, calling for more realism in objective setting and project design (World Bank 2020e, 2023d). Moreover, IEG evaluations have indicated the importance of World Bank efforts in, for example, the Geo-Enabling Initiative for Monitoring and Supervision in reinforcing elements of performance that are relevant for the design of operations (World Bank 2021d).

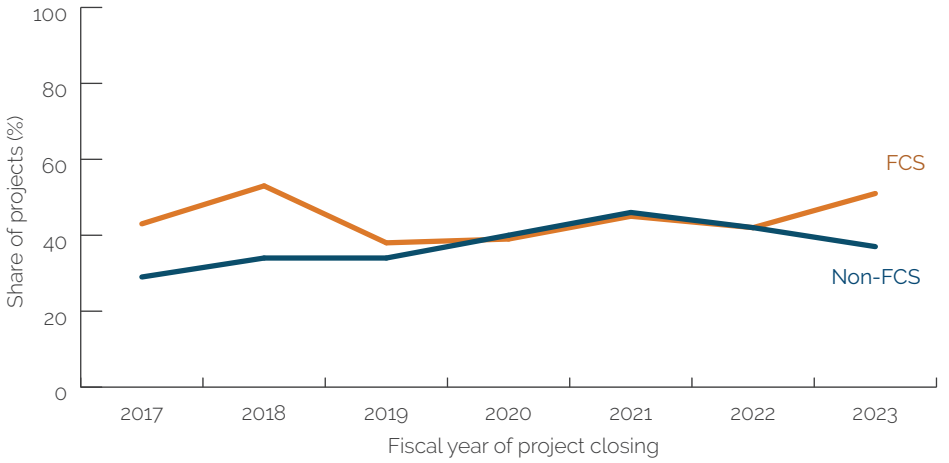
The analysis of ICR text revealed successful operations that leveraged existing local networks and delivery systems to maintain essential services despite ongoing conflict. Operations emphasizing quick wins—achievable, high-impact activities—helped maintain momentum and demonstrated immediate benefits to affected populations.

The *RAP* analysis also observed a growing share of projects in FCS with objectives incorporating expanded access to services since FY20, with these projects receiving higher outcome ratings on average (figure 2.12). Improving access to services is important, as it can support the building of trust in institutions (World Bank 2020d). The percentage of operations incorporating access to services in their objectives has grown from 39 percent in FY20 to 51 percent in FY23. In contrast, operations not exposed to FCS more frequently incorporate elements of quality of service in objectives. For both FCS and non-FCS operations, enhancing the capacity of institutions remains the most frequent area of focus. The FCS operations that incorporated expanding access received significantly higher average ratings. The FCS operations whose focus did not include access—but rather incorporated only improving quality of services or institutional capacity—received lower ratings. This finding suggests that consideration needs to be given to framing project objectives that balance the need for expanding access to services while also preserving institutional strengths that are needed for long-term development, which other analyses in this report point to as important.

Mitigating fiduciary risks—encompassing both financial management and procurement—is crucial for improving project outcomes. Correlations between outcome ratings and fiduciary risk at project completion, as well as shifts in fiduciary risk ratings, exist across the portfolio, with particular importance in FCS and African contexts, where fiduciary risks are higher than average. IEG's evaluation of procurement highlights its critical role in achieving development outcomes, noting it as a major challenge in project performance, especially in fragile, conflict-affected, and low-capacity countries (World Bank 2024f). The analysis of ICR text revealed that successful projects mitigated financial management risks through strengthening institutional capacity, enhancing financial controls, and adopting robust systems. Actions included training financial management staff, establishing dedicated financial management units, implementing regular external and internal

audits, and adopting computerized management systems. For procurement, effective risk mitigation strategies focused on capacity building, strengthening oversight, and implementing adaptive approaches. Key actions included extensive training programs, developing detailed procurement manuals, establishing multilayered supervision mechanisms, and adopting flexible procurement strategies. Both financial management and procurement areas emphasized hiring specialized staff, creating independent oversight bodies, and providing hands-on support to address delays and improve execution. Similar issues have been highlighted in the recent IEG evaluation on procurement (World Bank 2024f).

Figure 2.12. Share of Projects with an Objective on Expanding Access to Services



Source: Independent Evaluation Group.

Note: FCS = fragile and conflict-affected situations.

Adaptive management can be undertaken to support risk management and enhance project outcomes. Adaptive management is an iterative approach to decision-making, whereby interventions and portfolios are adjusted based on evidence and evolving context (World Bank 2020d). The World Bank recognizes the importance of adapting to local challenges and allowing operational flexibility (World Bank 2020d, 2022c, 2023e). The *RAP* analysis coded text discussing adaptive management in the Key Factors That Affected Implementation and Outcome section of ICRs. The analysis found a statistically significant correlation between positive sentiment in the text on

adaptive management and higher outcome ratings. Projects can implement adaptive management to mitigate risk or to take advantage of new opportunities (for example, through restructuring). Risk-responsive adjustments may involve changes to scope, timelines, activities, results frameworks, budgets, and procurement. ICRs provide examples such as shifting to virtual training during the COVID-19 pandemic, relying more on local partners in deteriorating security situations, and revising procurement strategies to attract qualified contractors. Adaptive management can also occur outside of risk management. Examples include scaling up projects based on lessons learned or implementing new activities through cost savings and exchange rate gains. Overall, adaptive management demonstrates a project's ability to improve outcomes by addressing challenges and capitalizing on enhancement opportunities.

[1] FCS exposure means the country of operation was classified as FCS for at least one year during the operation's lifetime. It is a more accurate capture of the context of the country in which the operation was implemented. Previous *RAPs* tagged only FCS as the country status of the operation's closing fiscal year. FCS exposure represents a change from the previous methodology because operations in countries that experienced FCS conditions during implementation but were non-FCS at closing are now properly captured in the portfolio analysis. The full performance implications of this vary as some projects and Global Practices better design for adaptation to conflict situations (World Bank 2021f). The change in method leads to an additional 187 operations being captured under the FCS analysis. The change to FCS exposure method produces a small increase in average ratings of 0.02 during FY13–23 and 0.06 in FY23.

[2] It is worth noting that Afghanistan accounts for the largest share of operations among FCS countries during FY13–23 at 9 percent, with this share reaching a peak of 15 percent in FY23.

[3] Other analyses of World Bank operations have identified similar institutional challenges in ICRs and Implementation Status and Results Reports (for example, World Bank 2022h, 2023a; World Bank Group 2017).

[4] Institutional capacity-building efforts are a feature of the Bank Group's evolution process, with papers on the Knowledge Compact and Global Challenge Programs identifying the importance of addressing institutional capacity gaps.

3 | International Finance Corporation

Highlights

In the medium term, the development outcomes of International Finance Corporation (IFC) investment projects improved. In the long term, however, they declined slightly. Investment project ratings in Latin America and the Caribbean and in challenging contexts also declined over the long term. Notably, the development outcomes of projects in countries classified as fragile and conflict-affected situations and International Development Association and blend countries declined from 50 percent rated mostly successful or better to 18 percent mostly successful or better, whereas the share of projects grew from 17 percent to 30 percent.

IFC's front-end work quality is a strong determinant of development outcomes in investment projects, whether it is through client quality, market analysis or assumptions, financial models, or project costs.

In the medium term, the development effectiveness of IFC advisory services projects improved. In the long term, however, it declined significantly. IFC advisory services have been delivering fewer projects and more low-effectiveness projects in International Development Association and blend countries over the long term. South Asia contributed more than any other region to the decline in IFC advisory services.

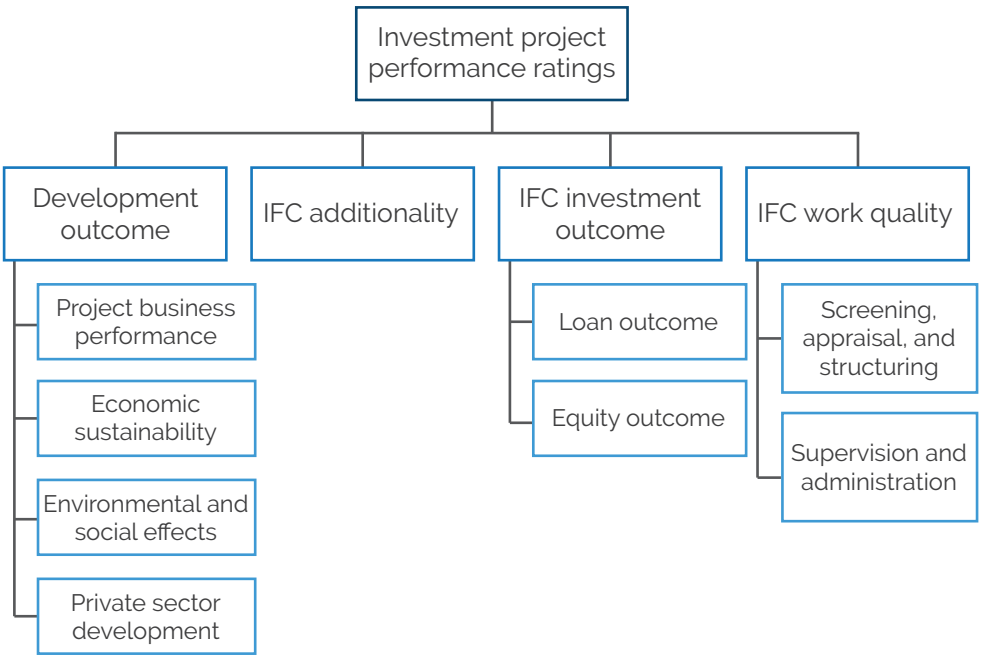
Low work quality contributes to weak development effectiveness. In FY21–23, 59 percent of IFC advisory services projects had a work quality rating of satisfactory or better. Only 9 percent of IFC advisory services projects had low work quality and high development effectiveness. In calendar years 2021–23, 55 percent of IFC investment projects were rated satisfactory or better on work quality.

COVID-19 has affected the development effectiveness of IFC advisory services projects and some IFC investment projects in the Financial Institutions Group. COVID-19 was the most important factor in IFC advisory services projects where work quality was strong and development effectiveness was weak. It also affected 19 percent of IFC investment projects but was prevalent only in the Financial Institutions Group.

Investment Projects

The *RAP's* main source of evidence for IFC investment projects is a random sample that IEG evaluates and validates every year. IEG draws a random stratified representative sample (40 percent) annually from among IFC investment projects that were approved by the Board of Executive Directors five years earlier and that reached early operating maturity (so that sufficient information is available for the evaluation). During the calendar year, IFC investment staff self-evaluate all active IFC investment projects in the sample, and IEG independently validates them. IEG evaluates closed projects in the sample in lieu of IFC self-evaluations. In this section, when we refer to projects "evaluated and validated by IEG," we mean this sample. "IFC-wide" means across IFC investment projects as a whole.

Figure 3.1. Performance Ratings in International Finance Corporation Investment Projects



Source: Independent Evaluation Group.

Note: IFC = International Finance Corporation.

IFC investment projects' performance is assessed on four dimensions: development outcome, IFC additionality, IFC investment outcome, and IFC work quality. Figure 3.1 shows these dimensions and their respective indicators. The development outcome dimension is particularly important in this section. It synthesizes a project's performance across four indicators: project business performance, economic sustainability, environmental and social effects, and private sector development. It is rated on a six-point scale: highly successful, successful, mostly successful, mostly unsuccessful, unsuccessful, and highly unsuccessful.

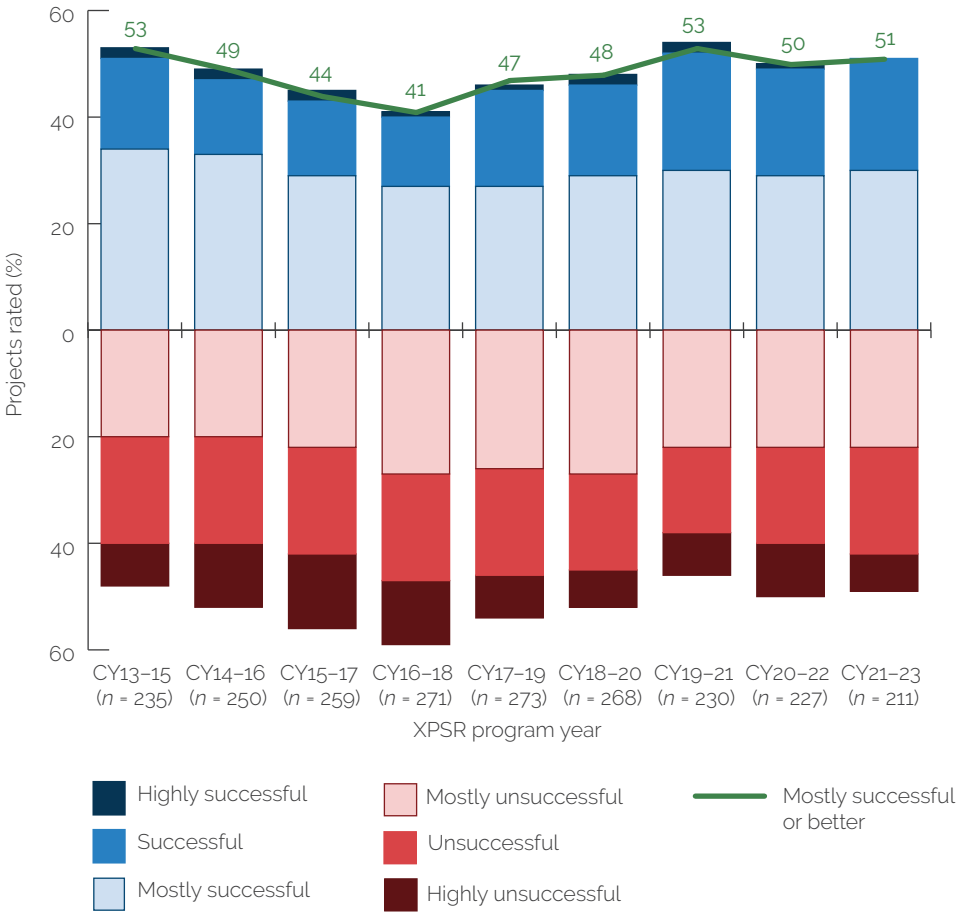
Trends

The development outcomes of IFC investment projects improved over the medium term but slightly declined over the long term. This subsection briefly touches on the IFC-wide improvements in development outcomes over the medium term but focuses on the long-term trend because (i) the organization can learn more from what did not work than from what worked, (ii) long-term trends encompass organizational changes and strategies that take effect over time, (iii) long-term trends capture a greater number of unique observations than shorter-term trends, and (iv) most of the lessons from the past have been reoccurring and remain crucial for IFC. We show that the development outcomes of some subgroups (Europe and South Asia) improved over the long term. However, ratings in Latin America and the Caribbean and in challenging contexts—IDA and blend countries, particularly FCS countries—declined over the long term. In addition, Africa's share of the active portfolio has increased more than any other subgroup. This shift could negatively affect IFC's overall development outcomes in the future, given Africa's declining development outcomes over the long term.

The development outcomes of IFC investment projects improved by 10 percentage points over the medium term but declined by 2 percentage points over the long term. The development outcomes of IFC investment projects improved over the medium term from 41 percent rated mostly successful or better (calendar year [CY]16–18) to 51 percent (CY21–23; figure 3.2).[1] The share of successful projects increased by 8 percentage points over the medium term. However, the development outcomes declined over the long term from 53 percent rated mostly successful or better (CY13–15) to 51 percent

(CY21–23). The increase of 4 percentage points in the share of successful projects over the long term was offset by a decline of 6 percentage points in the share of highly successful and mostly successful projects. In the remainder of this section, we analyze the development outcomes in five subgroups, two of which—Europe and South Asia—contributed positively to overall IFC development outcomes. The other three subgroups—Latin America and the Caribbean, IDA and blend, and IDA and blend projects in FCS—contributed negatively to overall IFC development outcomes.

Figure 3.2. Trends in Development Outcomes for International Finance Corporation Investment Projects, Calendar Years 2013–23

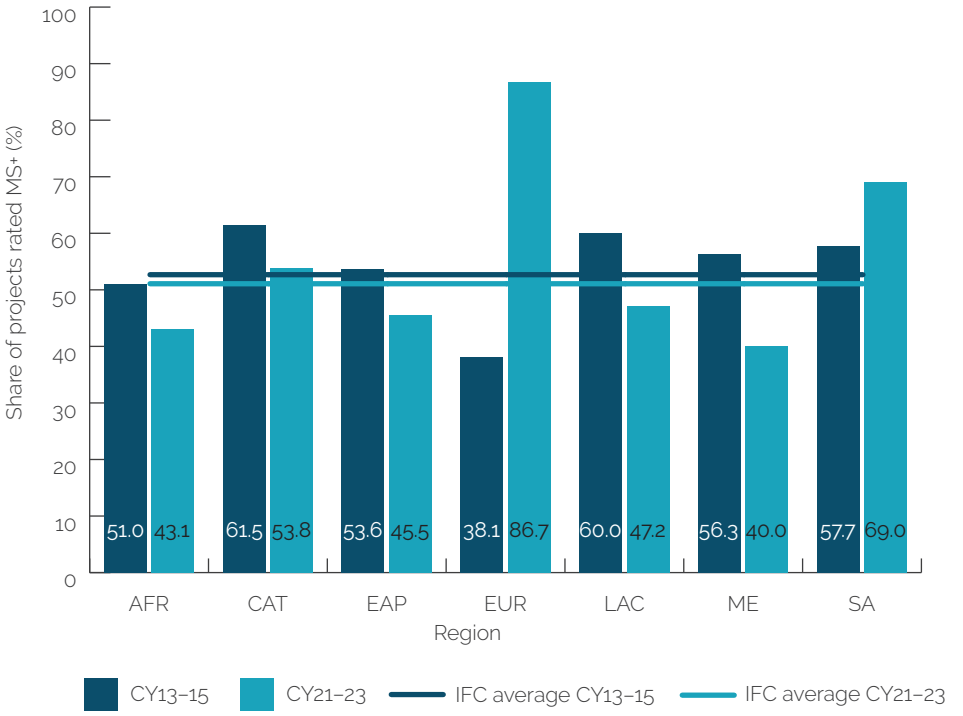


Source: Independent Evaluation Group, Expanded Project Supervision Report database.

Note: Trend line shows mostly successful or better. CY = calendar year; XPSR = Expanded Project Supervision Report.

The development outcomes of Europe and South Asia improved over the long term; moreover, South Asia was the biggest positive contributor to development outcomes among the regions. The development outcomes of Europe improved over the long term from 38 percent of projects rated mostly successful or better in CY13–15 to 87 percent in CY21–23 (figure 3.3). However, its share of projects evaluated and validated by IEG declined from 18 percent to 7 percent over the long term (figure 3.4). Among the regions, South Asia was the biggest positive contributor to development outcomes over the long term due to increases in both ratings (from 58 percent to 69 percent) and share of projects (from 11 percent to 14 percent) evaluated and validated by IEG.[2]

Figure 3.3. Development Outcomes by Region, Calendar Years 2013–15 Versus 2021–23



Source: Independent Evaluation Group, Expanded Project Supervision Report database.

Note: AFR = Africa; CAT = Central Asia and Türkiye; CY = calendar year; EAP = East Asia and the Pacific; EUR = Europe; IFC = International Finance Corporation; LAC = Latin America and the Caribbean; ME = Middle East; MS+ = mostly successful or better; SA = South Asia.

Figure 3.4. Share of Projects Evaluated and Validated by the Independent Evaluation Group by Region, Calendar Years 2013–15 Versus 2021–23

Source: Independent Evaluation Group, Expanded Project Supervision Report database.

Note: AFR = Africa; CAT = Central Asia and Türkiye; EAP = East Asia and the Pacific; EUR = Europe; LAC = Latin America and the Caribbean; ME = Middle East; SA = South Asia; WLD = World (multiregional).

Projects in Latin America and the Caribbean contributed to the long-term decline in IFC-wide development outcomes. Latin America and the Caribbean is one of the largest regions: it represents 25 percent of projects evaluated and validated by IEG in CY13–23. However, the development outcomes of the region declined from 60 percent of projects rated mostly successful or better to 47 percent over the long term. As a result of its large share of IFC-wide projects and the decline in its ratings, the region contributed more than any other to the decline in IFC-wide development outcomes (figures 3.3 and 3.4). This shift was mainly because IFC work quality in Latin America and the Caribbean, particularly preparation or front-end work (for example, screening, appraisal, and structuring), was the weakest among the regions over the long term. For example, according to self-evaluations and IEG validations, in some cases IFC missed or underestimated key risks such as client management quality and macro and market risk. Even if relevant risks were identified, adequate mitigation measures were not

designed. Continued contraction of major economies in Latin America and the Caribbean due to currency depreciation and increasing inflation, among other factors, also contributed to the decline in development outcomes in the region.

Projects in IDA and blend countries contributed to the long-term decline in IFC-wide development outcomes. The development outcomes of IDA and blend projects declined from 54 percent of projects rated mostly successful or better to 46 percent over the long term. The share of IDA and blend projects in IFC evaluated and validated by IEG also declined between CY13–15 (35 percent) and CY21–23 (27 percent), mainly because of India's graduation from IDA in FY14. Moreover, development outcomes of IDA and blend projects in recent years (CY21–23) are much weaker (46 percent rated mostly successful or better) than non-IDA projects (57 percent). The weak development outcomes in IDA and blend countries are especially striking because IDA is a strategic priority for IFC based on the IFC 3.0 strategy (2017) and the capital increase package (2018).

The long-term decline in development outcomes of IDA and blend projects is also partly attributable to a decline in the development outcomes of those in FCS contexts. The share of FCS projects in the overall IFC investment active portfolio has remained stable (at 11 percent) between FY21 and FY23 (as of April 2024); however, the number of projects in FCS has been increasing since FY21. The share of FCS in IDA and blend projects[3] in IFC evaluated and validated by IEG increased between CY13–15 (17 percent) and CY21–23 (30 percent). The development outcomes of FCS projects in IDA and blend countries declined significantly, from 50 percent rated mostly successful or better to 18 percent over the long term. In fact, sensitivity analysis showed that if the development outcome ratings of FCS projects were not considered, then those of IDA and blend would have increased by 2 percentage points rather than declining by 9 percentage points. IFC work quality rating for FCS in IDA and blend projects increased from 57 percent rated satisfactory or better to 59 percent over the same period. Economic issues were the most frequent factor linked to the performance of FCS in IDA and blend projects. Other factors (for example, asset quality, civil unrest, business risk, and client quality) also contributed to the overall decline in development outcomes of FCS in IDA and blend projects.

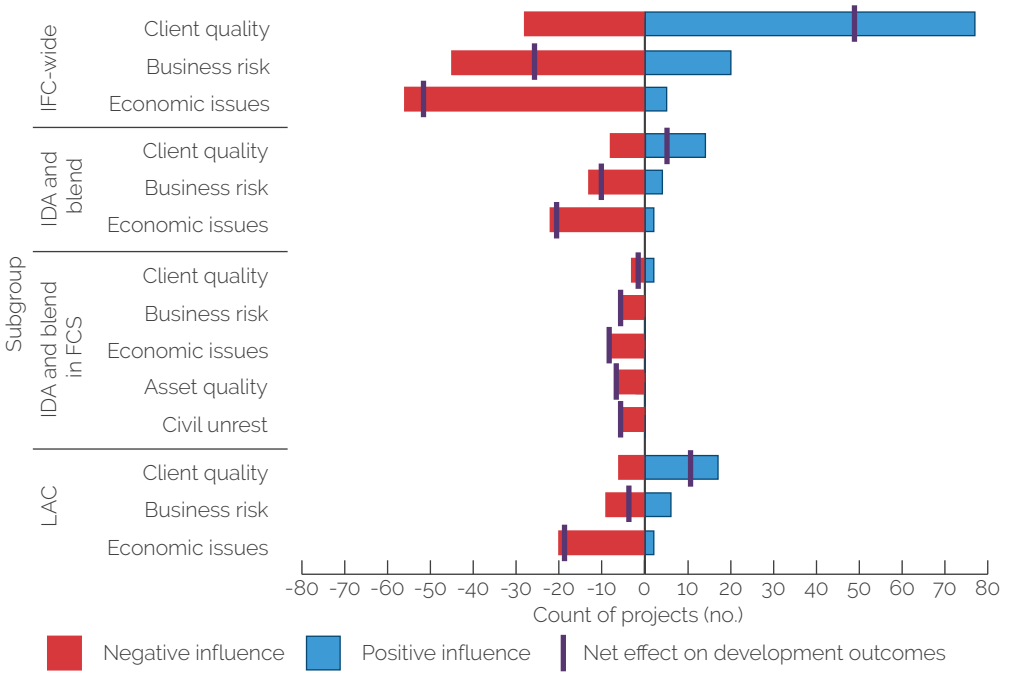
Looking forward, given its increasing share in the active portfolio, Africa's long-term development outcomes' decline may negatively affect IFC's overall development outcomes. An analysis of the active portfolio found that Africa is the only subgroup whose share increased by more than 5 percentage points—from 21 percent in FY21 to 29 percent in FY24 (as of May 2024). Its development outcomes improved from 35 percent rated mostly successful or better to 43 percent in the medium term, contributing positively to the IFC-wide improvement in development outcomes over that period. However, Africa's development outcomes declined from 51 percent rated mostly successful or better to 43 percent over the long term. Moreover, its recent development outcomes (43 percent rated mostly successful or better) are the weakest among the regions after the Middle East (40 percent). Given Africa's decline in ratings over the long term, the increase of its share of the active portfolio may contribute to a future decline in the overall IFC development outcomes. This shift is expected as projects have more complexities, particularly in the challenging context of IDA and blend countries in Africa.

Challenges

For IFC, challenges are prevalent factors that are usually negatively linked to development outcomes and are outside or only indirectly within the institution's influence. A factor can have either a positive or negative influence on development outcomes, but a challenge usually has a negative influence on development outcomes. Challenges may be linked to development outcomes either IFC-wide or for subgroups of interest. The subgroups of interest, which were discussed under the trend analysis, are Latin America and the Caribbean, IDA and blend, and IDA and blend projects in FCS.

The most prevalent challenges are business risk, asset quality, economic issues, and civil unrest. We conducted a deep dive of 256 IFC investment projects evaluated and validated by IEG (CY20–23), using the *RAP 2023* taxonomy to identify the top factors linked to development outcomes (figure 3.5). The most prevalent challenges are discussed in more detail in this section.

Figure 3.5. Most Prevalent Factors Linked to Development Outcomes, Throughout the International Finance Corporation and by Subgroup



Source: Independent Evaluation Group.

Note: The order of factors within subgroups follows the IFC-wide order. FCS = fragile and conflict-affected situations; IDA = International Development Association; IFC = International Finance Corporation; LAC = Latin America and the Caribbean; negative influence = the identified factor constrained the project performance; net effect = sum of positive influence and negative influence; positive influence = the identified factor aided the project performance.

Business risk is one of the most prevalent challenges. Business risk refers to risks related to a business model, cyclical business, or the operating environment. Business risk usually has a negative influence on IFC investment development outcomes, and it occurred in 25 percent of the projects reviewed. It appears across all the three subgroups. For example, a specialty fertilizer producer in Asia was expected to transform fertilizers. Its business model was built on a weak foundation because the client did not fully grasp the trends in fertilizer use in the country and based its strategy on overoptimistic assumptions. Consequently, the project activities were not fully completed given the limited market interest from local partners, mostly resulting from price competition with state-owned enterprises. In addition, because of both a substantial downscaling of the project and the adverse effects of COVID-19,

the financial and profitability results were substantially below the projections in the Board paper. IFC has indirect influence over business risk through client quality, as discussed in the Levers section in this chapter.

Asset quality is prevalent only in FCS and is mainly applicable to financial institutions. Low asset quality, such as a rise in the nonperforming loans of a client company, always has a negative influence on IFC investment development outcomes. This challenge occurred in 25 percent of IDA and blend projects in FCS that were reviewed. For example, an IFC client's nonperforming loans increased because of the deteriorating performance in a sector that was a key driver of growth for the country. Performance in this sector dropped because of the liquidation of one of the largest traders, which had systemic exposure to all top-tier banks in the country. This led to contagion across the banking sector. Ultimately, it significantly increased the nonperforming loans of the IFC client. As a result, the client did not achieve its target of increasing its small and medium enterprise loan portfolio.

Economic issues are the third most prevalent challenge IFC-wide. Economic issues are defined as risks related to the macroeconomic environment, inflation, monetary policy, or austerity measures. This factor almost always has a negative influence on IFC investment development outcomes. Economic issues occurred in 24 percent of projects reviewed. For example, an IFC project in Africa intended to allow the client to expand its healthcare services. However, the country's economy was heavily dependent on oil, and there was a prolonged drop in oil prices. A significant proportion of patients were government employees, and a contraction in public spending on health directly affected the hospital's results through reduced revenues and financial margins. IFC has no influence over economic issues because they are exogenous; however, IFC can improve its projections (for example, production, sales, and revenues) and scenario analysis by considering macroeconomic risks in combination with project risks during its front-end work. (Considering macroeconomic factors during IFC's front-end work can also help mitigate business risk, as discussed in the Levers section.)

Civil unrest was a prevalent challenge only in IDA and blend projects in FCS. This factor almost always has a negative influence on IFC investment development outcomes. Civil unrest occurred in 21 percent of IDA and blend

projects in FCS. For example, an IFC client was affected by the challenging macroeconomic situation created by the double impact of civil unrest and COVID-19. Consequently, the local currency depreciated by about one-third, with inflation soaring. As a result of the macroeconomic downturn, the outstanding loans declined significantly. IFC has no influence over civil unrest.

Levers

Levers are factors that are within IFC's influence and are positively linked to development outcomes IFC-wide or within subgroups of interest. This section focuses on how IFC can use levers to influence development outcomes. In particular, we argue that IFC can influence client quality, IFC work quality, IFC additionality, and outcome indicator tracking to improve development outcomes.

Client Quality

Client quality is a powerful lever. In this subsection, we show that selecting clients with proven business models, good financial standing, strong risk management frameworks, and flexible business strategies can mitigate business risk and help clients, particularly financial institutions, withstand epidemics like COVID-19. However, selecting high-quality clients may not be feasible in all contexts. Therefore, in challenging contexts where selecting high-quality clients may not be feasible, IFC can influence client quality by providing support for capacity building. In sum, client selection and complementary capacity building can help influence development outcomes.

IFC can influence development outcomes by selecting high-quality clients. Client quality includes the ability, technical expertise, or track record of the IFC client or its sponsor. It refers to the quality of the management team and their skills, contractor competency, familiarity, and acumen. This factor occurred in 41 percent of projects reviewed. It usually has a positive influence on IFC investment development outcomes, except for Africa and FCS projects. This is an important factor because client quality gives IFC indirect influence over challenges (such as business risk) that would otherwise be outside of its influence. IFC can influence this factor by selecting clients with firm commitments, sizable companies, proven business models,

strong management skills, and experience in the sector. For example, for an education project, IFC selected one of the largest private distance learning providers in Latin America and the Caribbean. The client's size and business model stand out in this example. The client adapted to the market by canceling the unused portion of its loans during a substantial downturn, then asked for financial support once the market improved. The affordable quality education provided by the client increased substantially. This improvement shows that IFC's selection of a high-quality client contributed to development outcomes, despite an external shock.

In challenging contexts where it may not be feasible for IFC to select high-quality clients, IFC can influence client quality by providing support for capacity building. In challenging contexts, such as Africa and IDA and blend projects in FCS, there may not be a robust pipeline of bankable projects and experienced clients with the capacity to successfully implement the projects. Nevertheless, IFC should be aware of client limitations and can build a client's capacity by providing nonfinancial additionalities, such as technical assistance through advisory services during supervision. For example, an IFC investment project in an FCS country benefited from client-facing advisory services before and during the investment period. The advisory services helped build staff capacity, develop new microfinance products, and train the client on responsible finance. The project introduced practices, systems, policies, and institutional capacity building intended to help the client achieve commercial and environmental sustainability. The institution building, including improvements to risk management and corporate governance, allowed the client to grow its microlending portfolio sustainably.

IFC can influence business risk indirectly by selecting clients with proven business models. IFC has an indirect influence over business risk through its role in selecting clients with high client quality. For example, an agribusiness project in an FCS country intended to build storage facilities and provide preharvest financing for purchase of fertilizers, among other inputs. The prefinancing was based on a model that had been successful in a non-FCS country. However, in this case, the farmers did not pay back preharvest credits, and the business model was untested locally and unsustainable. Ultimately, the project achieved none of its development targets. This example illustrates the point that it is important for IFC to select clients with

proven business models (which are tested locally and can adapt to local conditions when replicated from another country) to reduce business risk.

In addition to selecting clients with proven business models, IFC can also influence business risk by paying attention to combinations of market developments and economic factors. When doing risk assessment, IFC cannot narrowly focus on the client; it must also focus on combinations of broader factors related to market developments in the sector and the country's macroeconomic challenges. For example, IFC invested in the expansion of an animal feed and farming company in East Asia. IFC was aware of a potential macroeconomic factor (spreading animal disease) and a market factor (cost competition among the client's peers) that could have contributed to business risk. However, the client's innovative vertical farming technology reduced costs compared with traditional horizontal farming. The reduced costs enabled the client to compete with peers effectively despite the additional market pressures created by reduced consumption during the spread of an animal disease.

Financial institutions with strong risk management and flexible business strategies can better withstand epidemics like COVID-19. COVID-19 was linked to development outcomes in the IEG evaluations and validations of 19 percent of projects IFC-wide, particularly projects in the Financial Institutions Group. The approaches that worked for mitigating the COVID-19 factor were (i) a strong risk management policy that allowed maintaining good credit quality and positive financial results, (ii) sound credit underwriting and a strong business strategy that enabled the client to withstand the pandemic's effects, and (iii) a flexible strategy to stabilize the client's financial position in reaction to the pandemic. However, the consequences of the pandemic detracted from client success, including through drops in business volumes, banks being forced to curtail their small and medium enterprise businesses, and clients downsized. The Financial Institutions Group can help mitigate these negative effects of the COVID-19 factor by selecting clients that have strong risk management and flexible business strategies.

Work Quality and Additionality

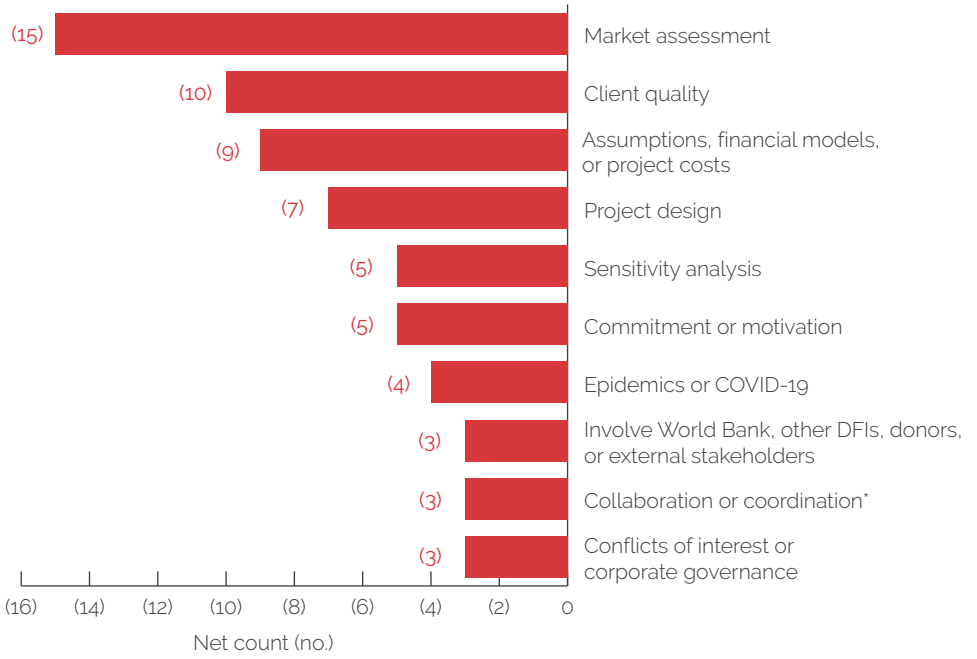
Enhancing front-end work quality and achieving nonfinancial additionality, in addition to providing financing, are important for improving development outcomes. In this subsection, we establish that the top three front-end work quality factors that could contribute to better development outcomes are (i) market assessment; (ii) client quality; and (iii) assumptions, financial models, and project costs.[4] We also suggest that spending sufficient time on these front-end factors may be particularly important for challenging and complex projects.[5] Finally, given the strong relationship between additionality and development outcomes, it is important for IFC to achieve nonfinancial additionality.

Improving front-end work quality can help improve development outcomes. The work quality of IFC investment projects declined from 62 percent rated satisfactory or better in CY13–15 to 55 percent in CY21–23. There is a strong relationship between work quality and development outcomes in projects evaluated and validated by IEG in CY21–23. There is an alignment between work quality and development outcomes in 77 percent of these projects (high work quality and high development outcomes or low work quality and low development outcomes). While 36 percent of the projects had low work quality and low development outcomes, only 9 percent of projects with low work quality delivered high development outcomes. IEG conducted a desk-based review of a universe of 19 projects whose work quality was rated unsatisfactory or whose development outcomes were rated highly unsuccessful in CY21–23 (with a cutoff date of December 31, 2023). The aim was to identify the top IFC work quality factors (figure 3.6) that, if addressed, could contribute to better development outcomes. The top three factors are as follows:

- Market assessment. Fifteen out of 19 projects reviewed had shortfalls in work quality related to market assessments. For example, product pricing on an internet platform was not market tested, leading to overestimating demand and profitability.
- Client quality. Ten out of 19 projects reviewed had work quality issues related to client quality. For example, clients lacked market entry and country experience in six large target markets in Asia. They were unable to enter these markets, and IFC's investment was written off.

» Assumptions, financial models, and project costs. Nine out of 19 projects reviewed had work quality issues related to assumptions, financial models, and project costs. For example, assumptions were not based on feasibility studies or market assessments.

Figure 3.6. Top 10 Work Quality Factors, Net Effect

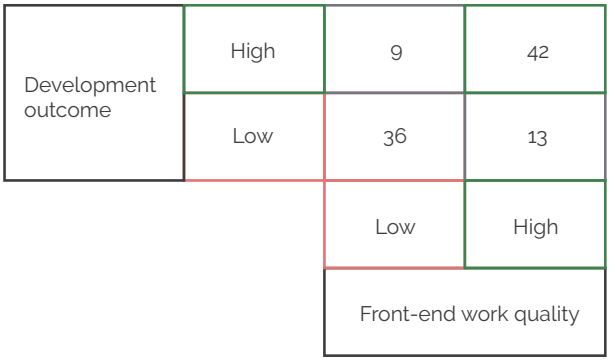


Source: Independent Evaluation Group.

Note: DFI = development finance institution; negative = the identified factor constrained work quality; net effect = sum of positive and negative; positive = the identified factor aided work quality.
*This item refers to collaboration or coordination between advisory services and investment projects.

All these factors fall under screening, appraisal, and structuring (also known as front-end work), which is one of the two components of IFC work quality (the other being supervision and administration). There is a strong relationship between front-end work quality and development outcomes (figure 3.7). In summary, IFC could focus more on front-end work quality processes related to market assessments; client quality; and assumptions, financial models, and project costs. These factors are all within IFC's influence, and efforts to improve them might contribute to better development outcomes.

Figure 3.7. Association Between Front-End Work Quality and Development Outcome, Calendar Years 2021–23 (share of projects, percent)

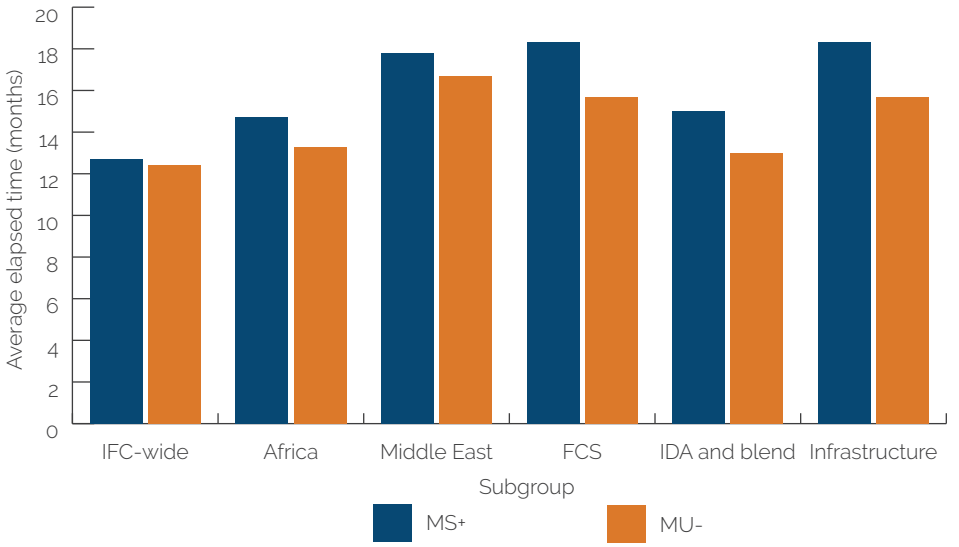


Source: Independent Evaluation Group, Expanded Project Supervision Report database.

Note: Front-end work quality = screening, appraisal, and structuring.

Spending sufficient time on front-end work may be particularly important in challenging contexts and projects that are especially complex. An exploratory analysis showed that processing mostly successful or better projects takes longer than processing mostly unsuccessful or worse projects in challenging contexts (specifically, in Africa, Middle East and North Africa, FCS, and IDA and blend) and projects that are especially complex (the Infrastructure industry group; figure 3.8). Delays could occur at any stage of processing between the Concept Note and first disbursement. However, a variety of project examples suggest that cutting short some of the key front-end work quality factors (such as market assessment; client quality; or assumptions, financial models, and project costs) may contribute to weak development outcomes.[6] As an example of a project that lacked market assessment, in one IFC project, a missing mining study led to the adoption of an unworkable technology. Deficiencies in client quality can be illustrated by a project in which the management team of an African health-care product company lacked experience in integrating large cross-border mergers. The two largest acquisitions were resold at large discounts, resulting in losses for IFC and all other shareholders. In another project, IFC's investment in a follow-on private equity fund was not based on a realistic valuation of the first fund, which distorted its projections.

Figure 3.8. Average Elapsed Time from Mandate to First Disbursement Versus Development Outcome



Source: Independent Evaluation Group.

Note: FCS = fragile and conflict-affected situations; IDA = International Development Association; MS+ = mostly successful or better; MU- = mostly unsuccessful or worse.

There is a strong relationship between additionality and development outcomes. There is an alignment between additionality and development outcomes in 72 percent of projects evaluated and validated by IEG in CY21–23 (low additionality and low development outcomes or high additionality and high development outcomes; figure 3.9). While 36 percent of the projects had low additionality and low development outcomes, only 15 percent had low IFC additionality and high development outcomes. This demonstrates that additionality is an important factor in development outcomes.

Delivering nonfinancial additionalities is important to improving development outcomes. IEG conducted a desk-based review of a universe of 18 projects whose additionality was rated unsatisfactory or whose development outcomes were rated highly unsuccessful in CY21–23 (with a cutoff date of December 31, 2023) to help identify key areas to improve outcomes.[7] We found the following:

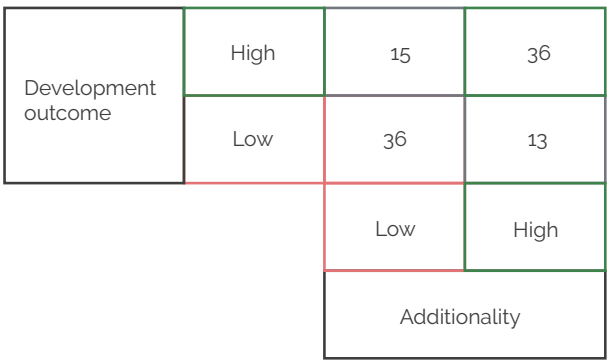
» Financial additionality occurred 20 times. Financial additionality is usually built into the design of IFC's financing in a project. Hence, in its two most common forms, it is realized as soon as the financing is disbursed (World

Bank 2022c). Among the projects reviewed, the most frequent financial additionality was financing (13 occurrences, 7 achieved fully or partially), followed by resource mobilization (4 occurrences, 3 achieved). Financial additionality was achieved (partly or fully) 65 percent of the time.

» Nonfinancial additionality occurred 46 times. In contrast with financial additionality, nonfinancial additionality is usually realized over time rather than at financial disbursement. Nonfinancial additionality goes beyond financial additionality in that it involves helping clients, particularly in challenging contexts, resolve weaknesses in specific areas identified at appraisal based on IFC's global knowledge and expertise. It is delivered during the life of the project and affects development outcomes. Thus, it often requires planning, monitoring, and continuing engagement (IFC additionality in middle-income countries; World Bank 2022c). Among the projects reviewed, nonfinancial additionality was achieved (partly or fully) 41 percent of the time. The most frequent nonfinancial additionality was improving environmental and social standards (15 occurrences, 9 achieved), followed by stamp of approval (8 occurrences, 4 achieved), knowledge sharing (5 occurrences, none achieved), network sharing (5 occurrences, 1 achieved), and corporate governance (4 occurrences, 2 achieved).

» Unforeseen additionality, defined as additionality realized but not anticipated at Board approval, occurred only 3 times.

Figure 3.9. Association Between Additionality and Development Outcome, Calendar Years 2021–23 (share of projects, percent)



Source: Independent Evaluation Group, Expanded Project Supervision Report database.

In summary, IFC may find it more difficult to achieve nonfinancial additionality than financial additionality. Specifically, in projects with weak development outcomes and weak additionality, nonfinancial additionality occurs more frequently than financial additionality but is achieved less often. The leading reason for IFC not achieving anticipated nonfinancial additionality is its failure to deliver anticipated support, often embodied in advisory services (World Bank 2022c).

Achieving nonfinancial additionality is important for development outcomes. Although IFC has more difficulty realizing nonfinancial additionality than it does realizing financial additionality, IEG suggests that several subtypes of nonfinancial additionality are significantly associated with a higher probability of success in several important outcomes (World Bank 2023c). For example, IFC invested in a private solid waste management company to develop three new greenfield municipal waste-to-fertilizer facilities. The project intended to establish a viable template that could attract greater private sector participation in the solid waste management sector in Central Asia. IFC's involvement in this transaction was expected to serve as a strong vote of confidence to municipalities, investors, and other financing institutions. However, the financing was partially disbursed because difficulties with the production process led to a slow ramp-up in volume, resulting in higher costs and flawed market and pricing strategies, which, in turn, led to low acceptance of the product by end users. As a result, the project came to a halt and the stamp of approval did not materialize, which contributed to the highly unsuccessful development outcome.

Nonfinancial additionality often helps improve other indicators that contribute to development outcome ratings. For example, supporting clients in improving their environment and social systems can help them meet IFC's environmental and social standards, ratings of which are one of the four indicators of development outcomes.[8]

Tracking Outcome Indicators and Measuring Results

IFC could improve the measurement of outcomes by recording more complete information about projects. IEG could not verify nearly 100 outcomes in CY21–23. IEG conducted a desk-based review of 173 IFC

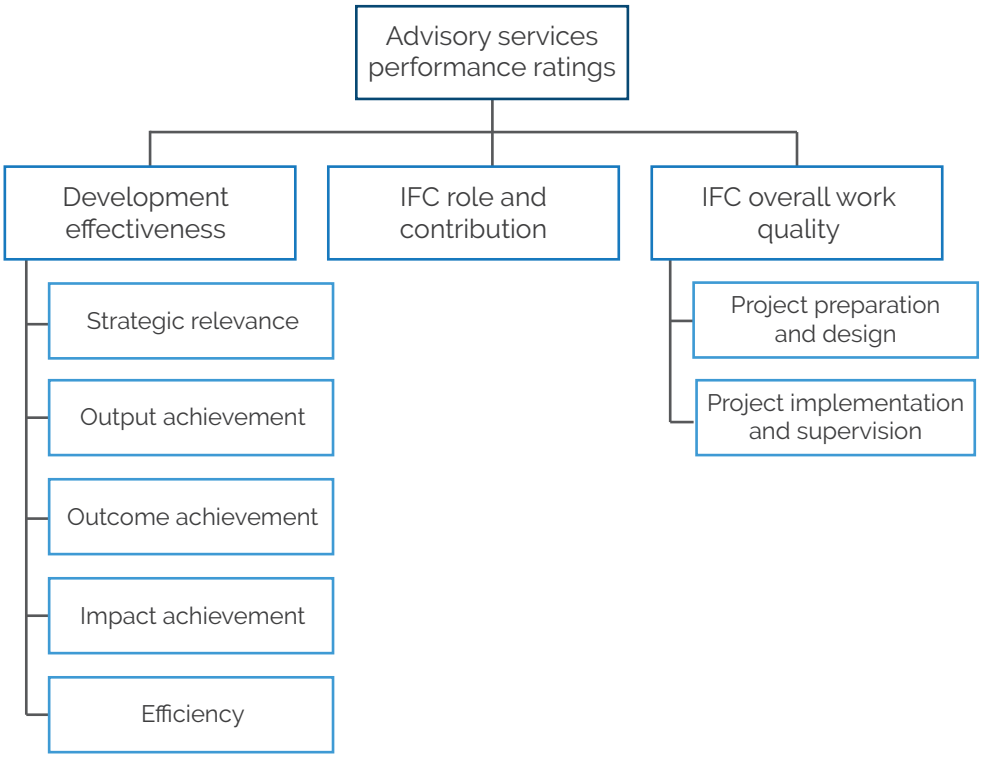
investment projects evaluated and validated by IEG during CY21–23. The review identified 842 individual outcomes (676 project-level outcomes and 166 market-level outcomes) based on the typology developed by *RAP 2021*. Of these, 96 outcomes (11 percent)[9] could not be verified or measured because of (i) lack of indicator, (ii) lack of client reporting, (iii) insufficient evidence, (iv) lack of clarity in how to measure outcomes, (v) issues in attributing results to projects, and (vi) too early to tell. Of the 96 outcomes that could not be verified, 76 were project-level outcomes and 20 were market-level outcomes. This represents 11 percent of the total project-level outcomes and 12 percent of the total market-level outcomes.

IFC could enhance the credibility of its development outcome ratings through improving results measurement via better identification and tracking of outcome indicators, particularly for market outcomes. To assess whether the introduction of the Anticipated Impact Measurement and Monitoring (AIMM) system in 2017 overcomes the issue of verifying outcomes discussed in the previous paragraph and confirms the claims from IFC management during the *RAP 2023* Board discussions, IEG conducted an analysis of a universe of 21 projects evaluated and validated by IEG with "live" AIMM scores as of December 31, 2023.[10] This analysis found continuing challenges in identifying and tracking outcome indicators.[11] Twenty-two percent of 170 outcomes (37 outcomes[12]—17 percent of project-level outcomes and 43 percent of market-level outcomes) did not have an indicator in the tracking system. For an analysis of tracking outcome indicators, we excluded 8 prepaid projects and projects IFC exited earlier than planned from the universe of 21 projects because they may not have been monitored. After excluding 62 outcomes associated with these 8 projects, there were 108 outcomes (94 project-level and 14 market-level outcomes) used for the analysis of outcome tracking. Of these, 26 percent of outcomes (18 percent of project-level outcomes and 79 percent of market-level outcomes) were never tracked (despite having indicators) or could not be tracked (because they did not have an indicator in the tracking system). That is, even with the introduction of AIMM, the identification and tracking of outcome indicators, particularly for market outcomes, remains a challenge. This is within IFC's influence, and addressing it could help facilitate including these outcomes in the new Scorecard, where appropriate.[13]

Advisory Services

RAP's main source of evidence on IFC advisory services projects is a random sample that IEG validates every year. IEG draws a random stratified representative sample (51 percent) annually from among IFC advisory services projects that have been delivered and self-evaluated by IFC. IEG independently validates the IFC self-evaluations. In this section, when we refer to projects "validated by IEG," we mean this sample. "IFC-wide" means across IFC advisory services projects as a whole.

Figure 3.10. Performance Ratings in International Finance Corporation Advisory Services Projects



Source: Independent Evaluation Group.

Note: IFC = International Finance Corporation.

IFC advisory services' performance is assessed on three dimensions: development effectiveness, IFC role and contribution, and IFC work quality. Figure 3.10 shows these dimensions and their respective indicators. Development effectiveness is particularly important in this section. It synthesizes a

project's performance across five indicators: strategic relevance, output achievement, outcome achievement, impact achievement, and efficiency. The rating is on a six-point scale: highly successful, successful, mostly successful, mostly unsuccessful, unsuccessful, and highly unsuccessful. Unlike development effectiveness ratings, IFC role and contribution and IFC work quality (including project preparation and design and project implementation and supervision) are based on a four-point scale: excellent, satisfactory, partly unsatisfactory, and unsatisfactory.

Trends

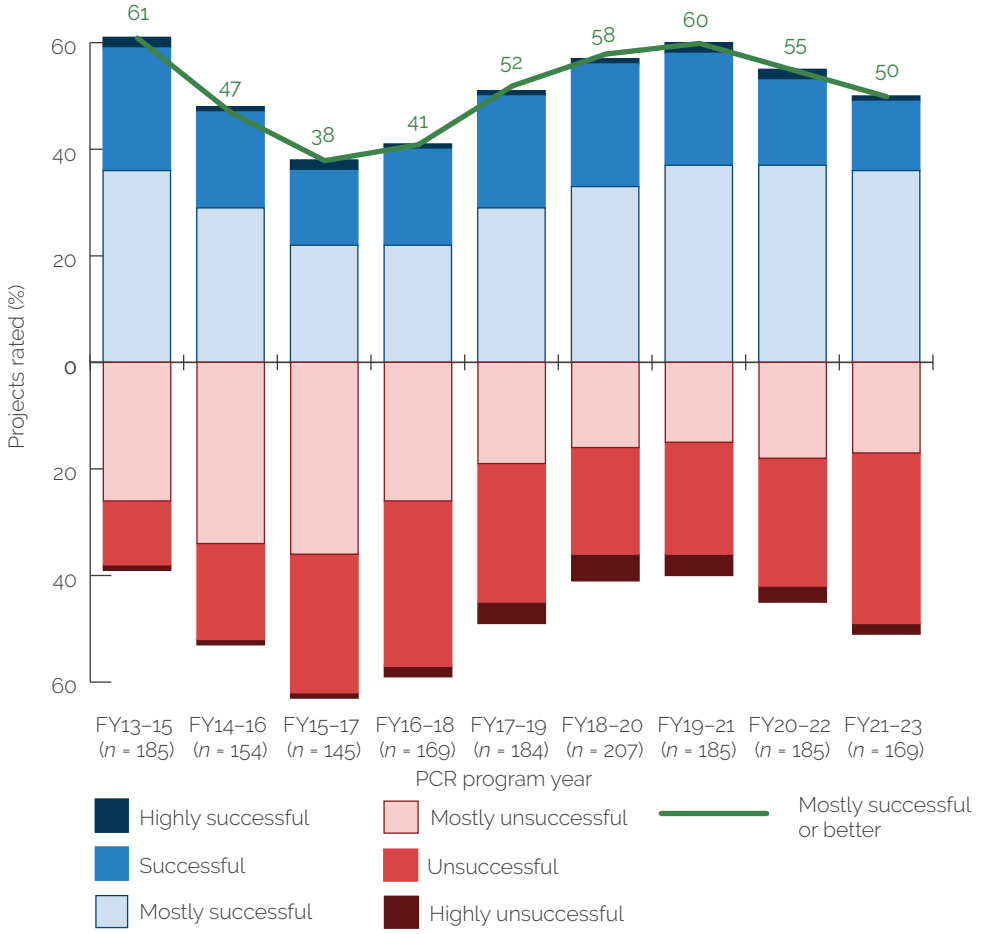
Declines in challenging contexts (IDA and blend) and South Asia have contributed negatively to a long-term decline in the development effectiveness of IFC advisory services projects. In this subsection, we show that IFC advisory services' development effectiveness declined by 11 percentage points over the long term. Projects in IDA and blend countries contributed the most to this decline. South Asia contributed more than any other region.

The development effectiveness of IFC advisory services projects has peaked at times, but it declined over the long term. The development effectiveness of IFC advisory services peaked at approximately 60 percent rated mostly successful or better in FY13–15 and FY19–21. Moreover, development effectiveness of IFC advisory services projects improved over the medium term, from 41 percent rated mostly successful or better (FY16–18) to 50 percent (FY21–23; figure 3.11). However, it declined over the long term, from 61 percent of projects rated mostly successful or better (FY13–15) to 50 percent (FY21–23), and over the short term, from 60 percent of projects rated mostly successful or better (FY19–21) to 50 percent (FY21–23). In the remainder of this subsection, we analyze the development effectiveness in two subgroups—(i) IDA and blend countries and (ii) South Asia—because these subgroups contributed negatively to overall IFC-wide development effectiveness.

Projects in IDA and blend countries contributed the most to the long-term decline in IFC-wide development effectiveness. IDA and blend represent close to 60 percent of IFC advisory services projects validated by IEG in FY13–23. The development effectiveness of projects in IDA and blend countries followed the IFC-wide trend and recently has been below the IFC-wide

average. The percentage of projects rated mostly successful or better declined from 59 percent in FY13–15 to 43 percent in FY21–23, which is below the IFC-wide average of 50 percent for FY21–23 (figure 3.12). This dip in development effectiveness is mainly due to the lower development effectiveness ratings for IDA and blend projects validated by IEG, although the share of these projects also decreased. That is, IFC advisory services projects have been delivering fewer projects and more projects with low development effectiveness in IDA and blend countries. In fact, sensitivity analysis showed that if the development effectiveness of IDA and blend was not considered, then the development effectiveness of IFC would have declined by only 5 percentage points rather than 11 percentage points.

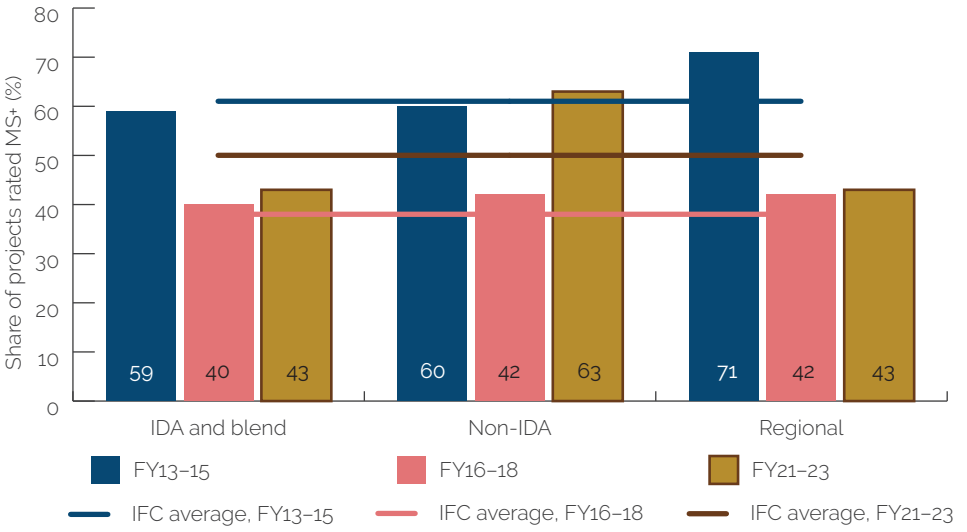
Figure 3.11. Trends in Development Effectiveness for International Finance Corporation Advisory Services Projects, FY13–23



Source: Independent Evaluation Group, PCR database.

Note: Trend line shows mostly successful or better. PCR = Project Completion Report.

Figure 3.12. Performance on Development Effectiveness by International Development Association Status, FY13–15, FY16–18, and FY21–23



Source: Independent Evaluation Group.

Note: IDA = International Development Association; IFC = International Finance Corporation; MS+ = mostly successful or better.

South Asia contributed more than any other region to the long-term decline in IFC-wide development effectiveness. South Asia's development effectiveness declined from 65 percent of projects rated mostly successful or better in FY13–15 to 50 percent in FY21–23. Simultaneously, its share of IFC-wide projects validated by IEG shrank from 17 percent to 11 percent during that time frame.[14] As a result, South Asia's contribution to the decline in IFC's development effectiveness was larger than that of any other region. The decline in the development effectiveness of South Asia was partly due to project preparation and design, the rating of which declined by 8 percentage points between FY16–18 and FY21–23. It was also partly due to implementation and supervision, the rating of which declined by 18 percentage points over the same period. These were the largest declines in these two components of work quality across the regions.

Challenges

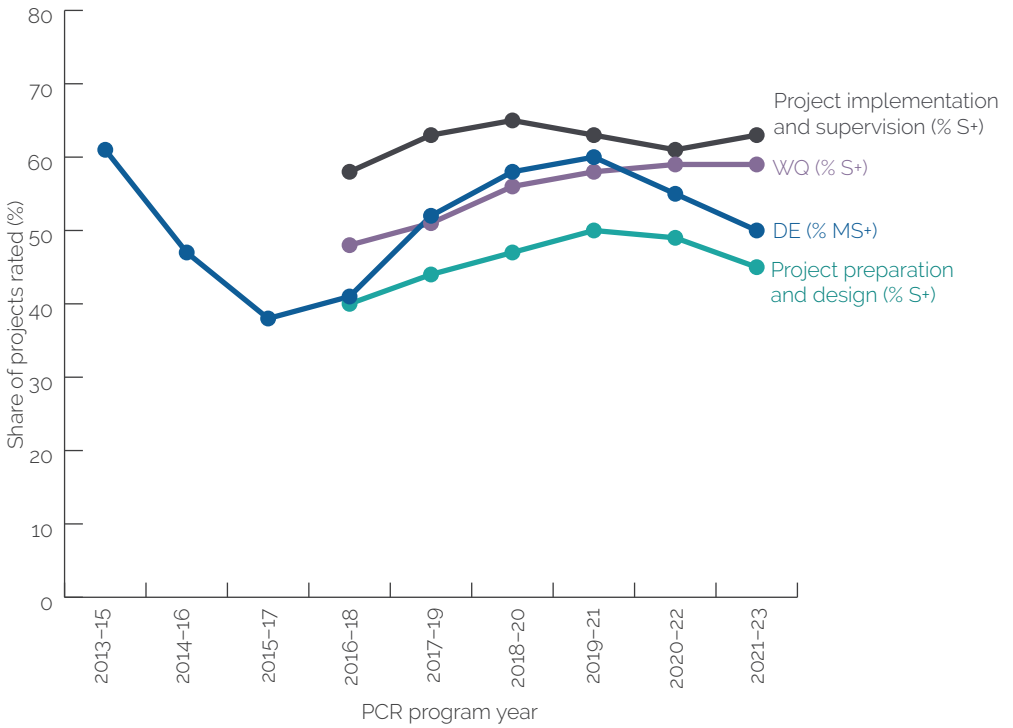
Factors beyond IFC's influence, including COVID-19, have contributed to a recent decline in development effectiveness. In this subsection, we show, first, that despite IFC's efforts (for example, stable work quality ratings), development effectiveness ratings have declined over the past three years of projects validated by IEG. Second, factors beyond IFC's influence have played an important role. Finally, through an analysis of factors linked to inadequate development effectiveness in a sample of projects rated highly on work quality, we show that the most important factor has been COVID-19. There is limited analysis in this subsection because IEG does not have a fully tested taxonomy on factors linked to the development effectiveness of IFC advisory services projects.

Development effectiveness has declined in recent years because of factors outside of IFC's influence. IFC has maintained stable work quality ratings, but development effectiveness ratings have declined. Work quality assesses the extent to which IFC's services ensure quality of preparation and design and support effective implementation. The work quality rating has two dimensions: (i) project preparation and design and (ii) project implementation. Work quality and development effectiveness almost coincided over the FY17–21 period (figure 3.13), but IFC-wide development effectiveness ratings have fallen in recent years (FY21–23), despite IFC doing its part on project preparation, design, and implementation. This shift indicates that factors outside of IFC's influence are dragging down development effectiveness ratings.

COVID-19 is the most important factor in explaining why development effectiveness has declined in recent years despite sustained IFC work quality. IEG tested the hypothesis that external factors beyond IFC's influence contributed to the decline in development effectiveness in FY21–23 despite IFC doing its part on work quality. IEG reviewed 31 IFC advisory services projects validated by IEG for which development effectiveness was rated mostly unsuccessful or worse, but work quality was rated satisfactory or excellent. The review found that in 25 out of 31 projects, COVID-19 strongly contributed to weak development effectiveness.[15] Nineteen of these projects were restructured, 11 of them explicitly due to COVID-19. Seven of these

were restructured early (between March 2020 and July 2020). Regardless, the restructuring was insufficient because all 11 received weak development outcome ratings. The time frame for adaptation may have been a factor: IFC advisory services projects are typically implemented in three years. One example of an IFC advisory services project restructured during COVID-19 is a platform to bring innovative water technologies to Africa, with the objective to make water utilities more efficient, effective, and resilient to shocks. During implementation, the IFC team restructured the project because of uncertainties surrounding the COVID-19 pandemic. However, the project was nevertheless terminated early as a result of a shift to focus on revenue preservation, which was crucial during the pandemic.

Figure 3.13. International Finance Corporation Advisory Services Development Effectiveness, Work Quality, and Project Preparation and Design



Source: Independent Evaluation Group.

Note: DE = development effectiveness; IFC = International Finance Corporation; MS+ = mostly successful or better; PCR = Project Completion Report; S+ = satisfactory or better; WQ = work quality.

Levers

There are two crucial levers for IFC advisory services: improving IFC work quality and systematically recording Concept Note dates. In this subsection, we show that improving work quality could improve development effectiveness. We also argue that systematically recording Concept Note dates would allow IFC to understand the connection between client responsiveness and performance. We did not conduct outcome-type analysis in this subsection because IEG does not have a fully tested taxonomy on outcome types for IFC advisory services projects.

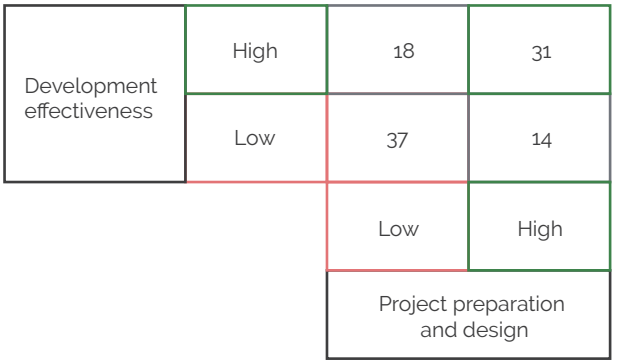
IFC work quality is linked to development effectiveness and is within IFC's influence. There is a strong relationship between IFC work quality and development effectiveness in projects validated by IEG in FY21–23. Work quality and development effectiveness are aligned in 73 percent of these projects (low work quality and low development effectiveness or high work quality and high development effectiveness; figure 3.14, panel a). Although 32 percent of the projects had low work quality and low development effectiveness, only 9 percent had low work quality and high development effectiveness. This evidence suggests that improving work quality could enhance development effectiveness.

The two components of IFC work quality are also linked to development effectiveness. Project preparation and design and development effectiveness are aligned in 68 percent of the projects validated by IEG (figure 3.14, panel b). Although 37 percent of the projects had low project preparation and design and low development effectiveness, only 18 percent had low project preparation and design and high development effectiveness. Project implementation and supervision and development effectiveness are also aligned in 63 percent of the projects validated by IEG (figure 3.14, panel c). Although 25 percent of the projects had low project implementation and supervision and low development effectiveness, only 12 percent had low project implementation and supervision and high development effectiveness. This evidence suggests that improving project preparation and design and project implementation and supervision could enhance development effectiveness.

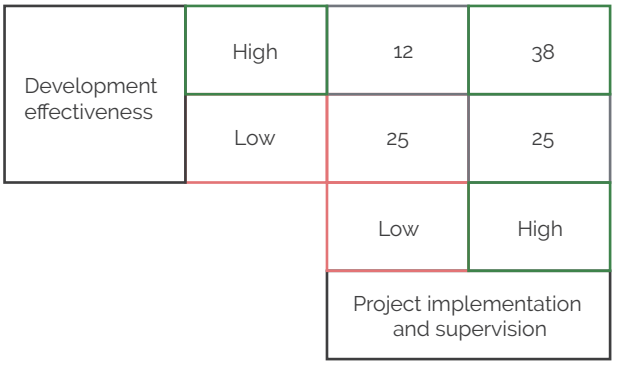
Figure 3.14. Associations Between Work Quality, Project Preparation and Design, and Development Effectiveness, FY21–23 (share of projects, percent)

a. Work quality and development effectiveness

b. Project preparation and design and development effectiveness



c. Project implementation and supervision and development effectiveness



Source: Independent Evaluation Group.

Developing an approach to measure preimplementation scoping time and recording Concept Note dates in the system would allow IFC to understand the relationship between client responsiveness and performance. The Bank Group's senior management has prioritized responsiveness to clients. However, IFC cannot accurately measure preimplementation scoping time (which is a measure of responsiveness to clients) because not all advisory services projects go through the Concept Note stage. For example, Concept Notes may not be needed for some subprojects of approved programmatic umbrellas in which the scoping work was conducted under their respective umbrellas. In addition, Concept Notes may not be needed for "fast-track" projects that were follow-ons from previous engagements wherein additional scoping work was not required. Out of the 411 standard advisory services projects evaluated and validated by IEG (FY13–23), 67 projects (16 percent) did not have a Concept Note date. Developing an approach to measure preimplementation time and recording Concept Note dates in the system would allow IFC to test associations between client responsiveness and performance indicators, such as development effectiveness, IFC work quality, and IFC role and contribution.

[1] Among the regions, development outcomes of Europe improved from 50 percent to 87 percent over the medium term. Among the industry groups, development outcomes of Disruptive Technologies and Funds improved from 14 percent to 48 percent over the medium term.

[2] Strong performance in South Asia was driven by improvements in project profitability and environmental and social effects.

[3] FCS projects in IDA and blend countries refer to a subset of projects in IDA-eligible countries that are FCS.

[4] IEG conducted a desk-based review of a universe of 19 projects whose work quality was rated unsatisfactory or development outcomes were rated highly unsuccessful in CY21–23 (with a cutoff date of December 31, 2023). The aim was to help identify the top three work quality factors that are associated with weak development outcomes and provide a synthesis of evidence on these issues. Because there is only one IFC investment project rated excellent on work quality or highly successful on development outcome, this project was excluded from the analysis. Hence, the cohort consisted of 19 projects with projects rated highly unsuccessful on development outcomes or unsatisfactory on work quality. This analysis will help set the stage for future analysis. Based on this analytic framework, a more balanced approach (in terms of selecting projects on both ends of the spectrum) can be followed in future analysis.

[5] The causality between work quality and development outcomes is not possible because there are factors beyond work quality (for example, macroeconomic factors) that can affect development outcomes.

[6] Assumptions, financial models, and project costs should take into account all key risks embedded in the region and the country.

[7] The study on association between IFC additionality and IFC development outcomes for investment projects is based on a smaller cohort and that the analysis will help set the stage for future analysis in this area.

[8] According to IFC, the new AIMM Navigator has the function to systematically track additionality. In addition, it has dedicated programs (such as the Local Champions Initiative) that identify and support potential clients through upstream and advisory services in areas such as financial management and environment and social compliance and help build a pipeline of bankable projects in FCS. However, IEG has not evaluated these programs or initiatives and their results.

[9] *RAP 2023* reported that 8 percent of the outcomes could not be verified.

[10] Live AIMM scores refer to projects that were assigned ex ante AIMM scores at Board approval.

[11] This exercise included a small number of projects approved in the first year of AIMM implementation, when IFC was transitioning from the Development Outcome Tracking System (DOTS) to the AIMM system. During the transition, the outcome indicator table in the Board paper often maintained DOTS indicators and did not always include the full set of AIMM outcome indicators. In addition, the outcomes assessed in this analysis included not only AIMM outcomes in the Board paper's outcome indicator table but also non-AIMM outcomes in the Board paper and other relevant outcomes discussed in IEG Evaluative Notes. Some of the indicators related to these outcomes, such as enhanced environmental and social standards of clients, were often not included in DOTS.

[12] Among 37 outcomes that did not have indicators in the tracking system, 27 outcomes (73 percent) were specific AIMM outcomes and 10 outcomes (27 percent) were non-AIMM outcomes. According to IFC, indicators for environmental, social, and governance outcomes are recorded and monitored separately in a different system (Sustainability Rating Tool, previously Environmental and Social Review Document) other than the AIMM system and DOTS. Therefore, 10 environmental, social, and governance outcomes (8 on environment and social, 1 on greenhouse gas emissions, and 1 on improved living standards) in this analysis are considered to have indicators and are being tracked by IFC (although in a different system other than the AIMM system and DOTS).

[13] According to IFC, it has made progress with the launch of the AIMM Navigator—a technology platform that allows IFC to track development impact more systematically and consistently within the AIMM framework and that IFC management expects to see an improvement in tracking outcome indicators and ratings over time.

[14] The number of projects in South Asia evaluated and validated by IEG was 31 projects in FY13–15 and 18 projects in FY21–23.

[15] Factors within IFC's influence sometimes mattered. Other prevalent factors included commitment or motivation of the client in 11 projects, change in scope or premature termination of advisory services in 11 projects (5 were influenced by IFC, but 6 were not), and project design in 8 projects.

4 | Multilateral Investment Guarantee Agency

Highlights

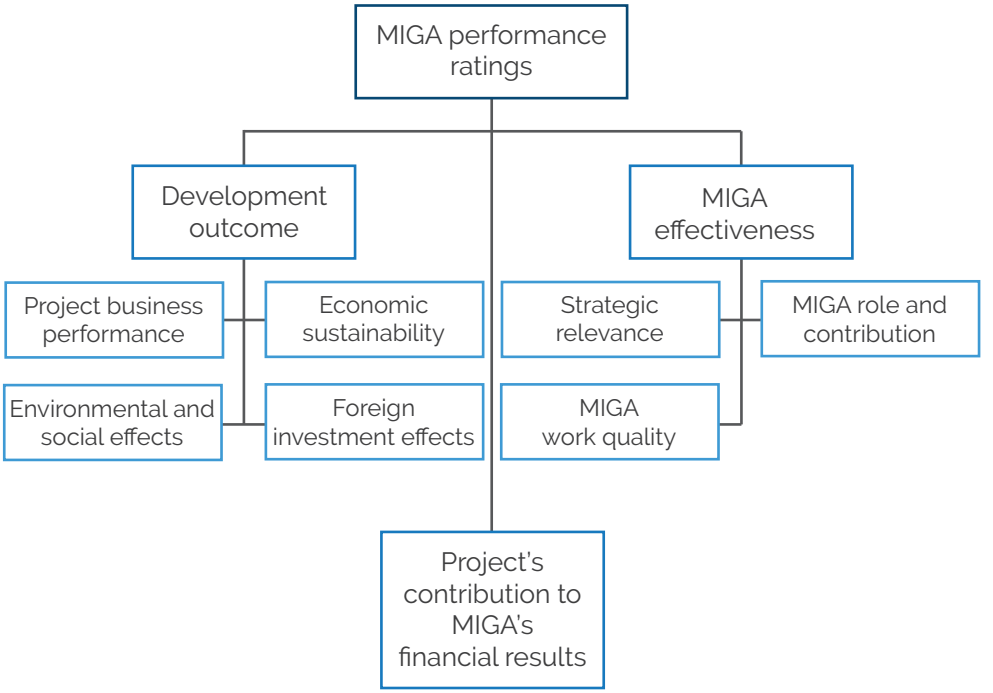
The Multilateral Investment Guarantee Agency (MIGA) is behind on submitting self-evaluations of its guarantee projects, which casts doubt on the validity of its development outcome ratings. Given the available data, development outcomes of MIGA guarantee projects declined slightly over the long term—from 69 percent of projects rated satisfactory or better in FY13–18 to 68 percent of projects rated satisfactory or better in FY18–23—as a result of a drop in ratings in challenging contexts and in infrastructure sector projects.

MIGA achieves foreign investment–level outcomes less often than project-level outcomes. This is a concern because MIGA's business model involves guaranteeing foreign direct investments, which implies that it should achieve foreign investment–level outcomes at least as often as project-level outcomes.

The new World Bank Group guarantee platform provides MIGA with the potential to enhance its supervision beyond environmental and social assessment and country risk assessment, but it also poses risks given MIGA's limited experience in monitoring project results during supervision.

RAP's main source of evidence for MIGA guarantee projects is MIGA's annual self-evaluation of almost all of its guarantee projects; IEG evaluates those that were canceled. For each MIGA guarantee project that has reached early operating maturity (so that sufficient information is available for the evaluation), MIGA underwriting staff conduct a self-evaluation that IEG independently validates. IEG evaluates canceled MIGA guarantee projects in lieu of MIGA self-evaluations. In this section, when we refer to projects "evaluated and validated by IEG," we mean this set of projects.

Figure 4.1. Performance Ratings in Multilateral Investment Guarantee Agency Guarantee Projects



Source: Independent Evaluation Group.

Note: MIGA = Multilateral Investment Guarantee Agency.

MIGA guarantee projects' performance is assessed on three dimensions: development outcome, MIGA role and contribution, and MIGA work quality. Figure 4.1 shows these dimensions and their respective indicators. Development outcome is particularly important for this section and measures performance across four indicators: project business performance, economic sustainability, environmental and social effects, and foreign investment

effects. Development outcomes are rated on a six-point scale: highly successful, successful, mostly successful, mostly unsuccessful, unsuccessful, and highly unsuccessful. Before FY20, the ratings were based on a four-point scale: excellent, satisfactory, partly unsatisfactory, and unsatisfactory.

Trends

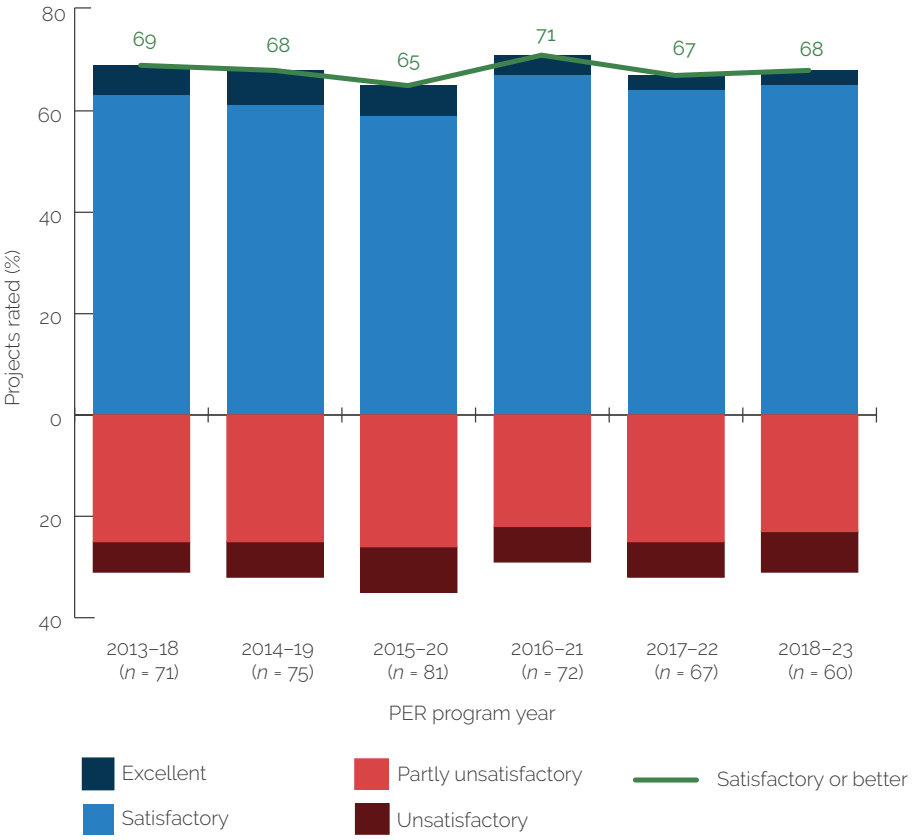
Development outcomes of MIGA guarantee projects declined slightly over the long term as a result of a drop in ratings in challenging contexts and in infrastructure sector projects. For MIGA guarantee projects, we analyzed the trends in development outcomes over a six-year rolling average because of the small number of projects evaluated annually. For this analysis, the long term refers to the period between FY13–18 and FY18–23. In this subsection, we report a slight long-term decline in MIGA's overall development outcomes. MIGA classifies its guarantee projects into four sectors: Agribusiness and General Services, Energy and Extractive Industries, Finance and Capital Markets, and Infrastructure. We found that Infrastructure sector projects, particularly those in IDA and blend countries and in Sub-Saharan Africa, contributed negatively to MIGA's overall development outcomes. However, these projects are in challenging contexts.

Development outcomes of MIGA guarantee projects declined slightly over the long term. Development outcome ratings of MIGA guarantee projects declined from 69 percent of projects rated satisfactory or better in FY13–18 to 68 percent in FY18–23 (figure 4.2). In the remainder of this subsection, we analyze the development outcomes for two subgroups: (i) IDA and blend countries and (ii) Sub-Saharan Africa. These subgroups contributed negatively to MIGA's overall development outcomes. Projects in the Infrastructure sector in these subgroups also contributed negatively to their development outcome ratings.

Projects in IDA and blend countries contributed negatively to MIGA's overall development outcomes. Development outcomes of projects in IDA and blend countries declined from 74 percent of projects rated satisfactory or better to 50 percent over the long term. However, if ratings of projects in IDA and blend countries were not considered, then the percentage of MIGA projects overall with development outcome ratings of satisfactory or better

would have increased by 14 percentage points. This negative contribution to MIGA's overall development outcome rating is expected because IDA and blend countries are challenging contexts.[1]

Figure 4.2. Development Outcomes for Multilateral Investment Guarantee Agency Guarantee Projects, FY13–23



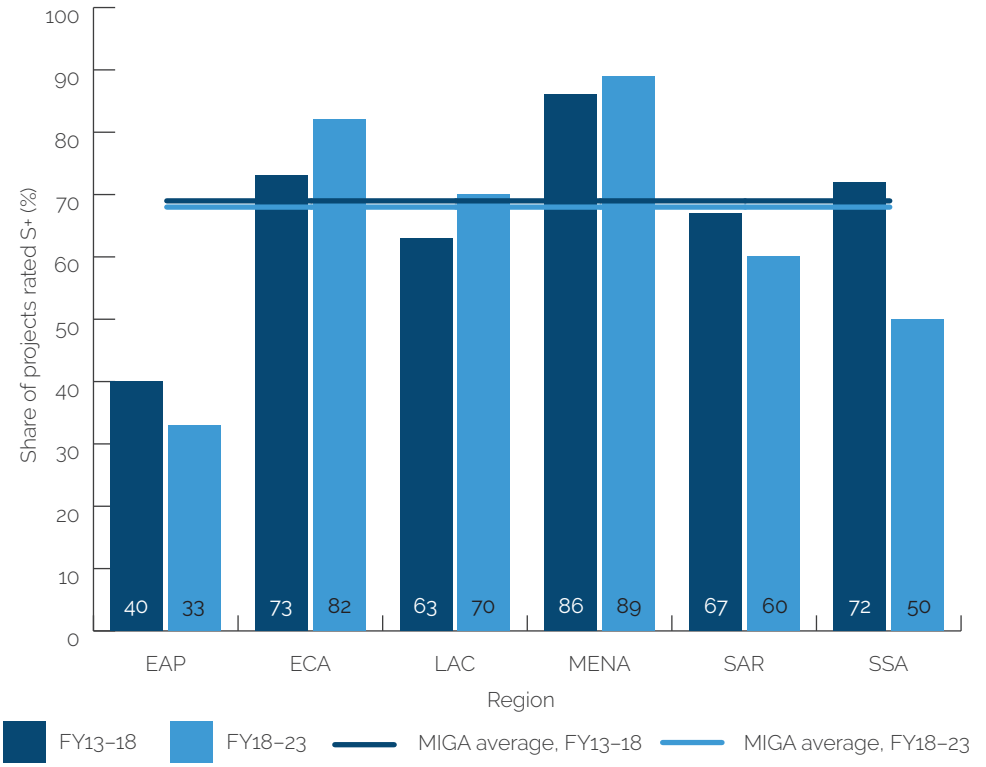
Source: Independent Evaluation Group, Multilateral Investment Guarantee Agency database.

Note: The Multilateral Investment Guarantee Agency Project Evaluation Report guidelines were changed in FY19, replacing a four-point scale for development outcome ratings with a six-point scale. The six-point rating scale, applied to projects starting in FY20, was converted to a four-point one as follows: highly successful = excellent; successful and mostly successful = satisfactory; mostly unsuccessful = partly unsatisfactory; and highly unsuccessful and unsuccessful = unsatisfactory. PER = Project Evaluation Report.

Similarly, projects in Sub-Saharan Africa contributed negatively to MIGA's overall development outcomes. Development outcomes of projects in Sub-Saharan Africa declined from 72 percent of projects rated satisfactory or better to 50 percent over the long term. However, if ratings of projects in

Sub-Saharan Africa were not considered, then the percentage of MIGA projects overall with development outcome ratings of satisfactory or better would have increased by 9 percentage points. Sub-Saharan Africa is a challenging context, so this negative contribution to MIGA's overall development outcomes is expected.

Figure 4.3. Development Outcomes for Multilateral Investment Guarantee Agency Guarantee Projects by Region, FY13–18 Versus FY18–23



Source: Independent Evaluation Group, MIGA database.

Note: EAP = East Asia and Pacific; ECA = Europe and Central Asia; LAC = Latin America and the Caribbean; MENA = Middle East and North Africa; MIGA = Multilateral Investment Guarantee Agency; S+ = satisfactory or better; SAR = South Asia; SSA = Sub-Saharan Africa.

Infrastructure sector projects are a secondary factor in the rating decline in challenging contexts such as IDA and blend countries and Sub-Saharan Africa. Examples of Infrastructure sector projects in challenging contexts include the construction of a dam and the construction of a megawatt power plant in IDA countries in Sub-Saharan Africa, both of which were rated mostly unsuccessful or worse on development outcomes. Development outcomes

of projects in the Infrastructure sector declined from 74 percent of projects rated satisfactory or better to 63 percent over the long term (figure 4.3). In addition, a substantial share of Infrastructure sector projects evaluated and validated by IEG in FY18–23 was in IDA and blend countries (52 percent) and in Sub-Saharan Africa (26 percent). However, if ratings of Infrastructure sector projects were not considered, then the percentage of MIGA projects in IDA and blend countries with development outcome ratings of satisfactory or better would have increased by 16 percentage points, and the percentage of MIGA projects in Sub-Saharan Africa with satisfactory or better ratings would have increased by 11 percentage points. These negative contributions to IDA and blend and Sub-Saharan Africa ratings overall are expected because of the challenging contexts and inherent complexities of Infrastructure sector projects.

Challenges

Critical challenges for MIGA include cost overruns and construction delays, project company quality, legal and regulatory risk, and achieving foreign investment–level outcomes. In this subsection, we identify the most prevalent factors linked to development outcomes: cost overruns and construction delays, project company quality, and legal and regulatory risk. (MIGA's direct client is its guarantee holder, which is the company benefiting directly from the MIGA guarantee. The guarantee holder has direct influence over a project company, which owns and implements the project. Project company quality refers to the ability, technical expertise, and track record of the project company.) We focus on factors that are not directly within MIGA's influence: cost overruns and construction delays, project company quality, and legal and regulatory risk. Finally, we show that MIGA guarantee projects achieve foreign investment–level outcomes less often than they achieve project-level outcomes.

The most prevalent factors linked to MIGA's development outcomes are related to cost overruns and construction delays, project company quality, and legal and regulatory risk. *RAP 2024* conducted a deep dive of 26 MIGA guarantee projects evaluated and validated by IEG (FY20–23) using the *RAP 2023* taxonomy to identify the top factors linked to development outcomes. A factor can have both positive and negative links to development outcomes

(figure 4.4). The most prevalent factor is cost overruns and construction delays, which occurred in 46 percent of projects reviewed. Project company quality occurred in 31 percent of the projects. Finally, legal and regulatory risk occurred in 27 percent of the projects. None of these factors are under MIGA's direct control, as discussed in this subsection. However, MIGA can indirectly influence cost overruns, construction delays, and project company quality through its guarantee holder. This is only true during the underwriting stage because MIGA limits its supervision to environmental and social assessment and country risk assessment.

Figure 4.4. Most Prevalent Factors Linked to Development Outcomes

Source: Independent Evaluation Group, Multilateral Investment Guarantee Agency database.

Note: "Positive direction" means the extent to which the factor contributes positively to the development outcomes of projects evaluated and validated by the Independent Evaluation Group. "Negative direction" means the extent to which the factor contributes negatively to the development outcomes of projects evaluated and validated by the Independent Evaluation Group.

MIGA can indirectly influence project company quality by selecting the right guarantee holder. Project company quality is defined as the quality and experience of the management team implementing the project and their technical skills, track record, contractor competency, familiarity, and acumen. MIGA uses factors such as management experience, sector experience, and commitment to select a guarantee holder, which, in turn, selects a project company to implement a project financed by the guarantee holder.[2] Project company

quality has a positive influence on MIGA's development outcomes in 75 percent of projects. For example, the selection of a guarantee holder with deep experience in the health-care sector that selected a strong project company led to the earlier-than-expected completion of a hospital construction project in Central Asia and Türkiye at a lower budget than expected.

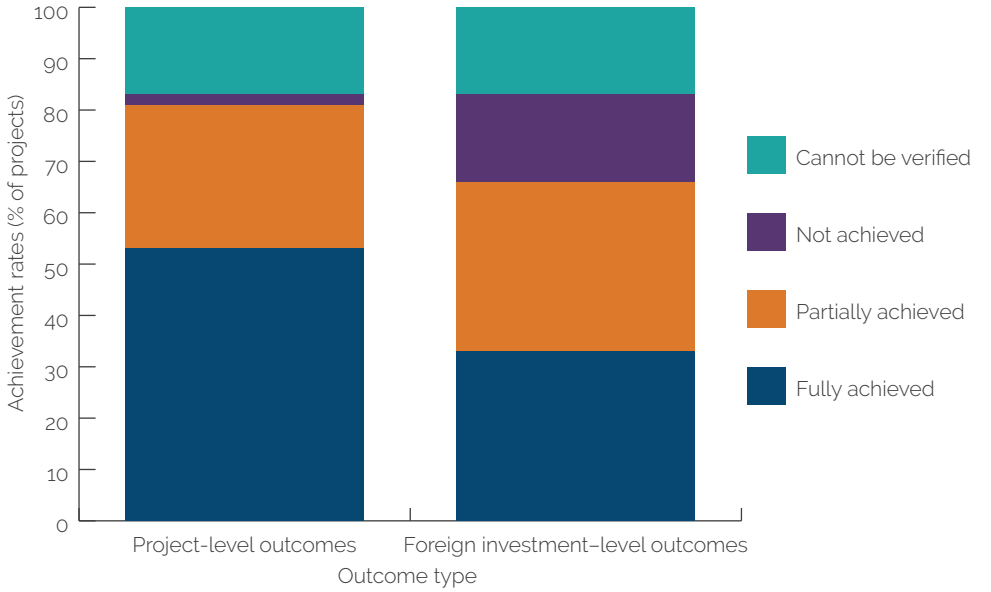
Cost overruns and construction delays are generally outside of MIGA's direct influence. This is a negative factor in two-thirds of the projects. Because MIGA's supervision is limited, it does not directly influence cost overruns or construction delays. However, guarantee holders with strong qualities tend to select project companies that prevent cost overruns and construction delays. Therefore, MIGA must carefully select the right guarantee holder during the underwriting stage to help prevent cost overruns and construction delays. For example, a large-scale power plant construction project in the Middle East was delayed by three years because MIGA's guarantee holder selected an unsuitable project company. Early diligence on selection of the guarantee holder is important because MIGA limits its supervision to environmental and social assessment and country risk assessment.

Legal and regulatory risk is outside of MIGA's influence. Legal and regulatory risk is defined as risk related to regulatory policies, government, legislation, and bureaucratic mechanisms. This factor has a positive influence on development outcomes in about 85 percent of MIGA projects. However, it has a negative influence in about 15 percent of projects. For example, financial performance of a bulk water treatment facility in Sub-Saharan Africa was negatively affected when the government refused to renegotiate a tariff after an unexpected increase in the project's construction costs. Given MIGA's role as a guarantor, it cannot influence governments to change laws, regulations, policies, or bureaucratic mechanisms.

Foreign investment–level outcomes are achieved less often than project-level outcomes for MIGA guarantee projects, which is a concern given MIGA's business model. *RAP 2024* conducted a deep-dive analysis of 15 MIGA guarantee projects evaluated and validated by IEG during FY21–23 to identify project-level and foreign investment–level outcomes and their achievement rates. Eighty-one percent of the 58 project-level outcomes were achieved (fully or partially), whereas 66 percent of the 18 foreign

investment–level outcomes were achieved (figure 4.5).[3] Given that MIGA's primary purpose is to support foreign direct investment, one would expect that it would be at least as good at achieving foreign investment–level outcomes as it is at achieving project-level outcomes. (MIGA launched IMPACT [Impact Measurement and Project Assessment Comparison Tool], an ex ante impact measurement tool, in FY20. Although both IMPACT and AIMM, IFC's ex ante monitoring tool, track project outcomes, IMPACT tracks foreign investment, whereas AIMM tracks market creation. This reflects MIGA's more focused mandate to promote development through foreign investment.)

Figure 4.5. Project-Level Versus Foreign Investment–Level Outcome Achievement Rates



Source: Independent Evaluation Group, Multilateral Investment Guarantee Agency database.

Levers

MIGA could learn more about development outcomes and factors linked to them by promptly delivering self-evaluations and enhancing supervision. In this subsection, we argue that MIGA could catch up on its self-evaluations, which are significantly backlogged, to accurately measure its development outcomes. We also point out that MIGA has the potential to go beyond

environmental and social assessment and country risk assessment and monitor project results in its supervision as it takes ownership of the new Bank Group guarantee platform.

MIGA is behind on submitting self-evaluations of its guarantee projects, which prevents us from having an accurate picture of its overall development outcome ratings. All MIGA guarantee projects are subject to self-evaluation. IEG reviews all completed self-evaluations and validates them to assess the ratings independently. IEG evaluations of MIGA projects enable IEG and MIGA to account to the Board for achievements and contribute to learning. However, 19 MIGA projects are pending self-evaluations during the FY21–23 period (45 percent of planned self-evaluations during this period). Of these 19 projects, 11 involved engagements from MIGA's legal team, including renegotiation of possible investment term modifications. As a result, self-evaluations of these MIGA projects were postponed.[4] The corrective actions and pending self-evaluations from MIGA could influence its development outcome ratings. This, in turn, could affect MIGA's reporting for the new Scorecard.

The new Bank Group guarantee platform provides MIGA with the potential to enhance its supervision beyond environmental and social assessment and country risk assessment, but it also poses risks given MIGA's limited experience in monitoring project results during supervision. Starting July 1, 2024, the Bank Group guarantee platform, housed at MIGA, has brought together guarantee products from the World Bank, IFC, and MIGA. It will serve as a one-stop shop for all Bank Group guarantees. This platform aims to boost Bank Group annual guarantee issuance from $10.3 billion in 2024 to $20 billion by 2030. Given that the platform will encompass guarantees from all three Bank Group institutions, MIGA now has an opportunity to harmonize its monitoring during supervision with the World Bank and IFC. In doing so, it could go beyond environmental and social assessment and country risk assessment and thereby take a more proactive approach to supervision by monitoring project results.

[1] While MIGA's performance in FCS was above the MIGA-wide average both in FY13–18 (78 percent versus 69 percent) and in FY18–23 (75 percent versus 68 percent), the results must be treated with caution because of a smaller cohort size of only 14 projects in total in FCS during FY13–18 (that is, an average of less than 2 projects per year and about 12 percent of MIGA projects evaluated and validated by IEG in FY13–23).

[2] MIGA's business of providing risk coverage to project sponsors, shareholders, or financiers to project companies is different from direct lenders such as the International Bank for Reconstruction and Development, IDA, and IFC. The difference of the business model has an impact on the level of influence on project company.

[3] Based on the consultations between MIGA and IEG, it was agreed to revise project maturity for project evaluations from three to five years, thereby giving more time for the outcomes to be realized and observed.

[4] MIGA does not conduct monitoring and supervision of its guarantees, except for environmental and social issues. As a standard practice before the pandemic, a field visit was conducted as part of MIGA self-evaluation to collect necessary evidence for outcome measurement, but such exercise was not undertaken during the pandemic. The lack of field visits constrained the data gathering from the project. MIGA restarted the self-evaluation based on field visit from the third quarter of FY24 and intensified country visits to cover multiple projects planned for FY25, including projects that were postponed in the previous fiscal years.

5 | Country Program

Highlights

There is no evidence of sustained improvements in country programs: development outcome ratings have increased, but these gains may not hold, and World Bank Group performance ratings have been flat. Between FY13 and FY20, the percentage of countries with moderately satisfactory or above development outcomes ratings rose from 68 percent to 78 percent. Meanwhile, countries with Bank Group performance rated either good or superior fluctuated between 59 percent and 63 percent.

The stagnation in Bank Group performance is concerning because the Bank Group directly manages it. The Bank Group plays a substantial role in its own performance as it makes choices about in which areas, and how, to engage with and support clients based on its own diagnostics, analytic work, and tools to support implementation.

There is 21 percentage point gap between development outcomes in countries classified as fragile and conflict-affected situations and non–fragile and conflict-affected situations (55 percent and 76 percent rated as moderately satisfactory or above, respectively), but the Bank Group performance is approximately the same in both groups (61 percent and 64 percent, respectively).

Results frameworks in Country Partnership Frameworks have persistent shortcomings that affect their ability to support implementation. Evidence on issues with results frameworks is relevant to the new Scorecard because management is considering ways to cascade the indicators into results frameworks across Country Partnership Frameworks (World Bank 2024b).

Country programs do not always deliver Bank Group collaboration that seeks to offer more complete development solutions to clients. In countries with two Completion and Learning Review

Validations between FY13 and FY23, only 40 percent of countries showed collaboration in more than one sector, and collaboration has materialized consistently only in 26 percent of countries.

Relevance, risk identification and mitigation, and support to implementation are the main factors that have strong influences on Bank Group performance in country programs. For example, in 25 out of the 37 countries where Bank Group performance improved or declined, relevance was aligned with the shift in rating.

Adaptive management is relevant for country engagement and needs to be further incentivized. The analysis of Country Opinion Surveys shows that adaptive management factors are rated at a statistically significant lower level for country programs with lower development outcomes and Bank Group performance ratings.

This chapter examines the performance trends of Bank Group country programs in Completion and Learning Review Validations (CLRVs) between FY13 and FY24. We use the term *Completion and Learning Review Validation* to refer to all IEG validations of self-assessments of country program performance.[1] In addition to rating trends, CLRVs delve into the challenges encountered during country program design and support to implementation, exploring their association with the Bank Group performance rating. For each of these analyses, we do not sample but use the population of CLRVs according to our selection criteria (appendix A). To further understand Bank Group performance from a country perspective, we incorporate analysis based on the COS conducted between FY12 and FY23. Typically, client countries will have completed two CLRVs and three rounds of COS within this time frame. Box 5.1 describes the key ratings from CLRVs and the COS that inform this chapter. The analysis also identifies levers within the Bank Group's control that could enhance ratings. Online dashboards can be accessed and enable interested readers to undertake their own breakdowns of the data (appendix B).

Box 5.1. Ratings and Perceptions at the Country Level Analyzed by *Results and Performance of the World Bank Group 2024*

At the end of each country strategy cycle, World Bank Group teams produce a self-evaluation—the Completion and Learning Review. All Completion and Learning Reviews are validated by the Independent Evaluation Group and contain two ratings: development outcome and World Bank Group performance:

» Development outcome. The extent to which the Country Partnership Framework (CPF) was successful in achieving its stated objectives. In determining the level of achievement of each CPF objective, the Completion and Learning Review and the Completion and Learning Review Validation examine the results chain running from the Bank Group interventions through the CPF objective, using evidence from the CPF results framework and additional evidence (if needed) to capture the full extent of each objective. Development outcomes are rated highly satisfactory, satisfactory, moderately satisfactory, moderately unsatisfactory, unsatisfactory, and highly unsatisfactory.

(continued)

Box 5.1. Ratings and Perceptions at the Country Level Analyzed by *Results and Performance of the World Bank Group 2024* (cont.)

» World Bank Group performance. Based on how well the CPF was designed and how well the Bank Group implemented the CPF program, including learning and adapting. Bank Group performance is rated using 13 factors, such as relevance of design, strength of results framework and intervention logic, risk identification, incorporation of lessons learned, quality of implementation support, Bank Group collaboration, use of knowledge services and filling knowledge gaps, learning and adaptation, cooperation with development partners, and attention to safeguard and fiduciary issues. Bank Group performance is rated superior, good, fair, and poor.

In addition to presenting trends in ratings, this chapter also draws on perceptions data from the Bank Group Country Opinion Survey Program. The questionnaire includes overall attitudes toward the Bank Group (for example, concerning trust, relevance, effectiveness in achieving development results, alignment with the country's development priorities, and ability to influence the development agenda) and opinions on the Bank Group's knowledge products, financial instruments, and activities in the country.

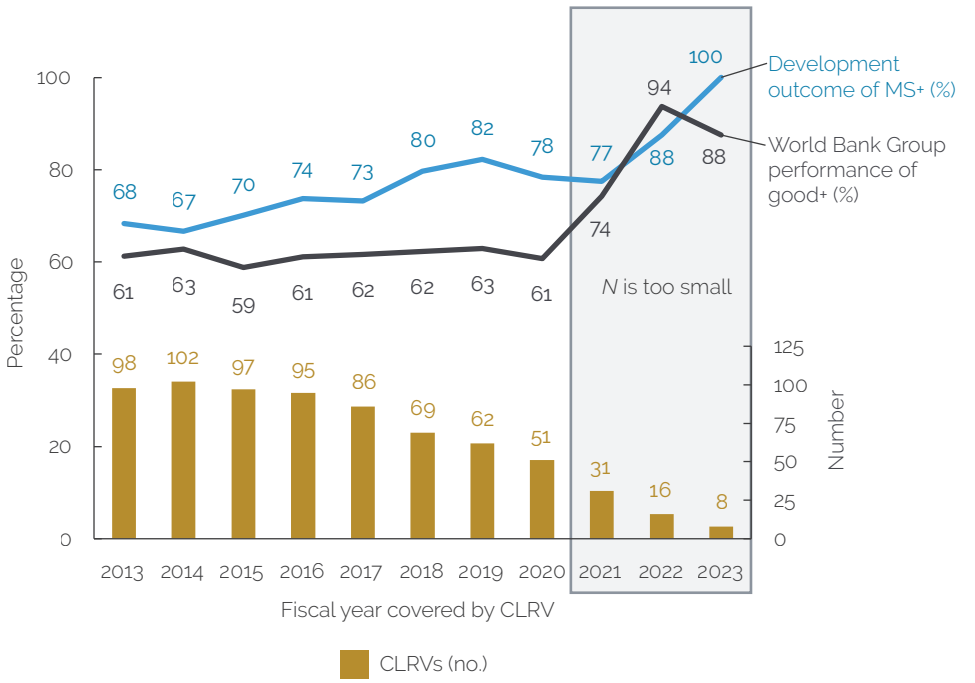
Sources: Independent Evaluation Group; World Bank 2021c, 2022c.

Trends

Development outcomes have risen, but this trend may not hold. In this chapter, "development outcomes" means the share of CLRVs where the development outcome rating is moderately satisfactory or above. Figure 5.1 shows an increase in development outcomes, rising from 68 percent of country programs with a rating of moderately satisfactory or above in FY13 to 78 percent of projects with the same rating in FY20, above the corporate target of 70 percent.[2] Although the share of countries with development outcome ratings remains above 75 percent, experience suggests that ratings for the most recent years fall as new CLRVs are validated. For example, all ratings for FY17–21 have fallen since they were first reported in *RAP 2022*. Consequently, it remains unclear whether the upward trend in recent years

will persist because of a challenging global environment characterized by multiple concurrent crises.

Figure 5.1. Country Program Ratings, FY13–23



Source: Independent Evaluation Group.

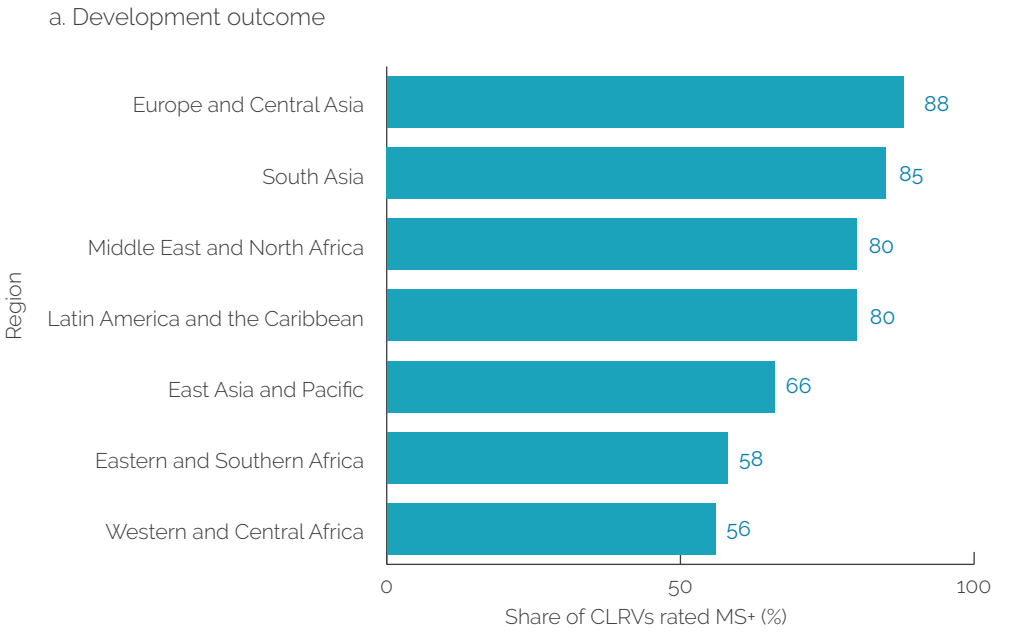
Note: The data are reported with the smoothing approach adopted since World Bank (2020b), in which a CLRV is included in each of the fiscal years covered by the CLRV. A Country Partnership Framework must have closed and its CLR must have been completed before the Independent Evaluation Group produces a CLRV, which often leads to substantial lags. Given the sparse coverage for recent years, we have stopped the analysis of ratings at FY20. The six Organisation of Eastern Caribbean States countries were reviewed together but graphed individually. The total number of included CLRVs is 209. The dashboard that enables further review of country program ratings is available (see also appendix B for more details). CLR = Completion and Learning Review; CLRV = Completion and Learning Review Validation; good+ = good or superior; MS+ = moderately satisfactory or higher.

Bank Group performance has remained stagnant, with low ratings observed in approximately 40 percent of country programs. In this chapter, "Bank Group performance" refers to the share of CLRVs where the World Bank Group performance rating is good or superior. Low ratings refer to country programs with Bank Group performance rated fair and poor. Between FY13 and FY20, overall Bank Group performance fluctuated between 59 percent and 63 percent (figure 5.1), consistently below the corporate target of 75 percent.[3] Given that this rating captures factors managed by the Bank Group, one would anticipate they could be addressed. Notably, Bank Group

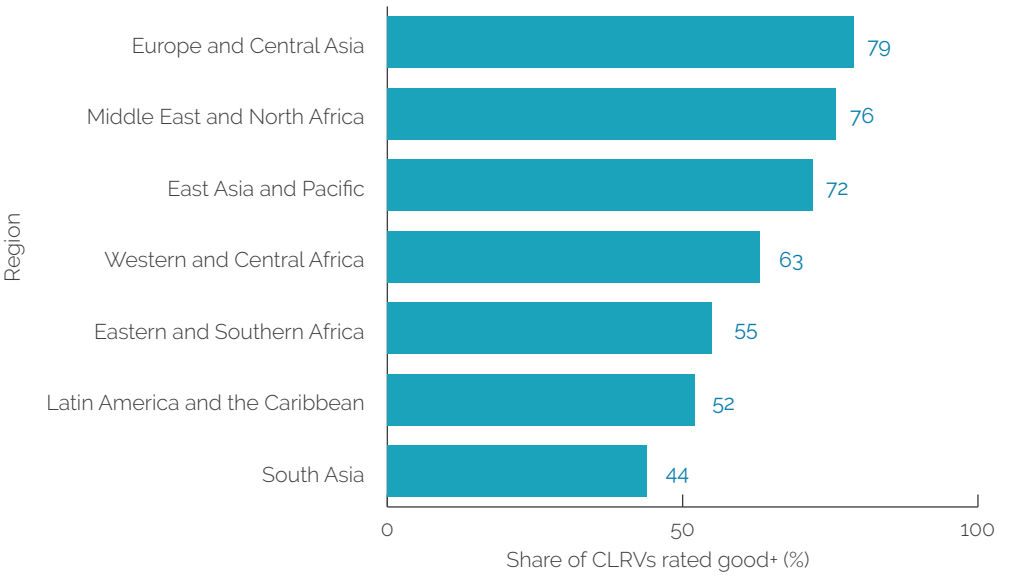
performance is more susceptible to change than development outcomes: more countries shift their Bank Group performance ratings (46 percent) than their development outcome ratings (32 percent) in two consecutive Country Partnership Framework (CPF) cycles. Moreover, there is approximately the same number of countries with improving or declining Bank Group performance ratings.

Country program ratings vary significantly among Regions (figure 5.2). The Europe and Central Asia Region exhibits the highest ratings in both development outcomes and Bank Group performance. Conversely, South Asia has high development outcomes (85 percent) but the lowest Bank Group performance, with only 44 percent of countries rated good or superior across the FY13–23 period.

Figure 5.2. Country Program Ratings by Region, FY13–23



b. World Bank Group performance

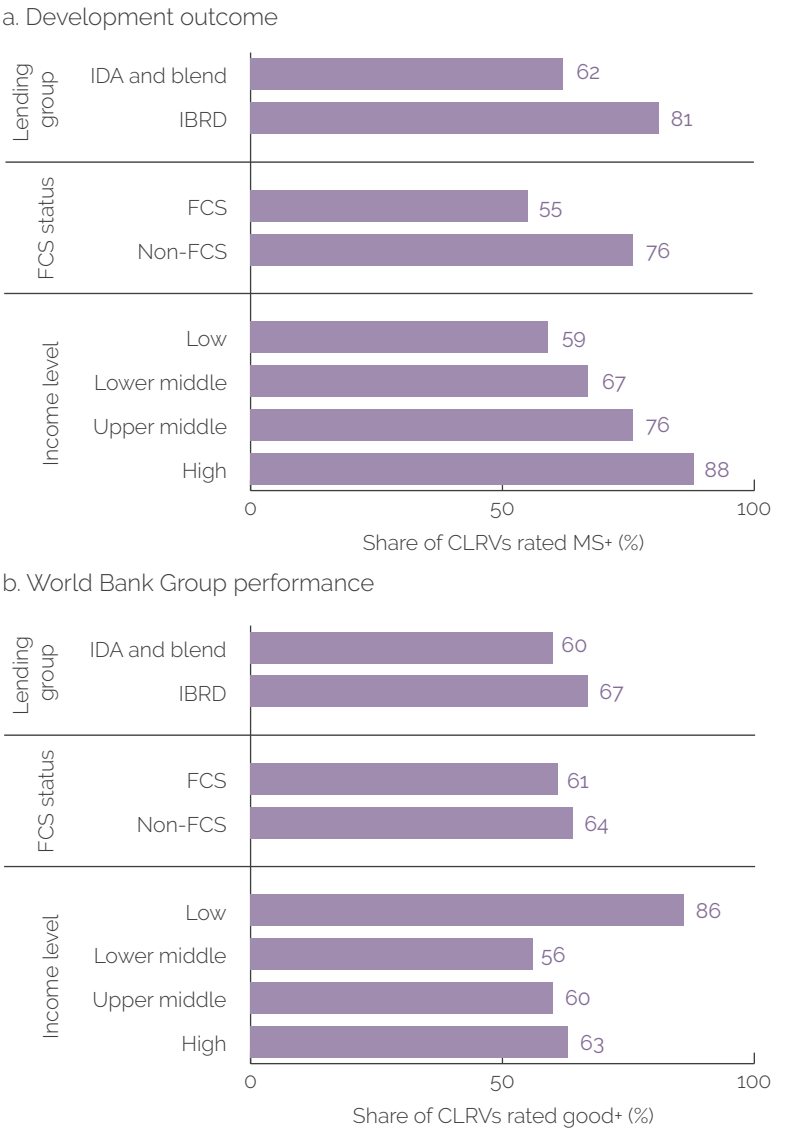


Source: Independent Evaluation Group.

Note: The data are reported with the smoothing approach adopted since World Bank (2020b), in which a CLRV is included in each of the fiscal years covered by the CLRV. CLRV = Completion and Learning Review Validation; good+ = good or superior; MS+ = moderately satisfactory or higher.

Development outcomes are more strongly associated with country characteristics than Bank Group performance (figure 5.3). There is a 21 percentage point gap between development outcomes in FCS countries (55 percent) and the outcomes in non-FCS countries (76 percent), but Bank Group performance is approximately the same in both groups (figure 5.3). Development outcomes rise alongside income levels, increasing from 59 percent in low-income countries to 88 percent in high-income countries. However, the situation reverses for Bank Group performance, with low-income countries outperforming high-income countries by 23 percentage points. The current evidence base does not allow us to explain these patterns in ratings in depth, although it indicates that some types of risks linked to country characteristics are likely more difficult to mitigate in achieving contributions to development effectiveness. For example, as discussed in chapter 2, operations in FCS encounter more country-level contextual obstacles and institutional capacity challenges.

Figure 5.3. Country Program Ratings in the Country's Latest Completion and Learning Review Validation



Source: Independent Evaluation Group.

Note: Income level and lending group are assigned based on the ending fiscal year of the latest CLRV, whereas FCS status considers all fiscal years of the latest CLRV. CLRV = Completion and Learning Review Validation; FCS = fragile and conflict-affected situations; good+ = good or superior; IBRD = International Bank for Reconstruction and Development; IDA = International Development Association; MS+ = moderately satisfactory or higher.

Challenges

Factors with a Strong Influence on World Bank Group Performance

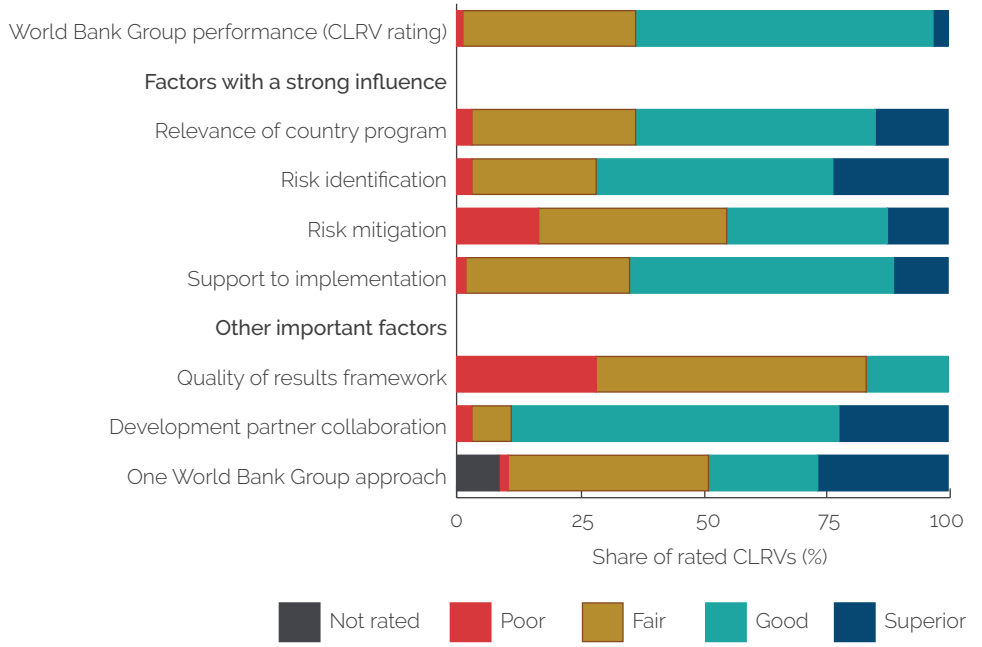
The stagnation in Bank Group performance is concerning because it directly relates to core factors under the Bank Group's control. The country engagement guidance provides a framework and process for the Bank Group to make choices about in which areas, and how, to engage with and support clients (World Bank 2021c). These choices are based on a range of core diagnostics and tools to support implementation.[4] The stagnation of Bank Group performance—despite the wide use of these tools—deserves careful attention. On the basis of a qualitative analysis of 162 CLRVs (the 2 most recent CLRVs for each of the 81 countries with multiple Completion and Learning Reviews validated by IEG in FY13–24), we have identified the four main factors defined in Completion and Learning Review and CLRV methodology that strongly influence,[5] with a high degree of confidence, the Bank Group performance rating: relevance of country program, risk identification, risk mitigation, and support to implementation (figure 5.4).

Relevance has a substantial influence on Bank Group performance, both positive and negative. Relevance reflects the "tailoring" of country programs and includes selectivity and framing of CPF objectives, choice of instruments and interventions, adaptiveness, and realism of program design. Using this definition of relevance, 64 percent of CLRVs pointed to good or superior country program relevance (figure 5.4). Moreover, in most countries where Bank Group performance improved or declined, relevance was aligned with the shift in one direction or the other. This alignment was observed in 25 out of the 37 countries with shifts in Bank Group performance. Examples of common relevance challenges identified in CLRVs include the following:

» Lack of selectivity. In addition to this analysis, IEG has consistently identified lack of selectivity in country programs as a key factor negatively affecting outcomes (World Bank 2015e). For example, the Tanzania FY18–22 program had 15 CPF objectives, including all Systematic Country Diagnostic recommendations plus additional government priorities, despite acknowledging capacity limitations.

» Inadequacy in the selection of instruments. Previously, an IEG synthesis of resource-rich countries found that instruments such as risk sharing, guarantees, and credit information did not receive the attention necessary when the credit risk seemed to have been a bigger constraint than the availability of finance (World Bank 2015f). In Bhutan, although the CPF for FY15–19 planned for a combination of lending instruments, the exclusive use of development policy financing during implementation may have compromised results.

Figure 5.4. World Bank Group Performance at the Country Program Level



Source: Independent Evaluation Group.

Note: The first bar represents the World Bank Group performance rating as validated by the Independent Evaluation Group in the CLRV. The remaining bars were coded by the *Results and Performance of the World Bank Group 2024* team, as elaborated in appendix A. A total of 162 CLRVs are represented, from all 81 countries with at least 2 CLRVs reviewed by the Independent Evaluation Group in FY13–24. Factors are considered to have a strong influence on Bank Group performance if they display a difference significant at the 5 percent confidence level in a *t* test of equality of means between the subgroups of CLRVs with Bank Group performance rating of good or superior and fair or poor. The dashboard that enables further review of factors linked to country program performance is available (see also appendix B for more details). CLRV = Completion and Learning Review Validation.

» Uneven or wavering government ownership of the country program. IEG evidence from operations confirms the importance of developing ownership across multiple interest groups (World Bank 2023a). In Tajikistan in FY15–18, an external shock took the government's attention away from the ambitious

reform agenda. In Mauritius in FY07–15, the initial commitment to the reform effort wavered in response to a newly elected government.

Most country programs did a good job of identifying risks, although sufficient risk mitigation did not always follow (figure 5.4). A review of CLRVs found that 71 percent described good or superior risk identification, whereas 55 percent described good or superior risk mitigation. When country programs were able to adapt to unidentified risks or challenges, they received more positive ratings. Both risk identification and mitigation played a role in countries with changes in Bank Group performance ratings (respectively, in 23 and 28 out of the 37 countries). Further analysis of the materialization of risks in country programs could offer valuable insights into opportunities to enhance performance. Among the reviewed CLRVs, the most common pitfalls regarding risk are as follows:

» Missed or underestimated political risks. With the benefit of hindsight, the Peru FY17–21 CLRV pointed to underestimated political risks, rated moderate both in the original CPF and in the Performance and Learning Review. The CLRV highlights this as a misreading of the political turmoil that was to follow, which had a significant impact on the country program implementation.

» Missed or underestimated capacity risks. In Kosovo in FY17–22, risks associated with institutional capacity constraints were not sufficiently acknowledged during program design. The risks were partly addressed through a portfolio improvement plan, which was underpinned by a thorough analysis of implementation bottlenecks.

Despite the limited depth of discussion of support to implementation in CLRVs, that support is a key factor influencing Bank Group performance ratings. Not all aspects of support to implementation are evenly discussed in CLRVs, which tend to focus on portfolio performance, advisory services and analytics (ASA) delivery, and safeguards and fiduciary issues. Support to implementation was positively assessed in most countries (67 percent) where Bank Group performance improved, with challenges relating to safeguards and fiduciary addressed, as well as marked improvement in the delivery of ASA. Common challenges in supporting implementation include the following:

- Insufficient attention to safeguards and fiduciary issues. During implementation, the Mozambique FY17–21 program faced challenges related to compliance with safeguards, with many difficulties stemming from project implementation units' low capacity to assess and mitigate environmental and social risks.
- Scattered or poorly delivered ASA. In Costa Rica in FY12–15, the ASA program was not strategic enough nor directly connected to the program, with many large analytic reports not followed up with program-related actions.
- Intermittent staff presence. The lack of a World Bank office constrained the scope and effectiveness of the Botswana FY09–13 and Djibouti FY09–13 programs. Similarly, the absence of a country manager and high staff turnover were challenges for the Mauritania FY07–12 program.

Quality of Results Frameworks and World Bank Group Collaboration: Other Important Factors

The quality of results frameworks, collaboration with development partners, and collaboration within the Bank Group approach are important challenges, even though they have limited influence on Bank Group performance. Our qualitative analysis of 162 CLRVs codified Bank Group performance on three additional factors: quality of results framework, development partner collaboration, and the One World Bank Group approach (figure 5.4). While these factors exhibit a similar distribution across CLRVs—irrespective of whether the Bank Group performance rating is fair or poor or good or superior—they remain critical. These factors are included as criteria within the country engagement guidance's methodology for assessing country programs. Furthermore, there is recognition in the literature of their role in enhancing project outcomes (World Bank 2015d) and mobilizing additional development resources (Eriksson 2001).

Results frameworks have shortcomings that affect their ability to support implementation. The evidence indicates that a weak results framework is "a key determinant of unsatisfactory outcome performance at the country program level" (World Bank 2015d, 1). Nevertheless, 83 percent of CLRVs reported major inadequacies in results frameworks, which disconnects from the overall Bank performance rating, where 36 percent of country programs

were rated weak. Moreover, there has been no substantial improvement; out of 81 countries, 16 improved their results frameworks to a good rating in their most recent CLRV, while 9 declined to fair or below. Frequent shortcomings are noted with the intervention logic and the chosen indicators. For example, indicators may focus on inputs or outputs rather than outcomes, or they may be overly reliant on indicators of operations that fail to capture the full extent of the country program objectives and do not properly account for the contributions of ASA, IFC, MIGA, policy dialogues, or the Bank Group's convening role. Although Performance and Learning Reviews often adjust the results framework, many weaknesses remain unresolved. Previous IEG reports have raised these concerns and also found that these practices generate incentives not aligned with an outcome orientation at the country level (World Bank 2020d, 2022c).

Insufficient collaboration with development partners still affects some country programs. Duplication of efforts and redundancies between donor programs occur only in 11 percent of country programs. Nevertheless, the Bank Group has historically acknowledged "the weaknesses of uncoordinated aid" (World Bank 1984, 57) and long understood the barriers that impede effective donor coordination (Eriksson 2001). Thus, it is reasonable to strive for effective collaboration between the Bank Group and development partners in all countries.

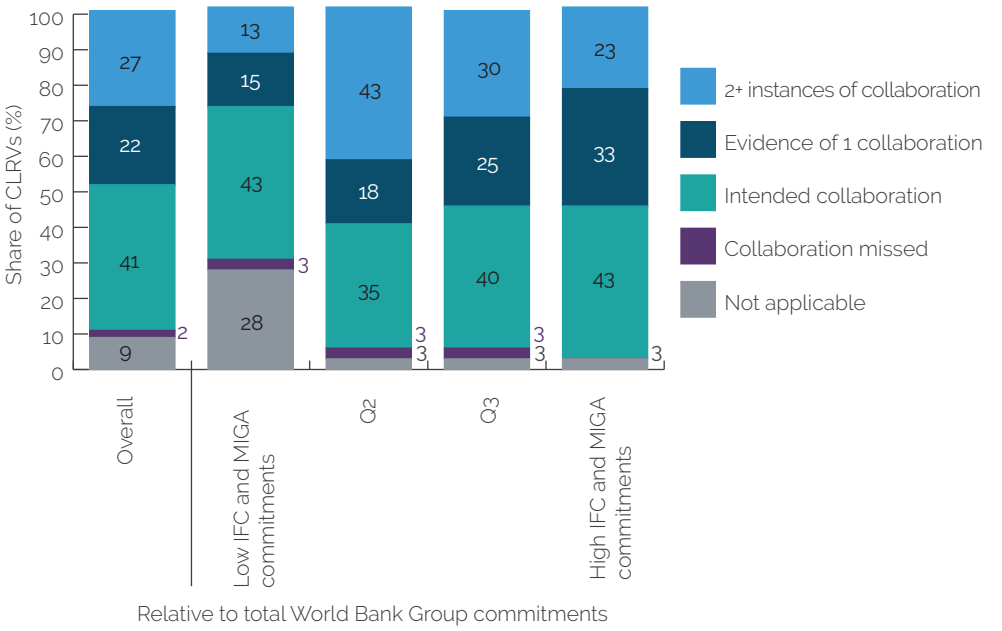
The intent to undertake Bank Group collaboration and offer more complete development solutions to clients is a feature of country programs. The intent to exploit synergies between Bank Group institutions to respond to client demands through enhanced collaboration at the country level has been sought over almost three decades. In 1996, the first joint country assistance strategies were defined; in 2013, the One World Bank Group strategy was released; the cascade approach was introduced in the 2017 IFC 3.0 strategy; and the Maximizing Finance for Development strategy was introduced in 2018. Currently, joint country representation has been introduced as part of the Better Bank initiatives to help enhance collaboration. Four types of Bank Group collaboration were identified within sectors in CLRVs: parallel and complementary, joint projects, sequenced interventions, and separate but coordinated work. In the reviewed CLRVs, an intent for collaboration among Bank Group institutions in one of these four types was identified in

90 percent of country programs in at least one CPF period. When this intent is realized, IEG evaluations have found that Bank Group collaboration can improve the performance of sectors in country programs because they are better able to address client needs with the tools and expertise of all three institutions at their disposal (World Bank 2016d, 2020a).

Bank Group collaboration has increased in country programs in a limited number of sectors and inconsistently over CPF periods. A review of CLRVs showed that for CPFs starting before FY16, less than half of country programs demonstrated collaboration in at least one sector. Since FY16, 28 out of 33 country programs undertook collaboration on interventions. Of these, just under two-thirds (18 out of 28) demonstrated collaboration with more than one sector. This collaboration mainly occurs in five sectors: the energy, financial, infrastructure, investment, and agriculture sectors account for just over three-quarters of the instances of collaboration. Moreover, collaboration as One World Bank Group has yet to arise consistently in countries across two CPF periods. Across the FY13–23 period, collaboration has materialized consistently only in 21 out of the 81 countries that had two Completion and Learning Reviews validated by IEG. The different forms of Bank Group collaboration, their instruments, and the sectors are discussed in more detail in appendix J.

Collaboration can occur even without joint financing. Collaboration occurs even when IFC and MIGA have small portfolios relative to the World Bank's—that is, financing is not always necessary. Figure 5.5 shows that outside the countries with the lowest share of MIGA and IFC commitments, more than half have examples of collaboration in Bank Group activities. A review of CLRVs found that 30 percent of the examples discussed involved joint financing. The remaining 70 percent of collaboration examples involved financing from one Bank Group entity and advisory services (for example, advisory services from IFC or ASA from the World Bank) from another or involved purely ASA. This finding suggests that countries can undertake Bank Group collaboration even when commitments from IFC and MIGA are low. IEG evidence suggests that collaboration can be enhanced through an alignment of objectives between Bank Group institutions and a clear understanding of priorities (box 5.2).

Figure 5.5. World Bank Group Collaboration at the Country Program Level



Source: Independent Evaluation Group.

Note: The data are based on the review of 162 CLRVs by the ***Results and Performance of the World Bank Group 2024*** team (all 81 countries with at least 2 CLRVs reviewed by the Independent Evaluation Group in FY13–24). The first bar has all CLRVs, while the remaining bars split country programs according to the relative size of IFC and MIGA programs as quartiles of the ration between IFC and MIGA commitments and total World Bank Group commitments. Country programs that were exclusively the International Development Association and the International Bank for Reconstruction and Development programs are marked as not applicable. CLRV = Completion and Learning Review Validation; IFC = International Finance Corporation; MIGA = Multilateral Investment Guarantee Agency; Q = quartile.

Box 5.2. Important Conditions for World Bank Group Collaboration

The likelihood of World Bank Group collaboration can be increased. For there to be opportunities for Bank Group collaboration, and for such collaboration to be successful, the two conditions should be met:

1. Alignment of objectives. For collaboration to be feasible and successful, there needs to be an opportunity for collaboration that aligns with the objectives of Bank Group institutions. For the International Finance Corporation and the Multilateral Investment Guarantee Agency, this means that there needs to be an opportunity with

(continued)

Box 5.2. Important Conditions for World Bank Group Collaboration (cont.)

a private sector business case in the short to medium term. For the World Bank, governments need to be willing to initiate reforms that improve private investment.

2. Clear and shared view of sector priorities. To collaborate, Bank Group institutions need a shared view of a sector's priorities. This condition includes understanding actors, their history, opportunities to grow the sector, constraints to realizing those opportunities, and what needs to be done to ameliorate those constraints.

To establish these conditions, three interconnected measures can be used:

1. The Bank Group can use analytics and advisory services to identify areas of alignment or clarify sector priorities. For example, in the Philippines, the International Bank for Reconstruction and Development and International Finance Corporation assessment of the competition environment helped develop an influential dialogue on competition policies at the national level that was integral to the preparation of the new development policy loan series.

2. A shared strategy that is more granular and flexible than a Country Partnership Framework helps make collaboration more effective. For example, the country team in Bosnia and Herzegovina prepared joint business plans that specified milestones for joint Bank Group cooperation.

3. Close coordination between staff from different Bank Group entities has been a factor in successful collaboration. Professional relationships between World Bank and International Finance Corporation staff facilitate knowledge exchange and readiness to work together. This has been facilitated by joint Global Practices, colocation of staff, and informal networks.

To develop these findings, we reviewed 147 Completion and Learning Review Validations, all *Results and Performance of the World Bank Group* reports, all Country Program Evaluations, and select evaluations completed by the Independent Evaluation Group since FY13.

Sources: Independent Evaluation Group; World Bank 2013a, 2013b, 2014, 2015a, 2015b, 2015c, 2015f, 2015g, 2016a, 2016b, 2016c, 2016d, 2016f, 2017b, 2018c, 2019a, 2019b, 2020b, 2020c, 2021a, 2022a, 2022d, 2022e, 2022f, 2022g, 2023f, 2023g, 2023h.

Levers

Extensive adaptive management practices can be found in CLRVs that often positively describe links to performance. IEG had previously identified several types of adaptive management used during country programs (adapted from World Bank 2020d), with some examples found in CLRVs for country programs closed between FY20 and FY23:

- Changes in practices or portfolio composition as context responsiveness. This is the most frequent adaptation (identified in 84 percent of the recent country programs). The North Macedonia FY19–23 Performance and Learning Review introduced a new objective to strengthen the program's focus on supporting private sector competitiveness, innovation, and resilience; activated the Contingency Emergency Response Component; and approved the Emergency COVID-19 Response Project to support efforts to meet the challenges of the pandemic.
- Attention of staff and management to specific projects and actions to resolve problems during implementation. This was the second most frequent adaptation (identified in 72 percent of the recent country programs). As a response to project delays in the Kosovo FY17–22 program, the World Bank, for example, increased attention to problem projects and restructured complex ones, enhanced fiduciary support through hands-on guidance and training, and strengthened portfolio monitoring and review meetings with the client.
- Resource allocation shift during implementation in response to client dialogue. The third most cited adaptation was identified in 70 percent of the recent country programs. The Peru FY17–21 program shifted from a relatively small portfolio focused on IPF to an ample program based on development policy financing to support reforms for economic recovery and to respond to the changing client demands after the COVID-19 outbreak.
- Results reporting and organizational learning. The fourth most cited adaptation was identified in 38 percent of the recent country programs. In the Uruguay FY16–20 program, the Bank Group adjusted two objectives to sharpen its focus and better reflect the government's support for the climate action agenda.

Client country perceptions indicated that adaptive practices are associated with Bank Group performance and development outcome ratings. Respondents to the COS provided perceptions that are linked to adaptive practices such as staff accessibility, responsiveness to country needs, and flexibility as circumstances change. When country programs had lower Bank Group performance and development outcomes, COS respondents averaged less favorable perceptions of these adaptive practices (table 5.1). In addition, adaptive processes are consistently identified as important to CPF design and implementation (World Bank 2020d, 2021c). This would suggest that focusing on adaptive practices to better meet country needs can also help improve ratings.

Teams can improve Bank Group performance by enhancing the way they use adaptive management, yet it is not incentivized in the country engagement guidance. According to *The World Bank Group Outcome Orientation at the Country Level*, Bank Group country teams practice adaptive management, but the country-level results system does not effectively support them in doing so (World Bank 2020d). Instead of using the tools in the country engagement model (for example, Performance and Learning Review and the CPF results framework), tacit knowledge, professional experience, and professional networks are relied on when making adaptive decisions. In the current country engagement guidance, adaptive practices are distributed across several factors considered in rating Bank Group performance. Moreover, the guidance does not specify beneficial approaches or provide explicit incentives for improving the focus on adaptive processes. In light of this evidence, defining incentives and describing important types of adaptive management could enhance Bank Group performance.

Table 5.1. Client Country Perceptions and World Bank Group Performance

Survey Question	Client Perceptions					Adaptive Practice
	Mean if World Bank Group performance is good or superior	Mean if World Bank Group performance is fair or poor	Gap between groups	*p* value		
World Bank Group staff accessibility	6.55	6.14	−0.41	.003	***	Yes
Alignment with the country's priorities	6.58	6.28	−0.31	.003	***	Yes
Flexibility as circumstances change	6.08	5.79	−0.29	.003	***	Yes
Timeliness of financial support	6.29	5.99	−0.29	.045	**	Yes
Financial instruments meet needs	6.36	6.07	−0.28	.005	***	Yes
Responsiveness to country needs	6.43	6.19	−0.24	.015	**	Yes
Collaboration with development partners	7.00	6.76	−0.24	.009	***	No
Effectiveness in achieving development results	6.64	6.44	−0.20	.020	**	No
Relevance of the Bank Group's role	6.78	6.64	−0.14	.110		No
Influence on the development agenda	6.74	6.67	−0.07	.320		No
Technical quality of knowledge work	6.93	6.88	−0.05	.340		No
Being a long-term partner	7.79	7.81	0.03	.600		No

Source: Independent Evaluation Group, based on Country Opinion Survey data collected between 2012 and 2022.

Note: All scores are measured with the Likert scale: 1 = to no degree at all; 10 = to a very significant degree. Survey questions are reported as averages based on country years that match the 162 CLRVs, coded by the ***Results and Performance of the World Bank Group 2024*** team, split by their Bank Group performance rating. The table reports the means, the difference between groups, and the one-sided *p* value in a *t* test of equality of means between the subgroups of Bank Group performance in the CLRV rating of good or superior and fair or poor. Rows are ordered by the gap between groups. The results from splitting country programs by their development outcome ratings, rather than Bank Group performance ratings, are similar. Significance level: * = 10 percent; ** = 5 percent; *** = 1 percent. CLRV = Completion and Learning Review Validation.

[1] The term *CLRV* is used to refer to all IEG validations of self-assessments of country program performance. Previously, these were labeled as Completion and Learning Review Reviews, Country Partnership Strategy Completion Reports, and Country Assistance Strategy Completion Report Reviews.

[2] The corporate target of 70 percent of moderately satisfactory or higher development outcome ratings in CLRVs was included in the World Bank Group Corporate Scorecard FY19–23. It has been superseded by the new World Bank Group Corporate Scorecard FY24–30, which does not include such a target.

[3] The corporate target of 75 percent of good or superior Bank Group performance ratings in CLRVs was included in the Bank Group Corporate Scorecard FY19–23. It has been superseded by the new Bank Group Corporate Scorecard FY24–30, which does not include such a target.

[4] Country programs rely on a range of core diagnostics (such as Systematic Country Diagnostics, Country Economic Memorandums, Country Private Sector Diagnostics, and Country Climate and Development Reports) and tools to support implementation (such as SORT, Implementation Status and Results Reports, and Country Portfolio Performance Reviews).

[5] We performed a qualitative analysis of Bank Group performance at the country program level based on the examination of 162 CLRVs (the 2 most recent CLRVs for each of the 81 countries with multiple Completion and Learning Reviews validated by IEG in FY13–24). Based on a similar analysis conducted for *RAP 2022* (World Bank 2022c) and incorporating descriptions from the country engagement guidance (World Bank 2021c), we rated the following seven factors: relevance of country program, risk identification, risk mitigation, support to implementation, quality of results framework, One World Bank Group approach, and development partner collaboration. We also examined text related to adaptative management practices, distributed across these seven factors. Factors are considered to have a strong influence on Bank Group performance if they display a difference significant at the 5 percent significance level in a *t* test of equality of means between the subgroups of CLRVs with Bank Group rating of good or superior and fair or poor. Detailed methods are elaborated in appendix A.

6 | Conclusions

The findings that emerge from this report are relevant to the Bank Group's evolution toward a Better Bank. The Better Bank initiatives implement changes in how the Bank Group operates, building on processes initiated by the Board of Governors at the annual meetings in 2022 as part of the Bank Group's evolution. The findings presented in *RAP 2024* are relevant to the cross-cutting issue of FCS,[1] the operational efficiency and the effectiveness initiative regarding preparation time and processing time, the new World Bank Group Scorecard, and the joint country representation initiative.

Fragile and Conflict-Affected Situations

World Bank operations in FCS could improve performance through simplified design and risk identification and mitigation of institutional capacity challenges. The share of operations in FCS, which historically have lower outcome ratings, has risen. The proportion of closed operations with full or partial exposure to FCS increased from 31 percent in FY20 to 37 percent in FY23. These operations have a more acute array of challenges. Appropriate design and risk management of operations in FCS correlate with improved outcome ratings. *Mid-Term Review of the World Bank Group Strategy for Fragility, Conflict, and Violence (2020–2025)* noted that crisis-related operations could perform well when they are focused, simple, and realistic in nature (World Bank 2023d). Evidence from *RAP 2024* reinforces this finding because operations in FCS whose objectives include a focus on access have significantly higher ratings. Yet a focus on access needs to be balanced with mitigating institutional risks and preserving institutional capacity. The share of operations reporting one or more institutional challenges and the average number of challenges per project are both higher in FCS than the overall portfolio.[2] An increase in reported challenges in these areas increases the likelihood of lower ratings.

In challenging contexts where selecting high-quality clients may not be feasible, IFC can influence client quality by providing support for capacity building. In Africa and IDA and blend contexts in FCS, it is particularly

important for IFC to select clients with proven business models (that is, those that are tested locally or that can adapt to local conditions when replicated from another country) to reduce business risk. Moreover, in these contexts, there may not be a robust pipeline of bankable projects and experienced clients with the capacity to successfully implement the projects. Nevertheless, IFC should be aware of client limitations and can build a client's capacity by providing nonfinancial additionalities (for example, technical assistance through advisory services during supervision).

Operational Efficiency and Effectiveness: Preparation and Processing Time

World Bank operations with long preparation time (above the 90th percentile) are associated with significant challenges at closure. The World Bank has recently made efforts to substantially shorten preparation times as part of the evolution (World Bank 2024c). Operations with the longest preparation times (above the 90th percentile) were more often linked to challenges with institutional capacity and design. These operations also had significantly lower Bank performance, outcome, and M&E ratings. While operations tackling difficult or complex challenges may take longer than average to design, these extended preparation periods can reflect more fundamental difficulties encountered during the design stage and also serve as an early warning of potential institutional capacity challenges that operations may face during implementation. Conversely, even though the *RAP 2024* analysis does not show a statistical association between shorter preparation time and project outcomes, other evaluation evidence highlights the importance of adequate preparation for successful outcomes (World Bank 2024f).

For IFC investment projects, spending sufficient time on front-end work may be particularly important in challenging contexts and projects that are especially complex. In challenging contexts (specifically, in Africa, Middle East and North Africa, FCS, and IDA and blend) and projects that are especially complex (the Infrastructure industry group), processing mostly successful or better IFC investment projects takes longer than processing mostly unsuccessful or worse projects. Delays could occur at any stage of processing between the Concept Note and first disbursement. However, cutting short key front-end work quality and preparation factors (such as market

assessment; client quality; or assumptions, financial models, and project costs) may contribute to weak development outcomes. IFC advisory services, in contrast to IFC investments, cannot accurately measure preimplementation scoping time because not all advisory services projects go through the Concept Note stage (for example, some subprojects of approved programmatic umbrellas or fast-track projects that were follow-ons from previous engagements). In addition, out of the 411 standard advisory services projects evaluated and validated by IEG (FY13–23), 67 projects (16 percent) did not have a Concept Note date. Developing an approach to measure preimplementation scoping time and recording Concept Note dates in the system would allow IFC to test associations between client responsiveness and performance indicators (such as development effectiveness, IFC work quality, and IFC role and contribution).

World Bank Group Scorecard

Challenges that affect the results monitoring of operations are relevant for the Bank Group Scorecard. The Scorecard aims to be a strategic management tool that drives action for results. Across Bank Group institutions, challenges were found with M&E quality, regular reporting, and indicators.

Many operations and country programs continue to have inadequate M&E, which could represent a challenge for implementation of the Scorecard. The share of IPF and Program-for-Results operations with M&E quality rated substantial or above increased from 29 percent in FY13 to 64 percent in FY23. However, this means that more than one-third of operations have inadequate M&E practices. World Bank operations with challenges in project data and monitoring have lower performance ratings. Recent evaluations, including *RAP 2021* and *RAP 2023*, have identified frequent challenges with indicators, data availability and baselines, and reporting and supervision (World Bank 2022c, 2023e, 2024a). Moreover, 83 percent of CLRVs reported major inadequacies in results frameworks. Evidence on issues with results frameworks is relevant to the Scorecard because management is considering ways to cascade the indicators into results frameworks across CPFs (World Bank 2024b). To improve M&E quality, consideration could be given to focusing on Global Practices that have half of their operations rated modest

and negligible over multiple years. These Global Practices may also have reporting issues on some of the new indicators.

MIGA could accurately measure its overall development outcome ratings by promptly delivering self-evaluations of all of its guarantee projects. A total of 19 MIGA projects are pending self-evaluations during the FY21–23 period (45 percent of planned self-evaluations during this period). The pending self-evaluations from MIGA might or might not sustain its development outcome ratings at a stable level. This, in turn, could affect MIGA's reporting for the new Scorecard.

IFC could improve the measurement of outcomes, particularly market outcomes, by recording more complete information about projects. IEG conducted a desk-based review of 173 IFC investment projects evaluated and validated by IEG during CY21–23. We could not verify nearly 100 outcomes. IFC introduced the AIMM system, an ex ante monitoring tool, in 2017. However, in an analysis of 21 projects with live AIMM scores (projects that were assigned ex ante AIMM scores at Board approval), IEG found that 22 percent of outcomes did not have an indicator in the tracking system (17 percent of project-level outcomes and 43 percent of market-level outcomes).[3] Moreover, most market-level outcomes were never tracked (despite having indicators) or could not be tracked (because they did not have an indicator). By identifying and tracking outcome indicators, IFC would be able to verify whether most of its outcomes, particularly market outcomes, have been achieved. This could help facilitate including these outcomes in the new Scorecard, where appropriate.

Joint Country Representation

Bank Group collaboration has for many years been a work in progress. The Bank Group has sought enhanced collaboration in country programs for almost three decades. The joint country representation initiative seeks to enable integrated solutions that span both the public and private sectors, leverage Bank Group knowledge and experience, and amplify collective impact. The joint country representation initiative could be constrained because Bank Group collaboration remains infrequent across two CPF periods and outside of a limited number of sectors.

The findings of this *RAP* on World Bank collaboration in country programs could be relevant for joint country representation. Important conditions that reinforce Bank Group collaboration are a shared view of sector priorities and objectives that are aligned across Bank Group institutions. Developing a shared view of sector priorities entails Bank Group institutions collectively understanding actors, opportunities to grow a sector, constraints on realizing those opportunities, and what needs to be done to ameliorate constraints. In addition, for collaboration to be successful, Bank Group institutions need to align their objectives. Different measures can be employed to establish these conditions. The Bank Group can use analytics and advisory services to identify areas of alignment or clarify sector priorities. A shared strategy that is more granular and flexible than a CPF can help make collaboration more effective. Measures to support close coordination between staff from different Bank Group entities, such as supporting the development of staff networks, have been factors in successful collaboration.

[1] In this chapter, we use the term *fragile and conflict-affected situations* for consistency with the rest of the report. The background documents of the evolution process use the term *fragility, conflict, and violence*.

[2] Andrews et al. (2017) highlight that building capacity in FCS is a challenging problem and does require the development of adaptive capacities of implementing agencies.

[3] According to IFC, indicators for environmental, social, and governance outcomes are recorded and monitored separately in a different system (Sustainability Rating Tool, previously Environmental and Social Review Document) other than the AIMM system and DOTS. Therefore, 10 environmental, social, and governance outcomes (8 on environment and social, 1 on greenhouse gas emissions, and 1 on improved living standards) in this analysis are considered to have indicators and are being tracked by IFC (although in a different system other than the AIMM system and DOTS).

Bibliography

Andrews, Matt, Lant Pritchett, and Michael Woolcock. 2017. *Building State Capability: Evidence, Analysis, Action*. Oxford University Press.

Denizer, Cevdet, Daniel Kaufmann, and Aart Kraay. 2013. "Good Countries or Good Projects? Macro and Micro Correlates of World Bank Project Performance." *Journal of Development Economics* 105 (November): 288–302.

ECG (Evaluation Cooperation Group). 2012. *Big Book on Evaluation Good Practice Standards*. ECG. https://www.ecgnet.org/sites/default/files/FINAL%20Big%20Book%20on%20GPS.pdf.

Eriksson, John. 2001. *The Drive to Partnership: Aid Coordination and the World Bank*. World Bank.

Geli, Patricia, Aart Kraay, and Hoveida Nobakht. 2014. "Predicting World Bank Project Outcome Ratings." Policy Research Working Paper 7001, World Bank.

OECD (Organisation for Economic Co-operation and Development). 2008. *Accra Agenda for Action*. OECD.

Ørnemark, Charlotte. 2016. "'Learning Journeys' for Adaptive Management: Where Does It Take Us." Global Partnership for Social Accountability Note 12, World Bank.

Otoo, Samuel, Natalia Agapitova, and Joy Behrens. 2009. *The Capacity Development Results Framework: A Strategic and Results-Oriented Approach to Learning for Capacity Development*. World Bank.

Ralston, Laura. 2014. "Success in Difficult Environments: A Portfolio Analysis of Fragile and Conflict-Affected States." Policy Research Working Paper 7098, World Bank.

Reed, Mark S., and Hannah Rudman. 2023. "Re-thinking Research Impact: Voice, Context and Power at the Interface of Science, Policy and Practice." *Sustainability Science* 18 (March): 967–81.

World Bank. 1984. *Toward Sustained Development in Sub-Saharan Africa: A Joint Program of Action*. World Bank.

World Bank. 2005a. *Building Effective States: Forging Engaged Societies*. World Bank.

World Bank. 2005b. *Capacity Building in Africa: An OED Evaluation of World Bank Support*. World Bank.

World Bank. 2013a. *Afghanistan Country Program Evaluation, 2002–11: Evaluation of the World Bank Group Program*. Independent Evaluation Group. World Bank.

World Bank. 2013b. *Results and Performance of the World Bank Group 2013*. Independent Evaluation Group. World Bank.

World Bank. 2014. *Tunisia—Country Program Evaluation, FY2005–13: Evaluation of the World Bank Group Program*. Independent Evaluation Group. World Bank.

World Bank. 2015a. *Kazakhstan—Country Program Evaluation, FY04–13*. Independent Evaluation Group. World Bank.

World Bank. 2015b. *Mongolia—Country Program Evaluation, FY05–13*. Independent Evaluation Group. World Bank.

World Bank. 2015c. *The Plurinational State of Bolivia: Country Program Evaluation, FY05–13*. Independent Evaluation Group. World Bank.

World Bank. 2015d. *Results Frameworks in Country Strategies—Lessons from Evaluations*. Independent Evaluation Group. World Bank.

World Bank. 2015e. *Selectivity in Country Strategies: The Evidence*. Independent Evaluation Group. World Bank.

World Bank. 2015f. *World Bank Group Engagement in Resource-Rich Developing Countries: The Cases of Bolivia, Kazakhstan, Mongolia and Zambia—Clustered Country Program Evaluation.* Synthesis Report and Management Response. Independent Evaluation Group. World Bank.

World Bank. 2015g. *Zambia—Country Program Evaluation, FY04–13.* Independent Evaluation Group. World Bank.

World Bank. 2016a. *Cluster Country Program Evaluation on Small States: Mauritius Country Case Study (FY07–15)—Enhancing Competitiveness and*

Private Sector Development. Independent Evaluation Group. World Bank. http://documents.worldbank.org/curated/en/390751475038142357/Cluster-country-program-evaluation-on-small-states-Mauritius-country-case-study-FY07-15-enhancing-competitiveness-and-private-sector-development.

World Bank. 2016b. *Cluster Country Program Evaluation on Small States: Pacific Island Countries Program Evaluation (FY05–15), Volume I.* Independent Evaluation Group. World Bank. http://documents.worldbank.org/curated/en/485891475064557163/Cluster-country-program-evaluation-on-small-states-Pacific-Island-countries-program-evaluation-FY05-15.

World Bank. 2016c. *Cluster Country Program Evaluation on Small States: Regional Program Evaluation of the Organisation of Eastern Caribbean States—Antigua and Barbuda, Dominica, Grenada, St. Kitts and Nevis, St. Lucia, and St. Vincent and the Grenadines.* Independent Evaluation Group. World Bank.

World Bank. 2016d. *Cluster Country Program Evaluation on Small States: Seychelles Country Case Study (FY07–15)—Enhancing Competitiveness and Private Sector Development*. Independent Evaluation Group. World Bank.

World Bank. 2016e. *Supporting Transformational Change for Poverty Reduction and Shared Prosperity: Lessons from World Bank Group Experience*. Independent Evaluation Group. World Bank.

World Bank. 2016f. *World Bank Group Engagement in Small States: The Cases of the OECS, PICs, Cabo Verde, Djibouti, Mauritius, and the Seychelles—Clustered Country Program Evaluation Executive Summary.* Independent Evaluation Group. World Bank.

World Bank. 2017a. "What's Next for Capacity Building in Africa." Discussion Note. World Bank.

World Bank. 2017b. *World Bank Group Joint Projects: A Review of Two Decades of Experience—Lessons and Implications from Evaluation*. Independent Evaluation Group. World Bank.

World Bank. 2018a. *CPIA Criteria 2017*. World Bank.

World Bank. 2018b. *Institutional Development Roadmap Handbook: Effective Design, Implementation and Evaluation of Project-Led Institutional Development in the Transport Sector*. World Bank.

World Bank. 2018c. *Mexico—Country Program Evaluation: An Evaluation of the World Bank Group's Support to Mexico (2008–17)*. Independent Evaluation Group. World Bank.

World Bank. 2019a. *Rwanda—Country Program Evaluation FY09–17*. Independent Evaluation Group. World Bank.

World Bank. 2019b. *The World Bank Group Partnership with the Philippines, 2009–18—Country Program Evaluation*. Independent Evaluation Group. World Bank.

World Bank. 2020a. *Renewable Energy: Evaluation of the World Bank Group's Support for Electricity Supply from Renewable Energy Resources, 2000–2017*. Independent Evaluation Group. World Bank.

World Bank. 2020b. *Results and Performance of the World Bank Group 2020*. Independent Evaluation Group. World Bank.

World Bank. 2020c. *State Your Business! An Evaluation of World Bank Group Support to the Reform of State-Owned Enterprises, FY08–18.* Independent Evaluation Group. World Bank.

World Bank. 2020d. *The World Bank Group Outcome Orientation at the Country Level*. Independent Evaluation Group. World Bank.

World Bank. 2020e. *World Bank Group Strategy for Fragility, Conflict, and Violence 2020–2025*. World Bank.

World Bank. 2021a. *Albania Country Program Evaluation.* Independent Evaluation Group. World Bank.

World Bank. 2021b. "Bank Guidance: Systematic Operations Risk-Rating Tool (SORT)." World Bank.

World Bank. 2021c. "IBRD/IFC/MIGA/IDA Guidance: Country Engagement." Guidance for Preparation of Country Engagement Products, World Bank.

World Bank. 2021d. *Managing Urban Spatial Growth: World Bank Support to Land Administration, Planning, and Development*. Independent Evaluation Group. World Bank.

World Bank. 2021e. *Results and Performance of the World Bank Group 2021*. Independent Evaluation Group. World Bank.

World Bank. 2021f. *World Bank Engagement in Situations of Conflict: An Evaluation of FY10–20 Experience*. Independent Evaluation Group. World Bank.

World Bank. 2022a. *Country Program Evaluation—Papua New Guinea: An Evaluation of World Bank Support FY08–22.* Approach Paper. Independent Evaluation Group. World Bank.

World Bank. 2022b. *CPIA Africa: Assessing Africa's Policies and Institutions*. World Bank.

World Bank. 2022c. *Results and Performance of the World Bank Group 2022*. Independent Evaluation Group. World Bank.

World Bank. 2022d. *The World Bank Group in Bangladesh, Fiscal Years 2011–20: Country Program Evaluation*. Independent Evaluation Group. World Bank.

World Bank. 2022e. *The World Bank Group in Chad, Fiscal Years 2010–20: Country Program Evaluation*. Independent Evaluation Group. World Bank.

World Bank. 2022f. *The World Bank Group in Madagascar, Fiscal Years 2007–21: Country Program Evaluation*. Independent Evaluation Group. World Bank.

World Bank. 2022g. *The World Bank Group in Ukraine, 2012–20: Country Program Evaluation*. Independent Evaluation Group. World Bank.

World Bank. 2022h. *The World Bank's Early Support to Addressing COVID-19: Health and Social Response—An Early-Stage Evaluation*. Independent Evaluation Group. World Bank.

World Bank. 2023a. *Elements That Enhance Institutional Capacity Development in World Bank Projects and Country Partnerships in Sub-Saharan Africa.* Evaluation Insight Note. Independent Evaluation Group. World Bank.

World Bank. 2023b. *Evaluation of the World Bank's Support to Procurement and Its Principles*. Approach Paper. Independent Evaluation Group. World Bank.

World Bank. 2023c. *International Finance Corporation Additionality in Middle-Income Countries*. Independent Evaluation Group. World Bank.

World Bank. 2023d. *Mid-Term Review of the World Bank Group Strategy for Fragility, Conflict, and Violence (2020–25)*. World Bank.

World Bank. 2023e. *Results and Performance of the World Bank Group 2023*. Independent Evaluation Group. World Bank.

World Bank. 2023f. *The World Bank Group in the Kyrgyz Republic: Country Program Evaluation, Fiscal Years 2014–21*. Independent Evaluation Group. World Bank.

World Bank. 2023g. *The World Bank Group in Mozambique, Fiscal Years 2008–21: Country Program Evaluation*. Independent Evaluation Group. World Bank.

World Bank. 2023h. *The World Bank Group's Engagement in Morocco, Fiscal Years 2011–21: Country Program Evaluation*. Independent Evaluation Group. World Bank.

World Bank. 2024a. *Confronting the Learning Crisis: Lessons from World Bank Support for Basic Education, 2012–22.* Independent Evaluation Group. World Bank.

World Bank. 2024b. *An Evaluation of World Bank and International Finance Corporation Engagement for Gender Equality over the Past 10 Years.* Independent Evaluation Group. World Bank.

World Bank. 2024c. "From Vision to Impact: Implementing the World Bank Group Evolution—Update to Governors." Board Report 188790, World Bank.

World Bank. 2024d. *Guidance Manual for Independent Evaluation Group Validators: Implementation Completion and Results Report Reviews for Investment Project Financing*. Independent Evaluation Group. World Bank.

World Bank. 2024e. *IDA21 Policy Package: The "Lenses" Paper*. World Bank.

World Bank. 2024f. *Making Procurement Work Better: An Evaluation of the World Bank's Procurement System*. Independent Evaluation Group. World Bank.

World Bank Group. 1992. *Effective Implementation: Key to Development Impact*. World Bank Group.

World Bank Group. 2017. *World Development Report 2017: Governance and the Law*. World Bank Group.

World Bank Group. n.d. "World Bank Group Country Opinion Survey (COS) Program." Accessed August 26, 2024. https://www.worldbank.org/en/programs/world-bank-country-opinion-surveys.

APPENDIXES

Independent Evaluation Group

Results and Performance of the World Bank Group 2024

Appendix A. Methodology

Approach and Structure

Results and Performance of the World Bank Group (RAP) answers one main question: What do Independent Evaluation Group (IEG) validations tell us about how the Bank Group's performance changed over time and across subgroups?

To answer this question, *RAP 2024* follows a structured design with two main steps of data set construction and then analysis. The three main data sets constructed relate to ratings, factors linked to performance, and outcome types. These data sets are the essential ingredients for the analysis of the *RAP*. Table A.1 provides an overview of the main data sets. In the second step, structured analysis is undertaken by drawing on single or multiple data sets. In *RAP 2024*, the data sets fed into the analysis of trends in the Bank Group's ratings over a 10-year period, challenges that may constrain performance, and levers that can be used to improve performance. This appendix outlines the methodology by institution, consistent with the chapters. It describes the construction of the data sets used for *RAP* and then (where not covered in separate appendixes) outlines the analytic techniques used to identify the findings presented in the report.

Table A.1. Overview of Types of Data Sets Compiled for *Results and Performance of the World Bank Group*

RAP Data Set	Overview Description
Ratings (compiled across all institutions)	The purpose of ratings data sets is to provide a single source of all relevant variables for IEG-rated operations. In generating this data set, combine institution-specific IEG ratings with relevant operations and country data.
Factors linked to performance (compiled across all institutions)	The purpose of the factors linked to performance data is to enable the identification of institution-specific issues linked to performance. The sources, definitions, identification, and analysis of factors are institution specific because each has a different operating model. In developing the data set, we capture the sentiment or rating of the factor and combine this with operations and country data, as relevant.

(continued)

RAP Data Set	Overview Description
Outcome type (compiled for World Bank, IFC, and MIGA)	The purpose of the outcome-type data set is to help understand the patterns of intended types of outcomes defined in the objectives of operations. In generating this data set, objectives are coded against institution-specific taxonomies. The coded outcome-type data are typically combined with relevant operations and country data.
Work quality and additionality	The aim of the analysis of association between work quality and development outcomes is to identify the top IFC work quality factors for IFC investment projects that, if addressed, could contribute to better development outcomes. The aim of the analysis of association between additionality and development outcomes is to understand the frequency and achievement rates of financial and nonfinancial additionality in IFC investment projects with weak development outcomes and weak additionality.
Factors beyond IFC influence contributing to development effectiveness for advisory services	The purpose of this analysis is to test whether factors beyond IFC's influence have played an important role in the decline of development effectiveness ratings of IFC advisory services projects in recent years (FY21–23).

Source: Independent Evaluation Group.

Note: IEG = Independent Evaluation Group; IFC = International Finance Corporation; MIGA = Multilateral Investment Guarantee Agency; *RAP* = *Results and Performance of the World Bank Group*.

The data sets defined for *RAP* are applied in different analyses, defined as follows:

» Trends. We describe overall performance trends by key ratings and important subgroups. Like a Scorecard report, *RAP 2024* presents trends, describes changes in trends, and provides a deeper analysis of specific issues based on the IEG validations of Bank Group self-evaluations.

» Challenges. Having described trends, we then identify challenges associated with performance, often in relation to subgroups. Challenges are persistent factors negatively linked to ratings. They are often drawn from institution-specific taxonomies. The definitions of challenges are institution specific because each institution responds to different kinds of challenges.

» Levers. Finally, we identify levers, which are actions within the Bank Group's influence that management can take to address challenges or improve performance.

The data sets were applied in different configurations for each of these analyses (table A.2). For the analysis of trends, only the rating data set was applied across all institutions. In considering challenges and ratings, factors linked to performance and outcome types were cross-analyzed to produce findings. To develop the findings on levers, the full range of applicable data sets was used.

Table A.2. Applications of Data Sets Compiled in *Results and Performance of the World Bank Group* Analyses

Analysis	Data Sets Used	Institutions
Trends	Ratings	World Bank, IFC, MIGA, and country programs
Challenges	Ratings Factors linked to performance	World Bank, IFC, MIGA, and country programs
	Outcome types	IFC and MIGA
Levers	Ratings Factors linked to performance	World Bank, IFC, MIGA, and country programs
	Outcome types	World Bank, IFC, and MIGA
	Work quality and additionality	IFC

Source: Independent Evaluation Group.

Note: IFC = International Finance Corporation; MIGA = Multilateral Investment Guarantee Agency.

Ensuring Validity of Findings

RAP 2024 includes several steps to guarantee a consistent approach to individual analyses. To support the replicability of the analyses, we built on protocols developed by previous *RAP* reports. For data set construction, we applied tested processes, drew on structured system data, maintained standardized rubrics, used consistent taxonomies for factors, and employed standardized coding protocols. In analysis of the data sets, we applied statistical tests appropriate to the data to indicate whether the results are unlikely to have occurred by chance, typically determined by a p value threshold (for example, $p < .05$). Where statistical tests were not possible, we sought to identify substantive trends in the data, usually over four data points. In addition, where needed, findings were cross-referenced or refined using existing

IEG studies. Finally, the limitations of each of the analyses were considered in delimiting the findings.

Methodology for the World Bank

Data Construction: Source and Coverage

Table A.3 lists data sources and sample coverage of World Bank lending projects used to construct the *RAP 2024* data sets.

Table A.3. *Results and Performance of the World Bank Group 2024* World Bank Data Sources and Coverage

Data Set	Data Sources	Coverage
World Bank performance ratings	IEG data on ICRR or Project Performance Assessment Report ratings	2,982 World Bank lending projects closed during FY13–23 and evaluated by IEG as of June 30, 2024
World Bank lending portfolio	World Bank Data Explorer	12,494 active and closed World Bank lending projects since 1986 as of May 2024
World Bank project outcome types	IEG ICRR rating data on project objectives, project data from World Bank Data Explorer, manually coded outcome types of objectives	1,336 World Bank investment project financing projects closed during FY17–23 and evaluated by IEG as of December 2024
Factors affecting operation of World Bank projects	World Bank project ICR documents, manually coded and machine learning–predicted factors	1,118 World Bank investment project financing projects closed during FY18–23 and evaluated by IEG as of December 2024
World Bank project Systematic Operations Risk-Rating Tool rating data	World Bank Data Explorer	World Bank investment project financing projects closed during FY16–23 and evaluated by IEG as of May 2024

Source: Independent Evaluation Group.

Note: ICR = Implementation Completion and Results Report; ICRR = Implementation Completion and Results Report Review; IEG = Independent Evaluation Group.

World Bank Performance Ratings

Performance ratings are assigned based on a shared approach defined between management and IEG. The Implementation Completion and Results Reports (ICRs) prepared by World Bank staff are essential self-evaluation tools to assess projects and operations. As part of its validation work, IEG conducts independent reviews of the ICRs, known as Implementation Completion and Results Report Reviews (ICRRs). These reviews critically validate the evidence, results, and ratings of the ICRs, aligning them with the project's design documents as necessary. IEG also conducts fieldwork to learn about selected World Bank lending operations annually through Project Performance Assessment Reports. Project performance ratings for World Bank projects are derived from an objective-based methodology that, together with performance rating scales and criteria, was agreed on with Operations Policy and Country Services. Ratings are rubrics for assessing performance relative to a project's or program's objectives. Ratings divide the World Bank's self-evaluation and IEG's validation narratives into categories or values that enable aggregation across operations. The ratings used by ICRRs and Project Performance Assessment Reports are listed in table A.4.

The data set covers 2,983 World Bank lending projects that closed during FY 2013–23 and were evaluated by IEG through either an ICRR or a Project Performance Assessment Report as of June 30, 2024. At this cutoff date, IEG's pipeline had 135 ICRRs for all lending types, less than 10 percent of which had been in the pipeline for 180 days or longer. Figure A.1 shows that ICRRs are still pending for a number of projects, particularly those closing in recent fiscal years, with FY23 being notably underrepresented. More ICRs of recently closed projects could arrive in IEG later. Consequently, *RAP* has an inherent selection bias in its coverage. Projects with ICRs and ICRRs completed relatively quickly after closure tend to have higher ratings than those with delayed evaluations. This pattern has been consistently observed in previous *RAPs*. Each *RAP's* successive update to rating trends typically results in a modest downward adjustment of ratings of the latest fiscal years relative to preceding *RAP* reports. Besides IEG rating data, the data set also included data on project characteristics, such as milestone dates and project volume, and country characteristics, including whether countries were

classified as fragile and conflict-affected situations (FCS), their income levels, and lending groups.

Table A.4. *Results and Performance of the World Bank Group 2024*
Independent Evaluation Group Project Performance Ratings

Rating	Definition	Scale
Outcome	The extent to which a project efficiently achieved, or was expected to achieve, its relevant objectives. The outcome rating integrates three underlying dimensions: relevance, efficacy (objectives achievement), and efficiency.	Six-point: » Highly satisfactory » Satisfactory » Moderately satisfactory » Moderately unsatisfactory » Unsatisfactory » Highly unsatisfactory
Relevance	The extent to which the project's objectives align with current World Bank country strategies at the time of project closing.	Four-point: » High » Substantial » Modest » Negligible
Efficacy	The extent to which the project's objectives were achieved, or are expected to be achieved, taking into account their relative importance and attributability to the activities or actions supported by the project. It is rated for each individual objective and at the overall level.	Four-point: » High » Substantial » Modest » Negligible
Efficiency	A measure of how economic resources and inputs are converted to results. It assesses whether the costs involved in achieving project objectives were reasonable in comparison with the benefits and recognized norms (value for money).	Four-point: » High » Substantial » Modest » Negligible
Bank performance	The extent to which World Bank services ensured quality project design and supported effective implementation through appropriate supervision to achieve development outcomes.	Six-point: » Highly satisfactory » Satisfactory » Moderately satisfactory » Moderately unsatisfactory » Unsatisfactory » Highly unsatisfactory

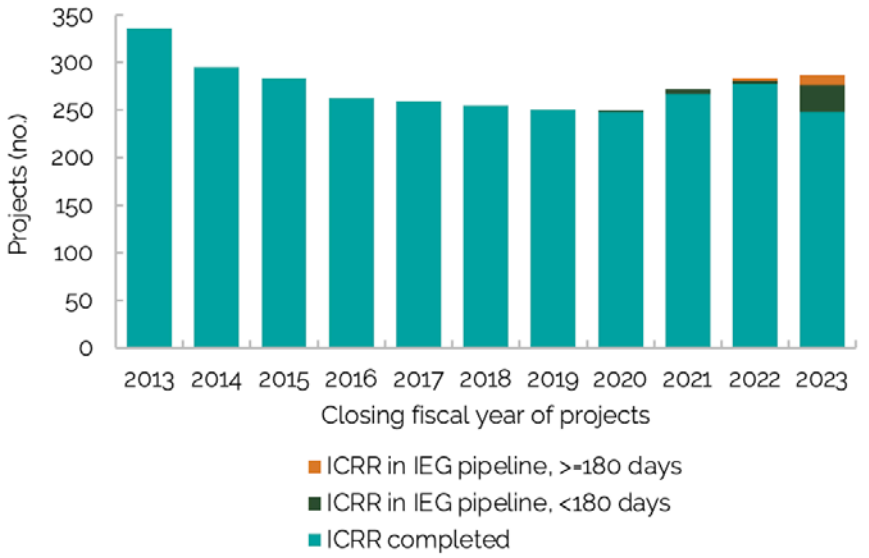
(continued)

Rating	Definition	Scale
Quality at entry Design	The extent to which the World Bank identified, facilitated preparation of, and appraised the project to maximize the likelihood of achieving planned development outcomes while maintaining consistency with the World Bank's fiduciary role. (The evaluation methodology for development policy financing projects changed in 2020, redefining this rating as "design.")	Six-point: » Highly satisfactory » Satisfactory » Moderately satisfactory » Moderately unsatisfactory » Unsatisfactory » Highly unsatisfactory
Quality of supervision Implementation	The extent to which the World Bank proactively identified and addressed threats to the achievement of relevant development outcomes and the World Bank's fiduciary role. (The evaluation methodology for development policy financing projects changed in 2020, redefining this rating as "implementation.")	Six-point: » Highly satisfactory » Satisfactory » Moderately satisfactory » Moderately unsatisfactory » Unsatisfactory » Highly unsatisfactory
M&E quality	The quality of the project's M&E design, implementation, and use of results to improve performance. (The evaluation methodology for development policy financing projects changed in 2020, eliminating this rating for development policy financing projects.)	Four-point: » High » Substantial » Modest » Negligible

Source: Independent Evaluation Group Implementation Completion and Results Report Review evaluation guidelines.

Note: M&E = monitoring and evaluation.

Figure A.1. Implementation Completion and Results Report Reviews Completed and in Pipeline



Source: Independent Evaluation Group.

Note: ICRR = Implementation Completion and Results Report Review; IEG = Independent Evaluation Group.

World Bank Lending Portfolio

RAP 2024 constructed a data set of 12,494 World Bank lending operations (as of May 2024) that were closed or active since 1986. To illustrate portfolio size over time, the original project-level data set was transformed into a longitudinal format. Each project was expanded into multiple entries, with each entry representing a fiscal year in the project's life cycle. For closed projects, this span covers the period from the approval fiscal year to the closing fiscal year. For active projects, this period extends from the approval fiscal year to the current fiscal year. This transformation resulted in a data set where each row represented a unique project-year combination.

To address missing values, the approval fiscal year was determined using a sequence of available dates: approval date, concept review date, and Activity Initiation Summary sign-off date, in that order. Similarly, for the closing fiscal year, the sequence included closing date, deactivation date, loan closing date, and note of cancellation date. Projects lacking an approval fiscal year,

and closed projects without a closing fiscal year, were excluded from the portfolio. The data set was supplemented with additional country-level data for each fiscal year, including income level, lending group, and FCS status.

World Bank Project Outcome Types

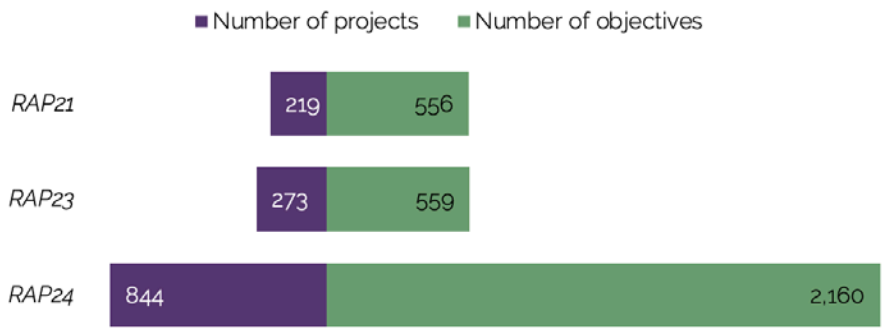
The data set covers all 1,336 World Bank investment project financing projects closed during FY17–23 and evaluated by IEG as of December 2024. To categorize the outcomes of these projects, *RAP 2024* employed the outcome typology developed for *RAP 2021* (World Bank 2021b). This typology includes 16 different outcome types, derived from typical project theories of change and select corporate objectives. Table A.5 provides a list of these outcome types. The *RAP 2024* data set incorporates outcome classifications from previous reports (*RAP 2021* and *RAP 2023*) and new coding done for *RAP 2024*. Figure A.2 shows the distribution of objectives coded across the three *RAPs*.

Table A.5. Project Outcome Typology

Outcome Types
1. Access to services expanded
2. Quality of services improved
3. Public assets improved
4. Natural capital sustained
5. Use of services of assets increased
6. Temporary relief to individuals provided
7. Awareness, attitudes, or behaviors changed
8. Human capital increased
9. Individual employability or livelihood improved
10. Citizen engagement or community participation enhanced
11. Legal or regulatory context improved
12. Capacity of institutions to perform institutional functions enhanced
13. Accountability, transparency, or governance enhanced
14. Enterprise or sectoral performance improved
15. Productive sector expanded
16. Equity or inclusion enhanced

Source: Independent Evaluation Group.

Figure A.2. Coding of Outcome Types



Source: Independent Evaluation Group.

Note: RAP = *Results and Performance of the World Bank Group*.

In constructing the *RAP 2024* portion of the data set, the coding was based on project objectives. The text for coding was extracted from ICRR documents, including statements of objectives, the theories of change underlying the objectives, and measurements of achieving the objectives. The coding was conducted at the objective level, with a maximum of three outcome types assigned per objective. This objective-level coding was then aggregated to the project level, resulting in projects being tagged with between one and six outcome types. The coding team comprised 10 expert evaluators working in pairs. Each pair was assigned to projects based on their area of evaluation expertise, with one coder and one reviewer per project to ensure consistent and specialized assessment.

Factors Linked to Performance of World Bank Operations

RAP 2024 builds on previous *RAP* reports by continuing to use an adapted version of the DeCODE (Delivery Challenges in Operations for Development Effectiveness) taxonomy. Developed by the World Bank's Global Delivery Initiative in 2016, DeCODE identifies typical challenges that may affect operational performance from design to closure. The taxonomy's validity is ensured through an iterative process involving literature reviews, text analytics, and practitioner consultations. For *RAP 2024*, the taxonomy comprises three main clusters—context, institutional capacity of stakeholders, and project—which are further divided into 12 categories and their respective

subcategories if applicable. Both categories and subcategories largely retain the preexisting definitions originally developed by DeCODE (table A.6). Notably, the cluster name *institutional capacity of stakeholders* is adapted from the original term *stakeholder*. The World Bank's influence varies across these clusters: it has limited influence on context, indirect influence on institutional capacity of stakeholders, and direct influence on project-related factors.

Table A.6. *Results and Performance of the World Bank Group 2024*
Taxonomy of Factors Affecting Project Implementation

Cluster	Category	Subcategory
Context	Legislation and regulations: delivery challenges stemming from an unsupportive legal environment caused by lack of appropriate legal or regulatory framework, inordinate delays in promulgating laws, or complicated and time-consuming regulatory processes.	
	Governance and politics: delivery challenges faced because of elections, opaque governance environment characterized by inadequate accountability, weak rule of law, political manipulation of projects, or corruption.	Political interference: challenges stemming from the abuse of public power for private gain and favoritism toward patrons, clients, and associates. Electoral cycles: challenges caused by elections and electoral processes.
	Conflict and instability: delivery challenges faced because of disruptions stemming from a conflict or postconflict situation, insecurity, or civil unrest.	
	Disasters and emergency response: delivery challenges caused by natural or manufactured disasters or other unexpected emergency situations.	Natural disasters: challenges stemming from natural disasters. Epidemics: challenges stemming from disruptions caused by epidemics.
	Business environment: delivery challenges caused by a weak private sector or weak sector regulations.	
	Macroeconomic environment: delivery challenges caused by instability, volatility, or interruptions in trade, market conditions, or financial systems.	

(continued)

Cluster	Category	Subcategory
Institutional capacity of stakeholders	Coordination and engagement: delivery challenges stemming from difficulty in coordination and engagement among stakeholders due to issues of administrative and bureaucratic structure, unclear definition of roles, or inadequate engagement and communication strategies.	
	Commitment and leadership: delivery challenges stemming from a change in leadership, shifts in priorities, or the absence of shared commitment and consensus among stakeholders.	
	Human resources and organizational capacity: delivery challenges faced because of constraints caused by lack of skilled human resources, difficulties in acquiring necessary skills, or limited organizational capacity.	
Project	Operational design: delivery challenges stemming from flaws in project design, including overly complicated design, overambitious objectives, inappropriate time allocation, or issues in identifying and selecting and targeting stakeholders and beneficiaries.	Appropriate objectives or project design: challenges caused by setting targets that are unrealistically ambitious or making the project design overly complex. Time allocation or task sequencing: challenges related to insufficient or excessive duration of a component or inappropriate timing and sequence of task. Stakeholder selection: challenges caused by problems identifying and selecting appropriate stakeholders to engage. Beneficiary targeting: challenges with ensuring that the appropriate beneficiary group is targeted.

(continued)

Cluster	Category	Subcategory
	Project finance: delivery challenges related to procurement or fiduciary arrangements such as planning and budgeting, financing mechanisms, financial reporting, and auditing.	Procurement: challenges caused by issues with procurement management systems, including ineffective contract management and delays. Financing mechanism: challenges related to the choice of financing mechanism or instrument. Budgeting: challenges related to insufficient or inappropriate budget allocation or caused by complex budget processes and management. Financial management and reporting: challenges related to disbursement, financial control, and financial reporting.
	Project data and monitoring: delivery challenges caused by ineffective monitoring and evaluation because of inadequate data collection and management, lack of or inappropriate indicators, or inadequate project supervision.	Indicators: challenges caused by lack of realistic indicators, duplication or overlapping indicators, or poorly designed indicators that are misaligned with project objectives. Data availability and baselines: challenges that stem from a lack of current or accurate data and inability to produce baselines. Reporting and supervision: challenges caused by obstacles in capturing relevant information and reporting it in a timely fashion.

Source: Global Delivery Initiative.

Note: The original DeCODE (Development Challenges in Operations for Development Effectiveness) taxonomy has 15 categories. Three were dropped (social and cultural, environment and geography, and basic infrastructure) because too few examples of these were identified in developing the training data for the machine learning model.

The data set for *RAP 2024* encompasses 1,118 World Bank investment project financing projects closed during FY18–23 and evaluated by IEG as of December 2024. Coding is based on the text from the self-reported ICR section titled Factors Affecting Implementation and Performance. This data set builds on the foundation laid by *RAP 2023*, which combined manual coding and machine learning techniques. Detailed information on the supervised machine learning models can be found in *RAP 2023* appendix A, "Methodological Approach" (World Bank 2023). To cover projects not included in the *RAP 2023* data, *RAP 2024* employed the machine learning model developed for *RAP 2023* for category tagging, along with the pretrained SiEBERT model for English-language sentiment classification. In this context, sentiment refers to how the ICR language characterizes specific factors as positively or negatively affecting project implementation. It is important to acknowledge the limitations of this approach, including the 75 percent accuracy rate of the *RAP 2023* machine learning model and the reliance on self-reported ICR narratives, which are subject to variations across different authors.

As an elaboration on the analysis of factors linked to performance, *RAP 2024* further developed a data set to analyze the World Bank's Systematic Operations Risk-Rating Tool (SORT). SORT is the World Bank's tool for operations to assess development outcome risks. It identifies specific inherent risks to the operation's development outcomes, reviews ongoing and planned mitigation measures, and assesses the residual risk by considering the likelihood of the risk materializing and the impact on development outcomes given the mitigation measures. Risk assessment via SORT is an ongoing and dynamic process throughout the lifetime of an operation. The SORT rating uses a four-point scale (high, substantial, moderate, and low), covering the following eight categories in addition to overall rating:

1. Political and governance
2. Macroeconomic
3. Sector strategies and policies
4. Technical design of project or program
5. Institutional capacity for implementation and sustainability

6. Fiduciary (financial management and procurement)
7. Environment and social
8. Stakeholders

The *RAP 2024* data include the initial and end SORT ratings of World Bank investment project financing projects closed during FY16–23 and evaluated by IEG as of May 2024.

Analysis

Trends

The trend analysis covered projects closed between FY13 and FY23, focusing on project ratings on outcome, Bank performance, quality at entry, quality of supervision, and monitoring and evaluation (M&E) quality. The trends were assessed both annually and using a three-year rolling average. The analysis examined the percentage of projects rated moderately satisfactory or above (substantial or above for M&E quality) and the average rating. In addition to overall trends, the study explored ratings across various subgroups, including Region, Practice Group, Global Practice, FCS exposure, project size, project duration, and COVID-19 exposure.

To assess shifts in ratings, decomposition analysis was used to break down the influence of changes within subgroups and shifts in portfolio composition across these subgroups. To determine whether differences in ratings between periods were statistically significant, we conducted the Mann–Whitney *U* test, both at the overall level and across subgroups. In addition, the Wilcoxon signed-rank test was used to identify statistically significant gaps between project outcome ratings and Bank performance ratings for each fiscal year of project closure.

To complement our trend analysis and provide context for the shifting composition of the portfolio, we examined the evolution of the World Bank's overall lending portfolio during FY13–24. This examination included an in-depth look at portfolio composition in terms of Regions, Practice Groups, Global Practices, and FCS content, offering insights into the changing focus of the World Bank's lending activities.

Challenges

To examine challenges encountered during project design and implementation, our analysis focused on 12 factor types across three clusters: context, institutional capacity of stakeholders, and project-specific issues. We evaluated 1,118 World Bank investment project financing operations closed between FY18–23, assessing both the prevalence of these factors in ICR and the proportion of projects that identified them as challenges. The analysis was conducted both at the overall portfolio level and across subgroups, such as by Region and FCS exposure.

To gain deeper insights into specific challenges (such as human resources, organizational capacity, coordination, engagement, leadership, project design, data monitoring, and financial management), the analysis applied text summarization techniques assisted by artificial intelligence tools. Initially, we uploaded the entire corpus of text from the Key Factors That Affected Implementation and Outcome section of ICRs used for coding the 12 factors to advanced language models such as GPT-4 and Claude 3. These artificial intelligence tools were then instructed to search for and extract all text relevant to specific challenges and to provide a summary of key points. The first round of output provided a broad range of insights. Following this initial extraction, we conducted a thorough review of the artificial intelligence–generated output by selecting or merging key points based on their relevance and significance. A second round of summarization by artificial intelligence tools was then conducted, focusing on the refined set of key points, which provided more detailed and targeted insights.

Additionally, to complement the analysis of design challenges, we examined project preparation time. Preparation time was defined as the number of days between the Activity Initiation Summary and the project approval date. The distribution of preparation time skewed highly to the right, with the 90th percentile exceeding 1,393 days. The relationship between prolonged preparation times and project challenges was further analyzed. Both chi-square tests and logistic regression revealed significant correlations between extremely long preparation times and challenges related to institutional capacity and project design.

To investigate the impact of extended preparation periods on project performance, the analysis analyzed the relationship between lengthy preparation times and various project ratings, including outcome, Bank performance, and M&E quality. We employed both Mann–Whitney *U* tests and ordinal logistic regression models, both of which revealed statistically significant correlations across all rating types. Furthermore, we examined whether the presence or absence of specific factors, and their identification as challenges, corresponded with differences in outcome and Bank performance ratings. These analyses also used Mann–Whitney *U* tests and ordinal logistic regression. To further explore the nuanced relationship between challenges and project outcomes, we employed Spearman rank correlation and locally weighted scatterplot smoothing (LOWESS) techniques. These methods allowed us to assess how the number of challenges within each cluster correlated with outcome ratings, providing insights into the potential nonlinear nature of these relationships.

Levers

The lever analysis used a multifaceted approach, using data from the World Bank's SORT ratings and from ICRs. We examined average SORT ratings across all categories, comparing overall and specific category ratings among various subgroups, including Regions and FCS exposure. To explore relationships between challenges, risk identification and mitigation, and project outcome, we used Spearman rank correlation analyses. These correlations were performed between identified challenges and end ratings of all SORT categories and between project outcome ratings and SORT ratings, including both end ratings and reductions during project life cycles.

Our method also included a sentiment analysis of text related to risk identification, mitigation, and adaptive management from the Key Factors That Affected Implementation and Outcome section of 1,118 ICRs. We then correlated the sentiment (positive or not) with project outcome ratings. To better understand risk mitigation and adaptive management strategies, we used a hybrid approach that combined manual review with artificial intelligence–assisted summarization techniques. This process helped us identify common types of risk identification, mitigation strategies, and adaptive management approaches.

Additionally, we analyzed outcome types across various subgroups, examining how different outcome types correlated with project ratings in different contexts. This analysis aimed to understand how a project objective's nature influences performance across diverse operational environments.

Methodology for the International Finance Corporation

Independent Evaluation Group Evaluation Methodology for International Finance Corporation Investment Projects

IEG draws a random stratified representative sample annually from among International Finance Corporation (IFC) investment projects approved by the Board of Executive Directors five years earlier that have reached early operating maturity. During the calendar year, IFC investment staff evaluate all active IFC investment projects selected in the sample using Expanded Project Supervision Reports (XPSRs), and IEG independently validates them using Evaluative Notes (EvNotes). For closed projects selected in the sample, IEG prepares a Project Evaluation Summary in lieu of an XPSR. To conduct the project evaluation and validation, IFC and IEG staff refer to XPSR guidelines, which provide the evaluation framework and performance rating criteria.

The evaluation system and performance ratings for IFC investment projects are both based on objectives and benchmarks. In addition to focusing on the achievement of expected objectives stated in the Board report at approval, IFC investment project performance is assessed against several benchmarks (such as performance of peer companies, the market, and similar industries) and considers unintended outcomes (both positive and negative).

The main performance assessment dimensions for IFC investment projects are development outcome, IFC additionality, IFC investment outcome, and IFC work quality. In addition, the XPSR assesses the sustainability of development and IFC investment outcomes in the longer term by examining project prospects and investment return expectations over the remaining life of the project.

» Development outcome synthesizes a project's performance across four dimensions: project business performance, economic sustainability,

environmental and social effects, and private sector development. It is rated on a six-point scale: highly successful, successful, mostly successful, mostly unsuccessful, unsuccessful, and highly unsuccessful.

- IFC additionality assesses the benefit or value addition IFC brings that a client would not otherwise have. It is rated on a four-point scale: excellent, satisfactory, partly unsatisfactory, and unsatisfactory.
- IFC investment outcome assesses the extent to which IFC has realized at the time of evaluation and expects to realize over the remaining life of the investment the loan income, equity returns, or both that were expected at approval. It is rated on a four-point scale: excellent, satisfactory, partly unsatisfactory, and unsatisfactory.
- IFC work quality assesses IFC's operational performance, including in relation to environmental and social aspects, with respect to precommitment work in (i) screening, appraisal, and structuring, and (ii) its supervision and administration after project approval by the Board and subsequent commitment. It is rated on a four-point scale: excellent, satisfactory, partly unsatisfactory, and unsatisfactory.

Independent Evaluation Group Evaluation Methodology for International Finance Corporation Advisory Projects

For all client and sponsor development projects and sector development and market creation advisory projects, the IFC advisory services operations staff conduct an evaluation at completion in the form of the Project Completion Report (PCR). IEG validates a random stratified representative sample of these reports each year through EvNotes. IEG annually draws a random stratified representative sample from among projects with PCRs prepared in the previous fiscal year. Both IFC and IEG staff refer to PCR guidelines when preparing these documents, which provide evaluation frameworks and performance rating criteria. The performance ratings for IFC advisory projects are derived from an objectives-based methodology that establishes minimum thresholds for rating and assessing project effectiveness.

The main performance assessment dimensions for IFC advisory projects are development effectiveness, IFC role and contribution, and IFC work

quality. As part of development effectiveness performance, PCRs assess the sustainability of results over the long term by examining a project's impact achievement beyond the immediate and intermediate outcome achievements.

- Development effectiveness synthesizes a project's performance across five indicators: strategic relevance, output achievement, outcome achievement, impact achievement, and efficiency. It is rated on a six-point scale: highly successful, successful, mostly successful, mostly unsuccessful, unsuccessful, and highly unsuccessful.
- IFC role and contribution assesses the extent to which IFC added value or made a special contribution to the advisory project. It is rated on a four-point scale: excellent, satisfactory, partly unsatisfactory, and unsatisfactory.
- IFC work quality assesses the extent to which services provided ensured quality at entry and supported effective implementation, through appropriate supervision and execution, toward the achievement of development objectives. IFC work quality and its two dimensions—project preparation and design and project implementation and supervision—are rated on a four-point scale: excellent, satisfactory, partly unsatisfactory, and unsatisfactory.

Analysis, Data Sources, and Sample Coverage

Table A.7 lists data sources and sample coverage of IFC investment and advisory projects used in the *RAP 2024* analyses.

Table A.7. Data Sources and Sample Coverage of International Finance Corporation Investment and Advisory Services Projects

Analysis	Data Source	Sample Coverage
IFC project performance ratings	IEG and IFC data	IFC investment projects in CY13–15 (long term), CY16–18 (medium term), and CY21–23 (recent); XPSR programs and IFC advisory projects in FY13–15 (long term), FY16–18 (medium term), and FY21–23 (recent); PCR programs validated as of June 30, 2024
Association between work quality and development outcomes	IEG and IFC data, XPSR Evaluative Notes	19 IFC investment projects in CY21–23 XPSR evaluation programs evaluated and validated as of December 31, 2023, that were rated highly unsuccessful on development outcomes or unsatisfactory on work quality

(continued)

Analysis	Data Source	Sample Coverage
Analysis of factors beyond IFC's influence contributing to weak development effectiveness	IEG and IFC data, PCR Evaluative Notes	35 IFC advisory services projects in FY21–23 PCR evaluation programs validated as of June 30, 2024, that were rated mostly unsuccessful or worse on development effectiveness and satisfactory or better on work quality
Association between additionality and development outcomes	IEG and IFC data, XPSR Evaluative Notes	18 IFC investment projects in CY21–23 XPSR evaluation programs evaluated and validated as of December 31, 2023, that were rated highly unsuccessful on development outcomes or unsatisfactory on additionality
Factors linked to development outcomes of IFC projects	IEG data and taxonomy, XPSR Evaluative Notes, IFC data	256 IFC investment projects in CY20–23 XPSR evaluation programs evaluated and validated as of December 31, 2023
Analysis of IFC investment project outcome types	IEG and IFC data, XPSR Evaluative Notes	173 IFC investment projects in CY21–23 XPSR evaluation programs evaluated and validated as of December 31, 2023
Analysis on identification and tracking of outcomes in the AIMM system	IEG and IFC data, XPSR Evaluative Notes, Board papers, AIMM Navigator, Development Outcome Tracking System	21 IFC investment projects with "live" AIMM scores in CY21–23 XPSR evaluation programs evaluated and validated as of December 31, 2023 (This is the complete universe of evaluated projects with live AIMM scores because these were the only projects that were assigned ex ante AIMM scores at Board approval and evaluated or validated by IEG as of December 31, 2023.)
Association between client responsiveness and performance	IEG and IFC data, IFC business metadata, management information systems, and iPortal	870 IFC investment projects in CY13–21 XPSR evaluation programs evaluated and validated as of June 30, 2024; 793 IFC advisory services projects in FY13–23 PCR evaluation programs validated as of June 30, 2024

Source: Independent Evaluation Group.

Note: AIMM = Anticipated Impact Measurement and Monitoring; CY = calendar year; IEG = Independent Evaluation Group; IFC = International Finance Corporation; PCR = Project Completion Report; XPSR = Expanded Project Supervision Report.

Sample Selection and Representativeness of International Finance Corporation Investment Projects

The XPSR system is based on a sampling of IFC investment projects that were approved five years earlier. The combined sample of calendar year

(CY)21–23 XPSR projects was drawn from the net approval population (NAP) of projects approved in CY16–18.

IEG selected the stratified random representative sample from investment projects meeting the early operating maturity criteria that had the best fit in terms of representing the population characteristics. In addition to active investment projects, the sample included closed investment projects to represent all mature projects. The overall XPSR sample size was determined to achieve representativeness of the population on a three-year rolling basis, with a sampling error of 5 percent or less at the 95 percent confidence level.

There were in total 223 projects in the combined CY21–23 XPSR programs chosen from a CY16–18 population of 538 projects (sampling rate of 41 percent). A principal goal of sampling is to achieve representativeness, which supports valid performance inferences about the population. Matching of the sample against the population was based on the number of investments. The sampling table compares the characteristics (such as IFC commitment, investment size, lending instrument, industry group, region, International Development Association [IDA] status, environmental category, indicative performance, and project status) of the combined sample of 223 CY21–23 XPSRs with those of 538 investment operations in the CY16–18 NAP. There was good fit between the sample and the population and no performance bias.

Sample Selection and Representativeness of International Finance Corporation Advisory Projects

At implementation completion, IFC prepares the PCR for all client and sponsor development projects and sector development and market creation advisory projects. Each year, IEG validates a random, stratified, representative sample of projects with PCRs prepared in the previous fiscal year. The coverage rate is determined to be sufficient to allow for statistical inference about (development effectiveness) success rates in the population and to achieve representativeness on a three-year rolling basis with a sampling error of 5 percent or less at the 95 percent confidence level. The stratified random sample has the best fit in terms of representing the population characteristics.

There were 176 PCRs in the combined FY21–23 samples, chosen from a population of 317 projects (sampling rate of 56 percent). As with XPSR sampling,

the principal goal of PCR sampling has been representativeness to support valid performance inferences about the population. The sampling table compares the characteristics (total funds managed by IFC, funding size, project duration, country borrower type, country FCS status, project type, primary business area, and region) of the combined sample of 176 FY21–23 PCRs with those of 317 advisory operations in the FY21–23 NAP. Overall, there was generally close alignment of characteristics between the sample and the NAP.

Factors Linked to Development Outcomes of International Finance Corporation Investment Projects

To identify the factors influencing IFC investment project implementation and performance, the *RAP 2024* team performed a qualitative review and content analysis of project evaluation documents. This included 256 IFC investment projects in CY20–23 for which the evaluation and validation were completed by the cutoff date of December 31, 2023. For each project, the *RAP* team identified the top three factors that positively or negatively affected project performance and classified them using the existing taxonomy of performance factors, consisting of 5 categories and 51 subcategories developed by IEG (table A.8).

The taxonomy used for this exercise was based on common challenges and issues faced in more than 1,000 evaluated IFC investment projects. For these projects, IEG had used machine learning in addition to human thinking to identify key performance factors and classify categories and subcategories. This machine learning model has been fully tested for IFC's Financial Institutions Group investment projects; has been partially tested for IFC's Infrastructure and Natural Resources industry group and its Manufacturing, Agribusiness, and Services investment projects; and is in the process of testing for Disruptive Technologies and Funds investment projects.

This *RAP* contributed to further training for the machine learning model with the data collected manually for the CY20–23 projects evaluated and validated by IEG. The current accuracy rate of the machine learning model is 69 percent, which means that the model is identifying the same top performance factors as the data collected manually in 69 percent of projects.

For CY20–23 projects, the factor identification and classification exercise involved two steps. First, the *RAP 2024* team conducted its factor analysis by identifying the top three factors for each project based on a review of project evaluation documents and classified them according to the taxonomy. Second, for each reviewed project, the Financial and Private Sector Micro Unit sector leaders reviewed and validated these key factors and their categories and subcategories. The second step ensured not only appropriate classification of categories and subcategories but also correct identification of factors that contributed to project performance. An additional review across industries made sure that classifications were consistent over the total portfolio of EvNotes analyzed.

Using the collected data, the *RAP* team analyzed the prevalence of key factors that contributed to some projects performing better or worse than others. The team analyzed the similarities and differences of the main factors across challenging contexts (for example, for IDA and blend and for IDA and blend in FCS) and for Latin America and the Caribbean.

Table A.8. Taxonomy of Factors Linked to Development Outcomes

No.	Subcategory	Definition
Country factors		
1	Civil unrest	Risk factors related to civil unrest and armed conflict
2	Epidemics/ COVID-19	Risk factors related to epidemics (human, animal, and plants)
3	Natural disasters	Risk factors related to natural disasters
4	Economic issues	Risk factors related to the macroeconomic environment, inflation, monetary policy, or austerity measures
5	Legal risk/ regulatory	Risk factors related to regulatory policies, government, legislation, and bureaucratic mechanisms
6	Political risk	Risks factors related to the political environment, including legislative and electoral dynamics
7	FX/local currency/ devaluation	Risk factors related to currency fluctuation, depreciation, devaluation, and other exchange risks
8	Expropriation/ nationalization/ transferability	Risk factors related to expropriation, nationalization, transfer, and convertibility

(continued)

No.	Subcategory	Definition
Market/sector/industry factors		
9	Business risk	Risk factors related to business model, cyclical business, or the operating environment
10	Competition	Issues of market competition: barriers to entry, monopolies, market dominance, and penetration
11	Customers	Challenges related to identifying correct target markets and clientele
12	Market share	Issues of market share
13	Pricing	Issues of price elasticity, supply, and marginal gains
Sponsor/client (management, sponsorship, and leadership)		
14	Ability/technical expertise/track record	Issues related to the quality and expertise of the management team and their technical skills, track record, contractor competency, familiarity, and acumen
15	Commitment/motivation	The strength and valence of strategic alignment, including issues of compatibility, motivation, and ownership
16	Conflicts of interest/corporate governance	Issues related to conflict of interest and corporate governance
17	Integrity/transparency/fairness/reputation	Issues of integrity and transparency, for example, in disclosures of sensitive ethical issues, irregularities, and negative public perceptions
18	Capacity/capitalization/leverage	Issues related to sponsor capacity, capitalization, and leverage
19	Organizational structure/operational risk	Issues related to organizational culture, institutional procedures, policies, and accountability
20	Succession/family-owned business/key-person risk	Issues related to succession, family-owned businesses, and key-person risk
Project inherent challenges		
21	Environment and sustainability	Factors related to environmental standards, social health and safety parameters, or other safety standards
22	Liquidity	Issues pertaining to liquidity
23	Asset quality	Issues pertaining to asset quality
24	Expansion	Issues pertaining to acquisition, modernization, and aggressive expansion
25	Greenfield	Issues related to greenfield projects
26	Gender	Issues related to gender
27	Earnings, profitability	Issues related to earnings, profitability
28	Cost overruns/construction delay	Issues regarding cost overruns or construction delays

(continued)

No.	Subcategory	Definition
29	Funding	Issues related to funding
30	Technology	Changes in technology affecting project performance
31	Project size	Issues pertaining to asset project size
32	Training and know-how/ implementation	Issues with training and know-how
Controllable by IFC		
33	Debt issues	Project design parameters related to debt issues: syndication, repayment, security, and refinancing
34	Equity issues	Issues related to equity, valuation, and shareholder rights
35	Financial risk mitigation	Risk mitigation mechanisms such as guarantees, securities, prepayment penalties, and restructuring mechanisms
36	Prepayments	Project design parameters related to prepayments
37	Subordinated debt/ quasi-equity issues	Issues related to quasi-equity such as conversions
38	Additionality principle/catalytic role	Issues related to additionality, demonstration effects, and added value
39	Collaboration/ coordination within IFC (for example, between AS and IS)	Use of advisory services to enhance IFC roles and contributions
40	Coordination and collaboration with World Bank, other DFIs, donors, and other external stakeholders	Issues related to combined partnership, collaboration with the World Bank, other DFIs, and external stakeholders
41	Assumptions/financial models/project cost	Financial modeling assumptions, including issues regarding overambitious objectives, deviations from forecasting estimates, and scaling
42	Market assessment	Issues related with market assessment, market analysis, and consumer preferences
43	Sensitivity analysis	Sensitivity analysis, worst-case scenario, stress tests
44	M&E	Issues related to compliance monitoring, including measurement, reporting, auditing, and work quality
45	Documentation	Issues pertaining to the quality of monitoring, documentation, and reporting
46	Supervision and reporting	Issues pertaining to supervision and administration
47	Relationship management	Issues pertaining to the quality and scope of relationship management, including fruitful and proactive engagements with on-site staff

(continued)

No.	Subcategory	Definition
48	Project design	Issues pertaining to project design (Was the project design tailored to client capacity and country or market conditions? Was adequate implementation support provided?)
49	Resources and timeline	Staffing, budget, timeline: Project was not adequately resourced (funding, staffing) with a realistic timeline. Staff turnover, lack of experience and knowledge; inadequate supervision of consultants. Staffing: Assess the extent to which (i) the project was adequately staffed and possessed sufficient set of skills and expertise to effectively manage the project work and (ii) staff turnover and transition arrangements were properly handled. Include implementing partner capacity and motivation.
50	Loan issues	Issues related to loan agreements, operating policies, breaches, or technical defaults
51	Other issues	Factors related to other issues

Source: Independent Evaluation Group.

Note: AS = advisory services; DFI = development finance institution; FX = foreign exchange; IFC = International Finance Corporation; IS = investment services; M&E = monitoring and evaluation.

Outcome-Type Analysis for International Finance Corporation Investment Projects

RAP 2021 developed a 13-category typology of intended outcomes that leveraged IFC's Anticipated Impact Measurement and Monitoring (AIMM) system. These intended outcomes were aligned with those defined by the AIMM sector frameworks. AIMM sector frameworks have been developed for more than 20 key sectors and subsectors (for example, small and medium enterprise finance, manufacturing, power, and private equity funds) of IFC's investment operations. The AIMM system identifies key development outcomes (defined as outcome claims), with specific indicators for each investment project, in accordance with the theory of change defined in each AIMM sector framework. Each sector framework identifies an expected theory of change that indicates how the projects in each relevant sector are expected to address development gaps. This is done by demonstrating typical outcomes to be achieved by each project at both the project and market levels. Each sector framework also includes a list of standard indicators and categorizes them under specific outcome types. Based on the impact

thesis and list of indicators, *RAP 2021* developed an outcome typology for 13 outcome categories and some subcategories. *RAP 2021* added additional categories that were not specified in the AIMM sector frameworks to compile 28 outcome types. *RAP 2021* identified outcome claims for projects based on their backfilled AIMM worksheet, doing this for all IFC investment projects that were self-evaluated by IFC and validated by IEG between CY12 and CY19.

RAP 2023 leveraged the outcome typology developed by *RAP 2021*. The *RAP 2023* team reviewed the AIMM sector frameworks, which have remained the same since *RAP 2021*, suggesting that the outcome typology developed in *RAP 2021* was still relevant. However, *RAP 2023* enhanced *RAP 2021's* outcome typology by adding new subcategories and revising definitions of some subcategories. *RAP 2023* identified 33 outcome types (28 at the project level and 5 at the market level).

RAP 2024 used the outcome typology used by *RAP 2023* (table A.9) and applied it to all IFC projects that were self-evaluated by IFC and validated by IEG between CY21 and CY23, but only those projects with XPSRs validated by December 2023 were included in the analysis. *RAP 2024* reviewed the text of IEG EvNotes and coded descriptions of project- and market-level development outcomes that the projects were intended to achieve. *RAP 2024* included only outcome claims that were clearly identified in the EvNote to capture key objectives based on what the IEG evaluator had already determined were the main intended objectives. A small number of outcome claims were not accompanied by specific indicators to measure their results. IEG shared its outcome analysis approach with IFC in the Concept Note and responded to IFC's questions about the process.

An outcome was considered fully achieved, partially achieved, not achieved, or cannot be verified based solely on the text of the project EvNote, which itself validated the project's self-evaluation XPSR. The *RAP 2024* team did not apply any additional judgment, assessment, or methodology.

Table A.9. Outcome Types of International Finance Corporation Investment Projects

Outcome Type	Definition
1.1—Access to goods and services	Increase in number of final beneficiaries of goods and services of the project or company; increase in volume of goods and services produced by project or company
1.1.1—Access to goods and services (MSMEs)	Increase in number of MSMEs as final beneficiaries of goods and services of the project or company; increase in volume of goods and services produced or provided by project or company
1.1.2—Access to goods and services (gender)	Increase in number of final female beneficiaries of goods and services of the project or company; increase in volume of goods and services produced or provided by project or company
1.1.3—Access to services (customers)	Increase in number of individual customers as final beneficiaries of goods and services of the project or company; increase in volume of goods and services produced or provided by project or company
1.1.4—Access to goods and services (miscellaneous)	Increase in number of final beneficiaries of goods and services of the project or company other than MSMEs, female beneficiaries, and individual customers or a mix of these final beneficiaries; increase in volume of goods and services produced by project or company
1.1.5—Access to goods and services (direct client level)	Increase in capacity of project or direct client company to produce goods and services because of IFC investment
1.2—Quality and affordability of goods and services	Improved quality of goods and services produced by project or company compared with baseline or with other producers or providers; lower production costs or process; reduced prices of goods and services compared with the baseline or other producers or providers
1.2.1—Quality of goods and services	Improved quality of goods and services produced by project or company compared with the baseline or other producers or providers
1.2.2—Affordability of goods and services	Reduced prices of goods and services compared with the baseline or other producers or providers
1.2.3—Increased efficiency of direct client company	Lower production costs or processes of project or company
1.3—Increased capacity of final beneficiaries	Enhanced capacity of final beneficiaries as a result of advisory services or training that is part of project scope
1.4—Improved living standards (earnings) of individuals	Increase in revenue or decrease in expenditures by final beneficiaries (individuals) of goods and services produced by the project or company

(continued)

Outcome Type	Definition
1.5—Improved sales or profitability of enterprises	Increase in revenue, decrease in expenditures, or increase in overall productivity by final beneficiaries (enterprises) of goods and services produced by project or company
2.1—Suppliers and distributors reached	Increase in number of suppliers who provide inputs to project or company or expansion of network of distributors of goods or services produced by project or company
2.2—Improved capacity of suppliers and distributors	Increase in capacity of suppliers or distributors as a result of advisory services or training that is part of project scope
2.3—Improved sales and profitability of suppliers and distributors	Increase in volume of inputs provided by suppliers or increase in the goods or services to be distributed by its distributors
3.1—Increased employment	Increase in direct employment of client company
3.2—Improved capacity and skills	Training provided to employees of project or company
3.3—Improved earning of employees	Increase in wages to employees of project or company
4.1—Increased transfers to the government	Increase in payments by project or company to government, such as in the form of taxes, royalties, fees, or dividends
5.1—Increased money spent or transferred to community	Increase in payments to communities around the project or company, such as on health, educational, or vocational programs
6.1—Enhanced environmental and social standards of the client	Improvement in environmental and social standards by IFC
6.2—Greenhouse gas emissions	Decrease in or avoidance of greenhouse gas emissions
6.3—Efficient use of resources	Decrease in use of water and other resources, improvement in solid waste management, implementation of waste-to-energy project
7.1—Gross value added	Gross value added to economy (calculated based on a multiplier and expressed in monetary value)
7.2—Induced or indirect employment	Induced and indirect employment based on multipliers
7.3—Exports	Increase in exports of goods and services, generating foreign currency
8.1—Governance	Improvement in corporate governance or increase in capacity of client company
9. Competition in the market	Increase in ability of firms to enter, exit, compete, innovate, and strive for efficiency under fair and good regulatory governance; price changes; new practices, technology, product innovation (first movers); product and business model differentiation, change in product offering, value addition; increase in efficiency under fair and good regulatory governance

(continued)

Outcome Type	Definition
10. Resilience in the market	Increase in market depth and improvement in market structure, regulation, and governance to help markets withstand physical, financial, economic, or climate-related shocks; improved corporate governance of direct clients; diversification (for example, energy sources or funding sources in sectors or products); increase in capacity to face shocks and stress; increase in market depth and improvement in market structure, regulation, and governance (capacity of regulator); decrease in domestic supply volatility; increase in energy security; increase in financial stability and consumer protection
11. Integration in the market	Increase in physical or financial connectivity to support greater market integration, greater integration with financial markets and domestic and global value chains, enhanced physical or financial connectivity, geographical integration, integration with financial markets (including capital mobilization), data integration, growing domestic and global value chains, trade diversification, economic complexity
12. Inclusiveness in the market	Increase in fair and full access to all goods, services, finance, and economic opportunities, including for underserved groups; increased inclusiveness and improved access; establishment of market-wide enabling framework or standards supporting inclusive business; increase in diversity
13. Sustainability in the market	Adoption of climate-related environmentally and socially sustainable products, technologies, and practices; increased ability of firms and industries to apply environmentally and socially sustainable approaches to mitigate risk, realize opportunities, and maximize operational efficiency; development of legal or regulatory framework that fosters sustainability; broad capacity and supporting institutions or sustainability practice

Source: Independent Evaluation Group.

Note: IFC = International Finance Corporation; MSMEs = micro, small, and medium enterprises.

Association Between Work Quality and Development Outcomes

This *RAP* identifies factors specific to work quality and additionality that are associated with development outcomes.[1] This analysis is particularly important given the strong correlation between work quality and development outcomes. We caution that while the empirical associations between IFC's work quality and development outcome ratings are well established through correlation analysis, causality between IFC's work quality and development outcome ratings may not be possible in *RAP 2024* because both positive and

negative factors beyond IFC's work quality can contribute to the development outcome ratings of IFC investment projects.

IEG conducted a desk-based review of a universe of 19 projects whose work quality was rated unsatisfactory or whose development outcomes were rated highly unsuccessful in CY21–23 (with a cutoff date of December 31, 2023). The aim was to help identify the top three work quality factors that are associated with weak development outcomes and to provide a synthesis of evidence on these issues. Because there is only 1 IFC investment project rated excellent on work quality or highly successful on development outcome, this project was excluded from the analysis. Hence, the cohort consisted of 19 projects with projects rated highly unsuccessful on development outcomes or unsatisfactory on work quality. This analysis will help set the stage for future analysis. Based on this analytic framework, a more balanced approach (in terms of selecting projects on both ends of the spectrum) can be followed in future analysis.

A desk-based review was conducted to identify the most prevalent work quality factors associated with weak development outcomes. For the purposes of this analysis, we used the same taxonomy that was used to identify factors linked to performance. The application of this taxonomy at the time of the desk-based review allowed us to identify the most prevalent work quality factors that are associated with weak development outcomes (table A.10). These work quality factors were peer-reviewed by the *RAP 2024* team to ensure consistency and accuracy.

Table A.10. Sample for Analysis of Association Between Work Quality and Development Outcomes

		Development Outcome					
		HU	US	MU	MS	SU	HS
Work quality	E	0	0	0	0	1	0
	S	0	8	9	41	25	0
	PU	10	18	25	11	8	0
	U	4	4	1	0	0	0

Source: Independent Evaluation Group, Expanded Project Supervision Report database.

Note: The shaded region indicates International Finance Corporation investment projects rated highly unsatisfactory on development outcomes or rated unsatisfactory on work quality by the Independent Evaluation Group. E = excellent; HS = highly successful; HU = highly unsuccessful; MS = mostly successful; MU = mostly unsuccessful; PU = partly unsatisfactory; S = satisfactory; SU = successful; U = unsatisfactory; US = unsuccessful.

Analysis of Factors Beyond IFC's Influence Contributing to Weak Development Effectiveness of IFC Advisory Services Projects

IEG tested the hypothesis that external factors beyond IFC's influence contributed to the decline in development effectiveness in FY21–23 despite IFC doing its part on work quality. We reviewed 31 IFC advisory services projects validated by IEG for which development effectiveness was rated mostly unsuccessful or worse, but work quality was rated satisfactory or excellent. For each project, the *RAP* team identified the top three factors that positively or negatively affected project performance and classified them using the existing taxonomy of performance factors (similar to the one used for IFC investment projects), consisting of 5 categories and 51 subcategories developed by IEG.

The factor identification and classification exercise involved two steps. First, the *RAP 2024* team conducted its factor analysis by identifying the top three factors for each project based on a review of project evaluation documents and classified them according to the taxonomy. Second, the identified factors were peer-reviewed by the *RAP* team. The second step ensured not only appropriate classification of categories and subcategories but also correct identification of factors contributing to weak performance.

Association Between Additionality and Development Outcomes

For IFC investment projects, we analyzed the association between additionality and development outcomes. We conducted a desk-based review of IEG evaluations and validations in CY21–23 (with a cutoff date of December 31, 2023) of 18 IFC investment projects rated unsatisfactory on additionality or highly unsuccessful on development outcomes. Because there were only 2 IFC investment projects rated excellent on additionality or highly successful on development outcome, these projects were excluded from the analysis (table A.11). This analysis will help set the stage for future analysis of the association between additionality and development outcomes. Based on the analytic framework used in *RAP 2024*, a more balanced approach (in terms of selecting projects on both ends of the spectrum) can be applied in future analysis.

This desk-based review enabled us to analyze additionality associated with weak development outcomes. Additionality analyzed can be (i) financial and nonfinancial anticipated at Board approval in terms of frequency of occurrences, (ii) realized (fully or partially) or unrealized, and (iii) missed (only those not anticipated in the Board paper but materialized). *RAP 2024* focuses only on cases that were missed but materialized. The results from the desk-based review were peer-reviewed to ensure consistency and accuracy.

Table A.11. Sample for Analysis of Association Between Additionality and Development Outcomes

		Development Outcome					
		HU	US	MU	MS	SU	HS
Additionality	E	0	0	0	0	0	0
	S	2	0	0	0	0	0
	PU	6	0	0	0	0	0
	U	6	2	1	1	0	0

Source: Independent Evaluation Group, Expanded Project Supervision Report database.

Note: The shaded region indicates International Finance Corporation investment projects rated highly unsatisfactory on development outcomes or rated unsatisfactory on additionality by the Independent Evaluation Group. E = excellent; HS = highly successful; HU = highly unsuccessful; MS = mostly successful; MU = mostly unsuccessful; PU = partly unsatisfactory; S = satisfactory; SU = successful; U = unsatisfactory; US = unsuccessful.

Analysis of Identification and Tracking of Outcomes in the Anticipated Impact Measurement and Monitoring System

IEG's outcome-type analysis in *RAP 2024* showed that about 100 outcomes could not be verified in terms of achievement of outcomes. This was also an issue in *RAP 2023*. However, IFC management claimed that the introduction of the AIMM system in 2017 addressed this shortcoming. To test this hypothesis, IEG conducted an analysis of a universe of 21 projects evaluated and validated by IEG with "live" AIMM scores (that is, projects that were assigned ex ante AIMM scores at Board approval and evaluated or validated by IEG as of December 31, 2023). Given the small cohort, this is a preliminary analysis with the intent of setting the stage for future analysis.

The aim of the analysis was to test the hypothesis about whether the introduction of AIMM had addressed the shortcoming of outcome verification.

First, the Board papers of the 21 projects were reviewed to identify outcomes (both project-level and market-level outcomes). Second, these outcomes were matched against the outcomes identified from the EvNotes in this year's *RAP* as part of the outcome-type analysis (described in the Outcome-Type Analysis for International Finance Corporation Investment Projects section). Third, the Board papers were reviewed to verify whether each outcome had at least one indicator assigned to it. Fourth, the tracking system was reviewed to assess whether the identified indicators had been added to the tracking system (AIMM Navigator and Development Outcome Tracking System). Fifth, we confirmed whether or not each indicator had been monitored at least once in the tracking system (by confirming if at least one data point was available in the system). The last step excluded projects that closed prematurely (projects canceled or prepaid early) because the tracking of indicators might be incomplete as a result of the premature closing.

Association Between Client Responsiveness and Performance

RAP 2024 aimed to explore the association between processing time or preimplementation scoping time and performance ratings for IFC. For IFC investment projects, the processing time was based on the elapsed time from mandate to first disbursement. For IFC advisory services projects, the preimplementation scoping was based on the elapsed time from the Concept Note to implementation plan. This analysis included IFC investment and advisory services projects evaluated and validated by IEG during the 10-year period covered by *RAP 2024*. This was an exploratory analysis that can be strengthened in future *RAPs*.

The data source for the analysis of IFC investment projects was IFC business metadata from management information systems. The *RAP 2024* cohort predated some of the efficiency initiatives implemented at IFC (for example, accountability and decision-making, streamlining of environment and social procedures). However, the analysis included other Bank Group efficiency initiatives undertaken in previous years. The data source for the analysis of IFC advisory services projects was iPortal.

This analysis aimed to analyze considerations for IFC's efficiency improvements, which are part of the Evolution Agenda. We first identified the association between processing time for IFC investment projects and performance ratings. Next, we explored how the association differs among various subgroups in the portfolio, such as region, industry group, IDA and blend, and FCS. However, we could not conduct this analysis for IFC advisory services projects because not all projects go through the Concept Note stage (for example, some subprojects of approved programmatic umbrellas or "fast-track" projects that were follow-ons from previous engagements). In addition, out of the 411 standard advisory services projects evaluated and validated by IEG (FY13–23), 67 projects (16 percent) did not have a Concept Note date.

Methodology for the Multilateral Investment Guarantee Agency

For each Multilateral Investment Guarantee Agency (MIGA) guarantee project that has reached early operating maturity, MIGA underwriting staff conduct a self-evaluation by preparing a Project Evaluation Report (PER) that IEG independently validates through a PER Validation Note (ValNote). To conduct the project evaluation and validation, both MIGA and IEG staff refer to PER guidelines, which provide the evaluation framework and performance rating criteria. The evaluation system and performance ratings for MIGA projects are both objectives and benchmarks based. In addition to a focus on the achievement of expected objectives stated in the president's report at approval, the performance of MIGA guarantee projects is assessed against several benchmarks (such as performance of peer companies, the market, and similar industries) and considers unintended outcomes (both positive and negative).

The main performance assessment dimensions for MIGA guarantee projects are development outcome, MIGA role and contribution, and MIGA work quality. The PER also assesses the sustainability of development outcomes in the long term by examining the project's prospects over its remaining life.

» Development outcome measures performance across four indicators: project business performance, economic sustainability, environmental and social

effects, and foreign investment effects. It is rated on a six-point scale: highly successful, successful, mostly successful, mostly unsuccessful, unsuccessful, and highly unsuccessful. Before FY20, the ratings were based on a four-point scale: excellent, satisfactory, partly unsatisfactory, and unsatisfactory.

» MIGA role and contribution assesses the benefits and value added that MIGA brings to the client, the project, or the political risk insurance industry. It is rated on a four-point scale: excellent, satisfactory, partly unsatisfactory, and unsatisfactory.

» MIGA work quality addresses due diligence and underwriting processes, including of risk assessment and mitigation, and monitoring after the issuance of the MIGA guarantee. It is rated on a four-point scale: excellent, satisfactory, partly unsatisfactory, and unsatisfactory.

Analysis, Data Sources, and Sample Coverage

Table A.12 indicates the data sources and sample coverage of MIGA guarantee projects used in the *RAP 2024* analyses. We did not conduct statistical significance test on the ratings because of a smaller cohort of MIGA projects evaluated and validated by IEG at the subgroup level (for example, IDA and blend, regions, and MIGA sectors).

Table A.12. Data Sources and Sample Coverage of Multilateral Investment Guarantee Agency Guarantee Projects

Analysis	Data Sources	Sample Coverage
MIGA guarantee project performance ratings	IEG and MIGA data	MIGA guarantee projects in FY13–23 PER programs evaluated and validated as of June 30, 2024
Analysis of MIGA guarantee project outcome types	IEG and MIGA data	15 MIGA guarantee projects in FY21–23 PER programs evaluated and validated as of December 31, 2023
Factors affecting MIGA guarantee project implementation and performance	IEG and MIGA data, PER Validation Notes	26 MIGA guarantee projects in FY20–23 PER programs evaluated and validated as of December 31, 2023

Source: Independent Evaluation Group.

Note: IEG = Independent Evaluation Group; MIGA = Multilateral Investment Guarantee Agency; PER = Project Evaluation Report.

Factors Linked to Development Outcomes

To identify the factors influencing development outcomes of MIGA guarantee projects, the *RAP 2024* team performed a qualitative review and content analysis of project evaluation documents. This review included 26 MIGA guarantee projects for which the evaluation and validation was completed by the cutoff date of December 31, 2023. For each project, the *RAP* team identified the top three factors that positively or negatively affected project performance and classified them using the existing taxonomy of performance factors, consisting of 5 categories and 51 subcategories developed by IEG. For the performance factor analysis of MIGA guarantee projects, the same methodology was used for *RAP* as is used for IFC investment projects (for the taxonomy and other details, see the Methodology for the International Finance Corporation section in this appendix).

Outcome-Type Analysis for Multilateral Investment Guarantee Agency Guarantee Projects

Similar to IFC projects, *RAP 2021* developed a 13-category typology of intended outcomes for MIGA guarantee projects by leveraging MIGA's IMPACT (Impact Measurement and Project Assessment Comparison Tool), the ex ante assessment and monitoring tool that was adapted from IFC's AIMM system. No IMPACT sector-specific frameworks had been developed, so the same outcome typologies developed for IFC projects were applied in *RAP 2021*, with adaptations to some outcome types. Because MIGA had not retroactively applied IMPACT to its portfolio, the text of the president's reports on MIGA guarantee projects for coding of expected development outcomes was used for *RAP 2021*. *RAP 2023* enhanced the outcome typology developed by *RAP 2021* by adding new categories and revising the definitions of some categories. *RAP 2023* identified 30 outcome types (25 at the project level and 5 at the foreign investment level).

RAP 2024 used the outcome typology developed by *RAP 2021* and *RAP 2023* and applied it to all MIGA guarantee projects that were self-evaluated by MIGA and validated by IEG between FY21 and FY23, but only those projects with PERs validated by December 2023 were included in the analysis (table A.13). The text of the Development Outcome at Approval section of IEG

ValNotes was reviewed for *RAP 2024*, including the expected development outcome and coded descriptions of the project-level and foreign investment–level development outcomes the projects were intended to achieve. Because IEG performed the coding manually, there is a risk of subjective assignment of outcome types for specific outcome claims. IEG discussed its outcome analysis methodology with MIGA.

For MIGA guarantee projects, *RAP 2024* followed the same approach to assess outcomes achieved as is used for IFC projects. *RAP 2024* assessed the extent to which expected outcome claims were achieved at evaluation by verifying the results presented in the project's ValNote. An outcome claim was considered fully achieved, partially achieved, not achieved, or cannot be verified based solely on the text of the project's ValNote, which itself validated the project's self-evaluation PER. The *RAP 2024* team did not apply additional judgment, assessment, or methodology.

Table A.13. Outcome Typology for Multilateral Investment Guarantee Agency Guarantee Projects

Outcome Type	Description
1.1—Access to goods and services	Increased number of final beneficiaries of goods and services of the project or company. Increased volume of goods and services produced by the project or company can be considered under this outcome type.
1.1.1—Access to goods and services (MSMEs)	Increased number of MSMEs as final beneficiaries of goods and services of the project or company. Increased volume of goods and services produced or provided by the project or company.
1.1.2—Access to goods and services (female)	Increased number of final female beneficiaries of goods and services of the project or company.
1.1.3—Access to goods and services (customers)	Increased number of individual customers as final beneficiaries of goods and services of the project or company. Customers of utility services are representative of this group. Increased volume of goods and services produced or provided by the project or company.
1.1.4—Access to goods and services (miscellaneous)	Increased number of individual customers as final beneficiaries of goods and services of the project or company. Customers of utility services are representative of this group. Increased volume of goods and services produced or provided by the project or company.

(continued)

Outcome Type	Description
1.2—Quality and affordability of goods and services	Improved quality of goods and services produced by the project or company, compared with the baseline or with other producers or providers. Lower production costs and process are covered by affordability. Reduced prices of goods and services compared with the baseline or with other producers or providers.
1.3—Enhanced capacity of final beneficiaries	Enhanced capacity of the final beneficiaries as a result of advisory services or training that is part of the project scope.
1.4—Improved living standards (earnings) of individuals	Increased revenue or reduced expenditure by the final beneficiaries (individuals) of goods and services produced by the project or company.
1.5—Improved sales and profitability of enterprises	Increased revenue, reduced expenditure, or increased overall productivity by the final beneficiaries (enterprises) of goods and services produced by the project or company.
1.6—Economic return	Economic rate of return.
1.7—Financial and business performance of direct clients	Financial and business performance of direct clients, mostly project-executing agencies.
2.1—Suppliers and distributors reached	Increased number of suppliers who provide inputs to the project or company, or the project expands the network of distributors of goods or services produced by the project or company.
2.2—Improved capacity of suppliers and distributors	Capacity of suppliers or distributors improved as a result of advisory services or training that is part of the project scope.
2.3—Improved sales and profitability of suppliers and distributors	The project increases the volume of inputs provided by its suppliers, or the project increases the goods or services to be distributed by its distributors.
3.1—Increased employment	Increased direct employment of the client company.
3.2—Improved capacity and skills	Training is provided to the employees of the project or company.
3.3—Improved earning of employees	Increased wages to employees of the project or company.
4.1—Increased transfers to the government	Payment from the projects or companies to the governments, such as tax, royalties, fees, and dividend.
5.1—Increased money spent and transferred to the communities	Payment to the communities around the project or company, such as health, educational, and vocational programs in association with infrastructure projects.
6.1—Enhanced E&S standards of the client	MIGA supports its clients to enhance their E&S standards.

(continued)

Outcome Type	Description
6.2—Greenhouse gas reduction	Projects such as renewable energy projects and energy efficiency projects that contribute to the reduction or avoidance of greenhouse gases.
6.3—Efficient use of resources	The project will reduce use of water and other resources, or the project will promote solid waste management and implement a waste-to-energy project.
7.1—Gross value added	The project brings gross value added to the economy, which is calculated based on a multiplier and expressed in monetary value.
7.2—Induced or indirect employment	Induced and indirect employment as a result of the project. This is also based on the multipliers.
7.3—Export sales	The project increases exports of goods and services produced. The economy's external balance from the generation and consumption of foreign currency.
8.1—Governance	Enhanced governance or capacity of MIGA's client company.
9. Business and sector practices	Potential to improve (financial or operational) performance of future investments through demonstration or transfer of new technologies, capabilities, practices, or business models.
10. Market development	Potential to enhance the market structure through increased competitiveness, resilience, integration, enhancements to the regulatory environment, and so on.
11. Development reach	Potential to stimulate future investments that increase inclusion and reduce inequality by reaching underserved populations (BoP, women, youth, and so on).
12. Sustainability	Potential to stimulate future investments to focus on climate change adaptation and mitigation, and adopting improved E&S standards and practices.
13. Signaling effects	Potential to stimulate further foreign investment in contexts where there are real or perceived barriers for domestic and foreign investors and lenders.

Source: Independent Evaluation Group.

Note: BoP = base of the pyramid; E&S = environmental and social; MIGA = Multilateral Investment Guarantee Agency; MSMEs = micro, small, and medium enterprises.

Limitations for the International Finance Corporation and the Multilateral Investment Guarantee Agency

Three limitations are identified for IFC investment and MIGA guarantee projects:

- For the outcome-type analysis, the specific intended outcomes are not rated in IEG project evaluation and validation documents. The objectives of IFC investment and MIGA guarantee projects are assessed but not rated because the IEG validations of the XPSRs are a benchmark-based (that is, market and industry standards) system, which considers long-term sustainability effects and trends rather than focusing solely on the achievement of the intended objectives. It can be challenging to assess the extent of the achievement unless the project evaluation and validation documents have explicit statements on the achievement of intended outcomes.
- For the analysis of the identification and tracking of indicators, it should be noted that AIMM and XPSR have different purposes and follow different methodologies for rating development outcomes and impacts. XPSR is an evaluation system in which the development outcome rating is based on four dimensions: project business success, economic sustainability, environmental and social effects, and private sector development. In contrast, AIMM is an ex ante analysis and monitoring system in which the overall impact score is based on two dimensions: project outcome rating (equivalent to the economic sustainability dimension in the XPSR) and market outcome rating (equivalent to the private sector development dimension in the XPSR). The AIMM system alone is not fully comparable with XPSR because IFC uses separate systems to monitor other dimensions (such as financial sustainability and environmental and social effects) of development outcome. AIMM assessments are provided throughout the project cycle until the project reaches the target year, whereas XPSRs are prepared for a sample of IFC investments once the project reaches early operating maturity. However, in rating the development outcome, XPSR also considers expected and unexpected effects based on extensive analysis of why things happened or did not happen and comparison with peer companies to draw lessons for future operations.
- Although the empirical associations among IFC's work quality, additionality, and development outcome ratings are well established through correlation analysis, causality among IFC's work quality, additionality, and development outcome ratings may not be possible in *RAP 2024* because both positive and negative factors beyond IFC's work quality and additionality can contribute to the development outcome ratings of IFC investment projects.

Methodology for Country Programs

This subsection presents the methodology applied for construction of the data sets and the analysis of trends, challenges, and levers of country programs. The country programs analysis uses two main data sources—namely, the IEG Completion and Learning Review Validations (CLRVs) and the Bank Group Country Opinion Survey (COS) program. The ratings and factors linked to implementation were built from these. Additionally, IEG Country Program Evaluations, relevant evaluations, and document reviews were undertaken as needed.

The term *CLRV* is used to refer to all IEG validations of self-assessments of country program performance. Since 2003, IEG has validated every Completion and Learning Review completed by the Bank Group. Previously, these were called Completion and Learning Review Reviews and Country Assistance Strategy Completion Report Reviews. These are all publicly disclosed after Board discussion, and the data set is publicly available on the IEG website.[2] The method for Completion and Learning Review and CLRV ratings is described in the country engagement guidance (World Bank 2021a).

Data Set Construction: Source and Coverage

The analysis of country program trends in *RAP 2024* includes the ratings of all CLRVs finalized by IEG between FY13 and FY24. The analysis of challenges and levers was undertaken on a subset of these CLRVs—the two most recent CLRVs for countries that have at least two CLRVs produced by IEG in FY13–24 with assigned ratings. This allows us to compare shifts in ratings between the first and the second CLRVs and thereby deepen the analysis of the factors and the associated challenges and levers linked to shifts in ratings. Table A.14 outlines the universe of CLRVs and the ones eligible for the analysis of trends, challenges, and levers.

Table A.14. Data Set Eligibility Criteria: Completion and Learning Review Validations

CLRV Data Set	CLRVs (no.)	Countries (no.)
Universe: all CLRVs completed by IEG during FY03–24	354	121
Trends analysis: all CLRVs completed by IEG during FY13–24	209	118
Challenges and levers (detailed analysis): the two most recent CLRVs for countries that have at least two CLRVs with assigned ratings during FY13–24	162	81

Source: Independent Evaluation Group.

Note: CLRV = Completion and Learning Review Validation; IEG = Independent Evaluation Group.

RAP 2024 used the results of the World Bank Group COS to identify perceptions associated with shifts in performance ratings.[3] The COS seeks to systematically measure and track the perceptions of the Bank Group's clients, partners, and other stakeholders who are development practitioners in client countries across the globe. It has been conducted regularly since 2012, with countries surveyed on a three-year cycle by an independent local research firm. From the universe of all COS rounds, we attempted to include only those conducted in years covered in the Country Partnership Framework (CPF) period of a CLRV. Among the 162 CLRVs in the detailed analysis of challenges and levers, 148 CLRVs (91 percent) had at least one corresponding COS round, while the remaining were not matched. Table A.15 outlines the universe of COS rounds and the ones eligible for the analysis of levers.

Table A.15. Data Set Eligibility Criteria: Country Opinion Survey

COS Data Set	COS Rounds (no.)	Countries (no.)
Universe: all COS rounds completed by the World Bank Group during FY12–23	425	131
Levers: COS rounds conducted in country years covered by one of the 162 CLRVs included in the detailed analysis	229	81

Source: Independent Evaluation Group.

Note: CLRV = Completion and Learning Review Validation; COS = Country Opinion Survey.

The core data in the country programs analysis are the publicly available CLRV ratings. In addition, two data sets were constructed by the *RAP* team on factors linked to performance for the 162 CLRVs included in the detailed analysis. The first data set captures eight factors linked to Bank Group performance at the country program level, and the second captures Bank Group perceptions from the COS, aligned with CPF periods. A detailed description of each of these three data sets and their corresponding data items follows:

1. Publicly available CLRV ratings. Each CLRV contains two ratings: development outcomes and Bank Group performance. These ratings capture the following information:

» Development outcomes. The extent to which the CPF was successful in achieving its stated objectives. It considers any unforeseeable shocks or events and highlights aspects of flexibility and adaptation that are relevant to the objectives but assessed as part of the Bank Group performance. Development outcomes are rated highly satisfactory, satisfactory, moderately satisfactory, moderately unsatisfactory, unsatisfactory, and highly unsatisfactory.

» World Bank Group performance. Based on how well the CPF was designed and how well the Bank Group implemented the CPF program, including learning and adapting. It includes separate discussions of World Bank, IFC, and MIGA performance where relevant. Bank Group performance is rated superior, good, fair, and poor.

2. Factors linked to Bank Group performance in country programs. All 162 CLRVs included in the detailed analysis of challenges and levers were downloaded from the IEG website. Then, we captured the CLRV text that justifies the Bank performance rating in a structured manner under eight factors and assigned four-point ratings to these factors using a rubric. The four-point scale (superior, good, fair, and poor) is based on the country engagement guidance (World Bank 2021a) used by CLRVs to rate the Bank Group performance and drew from a similar analysis conducted for *RAP 2022*. The rubric was updated following the testing phase based on consultations between the team. Table A.16 describes the eight factors, and box A.1 details the coding process.

3. Bank Group perceptions from the COS, aligned with CPF periods. We selected 12 questions from the COS questionnaire. These questions were selected because they were relevant for Bank Group performance and overlapped with country engagement guidance. Additionally, the text captured under the learning and adaptation factor of the 37 CLRVs coded by this *RAP's* team was extracted for recent CLRVs that covered at least one year since FY20. We analyzed this text using the types of adaptive management identified in IEG's outcome orientation evaluation (World Bank 2020b). The types of adaptive management found relevant in *RAP* are discussed in the Levers section in chapter 5.

Table A.16. Eight Factors for Examining World Bank Group Performance in Country Programs

Factor	Description
1. Relevance of country program	The extent to which a country program's objectives are relevant to its context and consistent with the national strategy (Are the binding constraints to development addressed? Is the program selective enough? Is there government buy-in to the selected priorities?). The extent to which the instruments and interventions were appropriate for achieving the Country Partnership Framework objectives (Are the interventions realistic and effective? Are the chosen instruments fit for purpose?).
2. Quality of results framework	The extent to which the results framework and intervention logic were strong (linking objectives and outcomes for interventions of all World Bank Group institutions), with appropriate indicators to track progress: (i) indicators capture well all dimensions of the objectives and (ii) indicators are clear, with baselines, targets, and target years. Whether there were weaknesses in the results framework design and whether they were corrected during implementation. Attribution and contribution to country program objectives are appropriately reflected.
3. Risk identification	Whether the Bank Group identified the critical risks to achieving the country program objectives and considered these risks appropriately at the design and PLR stages. Consideration of the trade-off between risk and development impact is expected, particularly in a fragile state.
4. Risk mitigation	The extent to which the Bank Group proactively mitigated risks to the achievement of relevant program objectives. Were risk mitigation mechanisms applied, and were they effective?

(continued)

Factor	Description
5. Learning and adaptation (not rated[a])	Responsiveness to changing circumstances, risks, priorities, and demands of the country, including introducing midcourse correction when needed and making substantive changes (as appropriate) to objectives and program at the PLR stage (or any other stage if substantial changes occurred after the PLR stage), with correspondingly appropriate adjustments to the results framework. Strategic use of knowledge products and evaluations to inform the adjustment to the program and the objectives.
6. Support to implementation	The extent to which the Bank Group supported effective program implementation through appropriate supervision (including timeliness of program implementation, mechanisms such as joint reviews with government, and attention to safeguard and fiduciary issues), technical assistance and knowledge products (applicability, quality, and dissemination). If there were efforts for improving alignment with country systems, it should be considered a positive aspect.
7. One World Bank Group approach	The extent to which there was appropriate collaboration and appropriate division of labor among the World Bank, the International Finance Corporation, and the Multilateral Investment Guarantee Agency (evidence of synergies, duplication of efforts, and coherence). Consider four different types of collaboration: (i) joint implementation, (ii) sequenced work, (iii) parallel work, and (iv) separate but coordinated work.
8. Development partner collaboration	The extent to which other development partners' programs were considered at design, and the strength of collaboration with development partners during implementation. Consider whether there was any duplication of efforts or redundancies and whether the Bank Group assumed a lead role among donors in any country program areas.

Source: Independent Evaluation Group.

Note: PLR = Performance and Learning Review.
a. The choice of not rating the learning and adaptation factor reflects that older Completion and Learning Review Validations do not discuss this dimension in depth and thus made the data inconsistent over time; therefore, rating would have been unreliable.

Box A.1. Coding Process for Completion and Learning Review Validations

The coding team consisted of four evaluators who were assigned between 40 and 50 Completion and Learning Review Validations (CLRVs) each. The coding process of capturing the relevant text and assigning a rating for each of the eight factors entailed three stages:

1. Guiding material and training sessions. A coding template in Microsoft Excel and a rubric were developed, pilot-tested, and refined. All four coders participated in a training session, and after each coder had individually reviewed six pilot CLRVs, the team met for a second session. Coders were then assigned a set of CLRVs for which they completed the coding.

2. Quality assurance. Each coder reviewed their set of assigned CLRVs individually. To ensure intercoder reliability, the team conducted an iterative series of checks that included the following:

- Comparing the frequency of assigned ratings by each coder;
- Double-coding and reviewing a limited subset of CLRVs;
- Flagging any factor in which coders were in doubt between a rating of good or fair; and
- Having final audits performed by an Independent Evaluation Group staff member who had served as the CLRV coordinator for more than four years and was not part of the coding team.

3. Focused review of World Bank Group collaboration. When coding the 162 CLRVs included in the detailed analysis, we undertook a focused review of the Bank Group collaboration using the following rubric:

- Not applicable: The country program was exclusively International Development Association, with no work intended or undertaken by the International Finance Corporation (IFC) or the Multilateral Investment Guarantee Agency (MIGA).
- Collaboration missed (poor): The country program had no evidence of intended collaboration between the World Bank, IFC, and MIGA, despite areas where there was room for collaboration identified in the CLRV.

(continued)

Box A.1. Coding Process for Completion and Learning Review Validations (cont.)

- Intended collaboration (fair): The country program alluded to intended collaboration between the World Bank, IFC, and MIGA at the design stage, but there was no example of successful collaboration reported.
- Evidence of one collaboration (good): There was some consideration of how each Bank Group institution would use its areas of strength and specific instruments at program design, with one successful collaboration reported.
- Two or more instances of collaboration (superior): There was thoughtful consideration of how each Bank Group institution would use its areas of strength and specific instruments at program design, with two or more instances of successful collaboration reported.

Additionally, we reviewed the text of each CLRV and parsed out each instance of reported collaboration between the World Bank, IFC, and MIGA, coding its sector, and the instruments used.

Source: Independent Evaluation Group.

Analysis

Trends

To analyze the trends in country program, we used the ratings of all 209 CLRVs from 118 countries completed by IEG in FY13–24. These data were merged with country characteristics such as Region, eligibility, income level, and FCS status. In analyzing ratings by year, we used the smoothing approach adopted since *RAP 2020* (World Bank 2020a), in which a CLRV is included in each of the fiscal years covered by the CLRV. This smoothing approach enables the analysis of the number of countries receiving a rating in each year. The analysis also included breakdowns by country characteristics. When reporting ratings by country groups, such as income level, we included only the latest CLRV available for each country (118 CLRVs from 118 countries). This allows us to represent each country group more accurately,

avoiding biasing results toward client countries that have completed more CPF cycles in the FY13–24 period.

Challenges

To identify which factors strongly influence the Bank performance rating, we used the data set generated after coding the 162 CLRVs. On this data set, we performed two types of analysis. First, we undertook a descriptive analysis of the frequencies of ratings, the factors associated with shifts in ratings at the country level, and the types and sectors of Bank Group collaboration. As a reference point, these frequencies were compared with Bank performance ratings to understand variations between the different factors. Second, we completed *t* tests of equality of means for each of the seven rated factors between the subgroups of CLRVs with Bank Group ratings of good or superior and fair or poor.

Through these analyses, we identified factors associated with the overall Bank performance rating. The four factors displaying a difference significant at the 5 percent level were considered to have a strong influence on Bank Group performance (that is, relevance, risk identification, risk mitigation, and support to implementation). The three remaining factors (quality of results frameworks, development partner collaboration, and One World Bank Group approach) remain important challenges, even though they have limited influence on Bank Group performance. For each factor, we identified challenges through a review of instances of lower ratings and the extraction of qualitative text examples.

A review of IEG reports was undertaken to identify specific factors that influence Bank Group collaboration. A content analysis and semistructured qualitative review of Bank Group collaboration were undertaken. For the content analysis, an expert reviewer coded extracted Bank Group collaboration text to the sector and type of collaboration. This coding was reviewed by a quality assurer. For the semistructured qualitative review, two coders independently collated text relating to Bank Group collaboration from a defined document list. The review included a range of documents for the FY13–24 period that contained analysis of Bank Group collaboration at the country level. This led to the inclusion of all the *RAP* series for the period, 21

Country Program Evaluations, one IEG evaluation (*World Bank Group Joint Projects: A Review of Two Decades of Experience—Lessons and Implications from Evaluation*), and one internal Just-in-Time Policy Note. Each reviewer interpreted the text extracts for patterns individually, and then they met as a group to discuss interpretations until they formed a consensus analysis of clusters of issues.

Levers

Statistical tests were conducted on the 12 selected COS questions. On further review, we classified the list into six questions related to adaptive management practices and six questions that were unrelated. All 12 COS questions included in table 5.1 were measured with the Likert scale (1 = to no degree at all; 10 = to a very significant degree). Survey questions were reported as averages based on country years that match the 162 coded CLRVs, split by their Bank Group performance rating. For each question, we examined whether the gaps were significant in a *t* test of equality of means between the subgroups of Bank Group performance in the CLRV rating of good or superior and fair or poor. For each type of adaptive management, we identified examples of practices undertaken in countries through a content analysis of text captured in 37 recent CLRVs (which covered at least one year since FY20) on learning and adaptation and the extraction of qualitative text examples.

Limitations

Limitations in the country programs analysis are mostly due to limitations in its main data sources, the IEG CLRVs and the Bank Group COS program.

Limitations related to CLRVs include the timeliness of data, completeness, and subjectivity of ratings. A CPF must have closed and its CLR must have been completed before IEG produced a CLRV, which often includes substantial lags. For example, the Jamaica FY14–19 CLRV was completed in FY24. Given the sparse coverage for recent years, we stopped the analysis of ratings at FY20. Within country programs, there is a wide diversity of project types, locations, and contexts, which means that a range of mechanisms that contribute to results and performance are not visible within the ratings and text. As described by previous IEG reports, CLRVs provide a partial picture

of country-level development outcomes because of their "overemphasis on those results that can be [quantitatively] measured and on results from lending projects" (World Bank 2020b, xiii). Moreover, changes to the CLRV format and a change in emphasis in the discussion over time limit the analytic depth for comparison.

The COS program has the usual limitations of opinion surveys, which is partly mitigated by a robust methodology. The survey is conducted by an independent research firm in each country. The collection process is guided by a standardized sample composition, where each stakeholder group is assigned a recommended proportion. More details are available on the COS website.

References

World Bank. 2020a. *Results and Performance of the World Bank Group 2020*. Independent Evaluation Group. World Bank.

World Bank. 2020b. *The World Bank Group Outcome Orientation at the Country Level*. Independent Evaluation Group. World Bank.

World Bank. 2021a. "IBRD/IFC/MIGA/IDA Guidance: Country Engagement." Guidance for Preparation of Country Engagement Products, World Bank.

World Bank. 2021b. *Results and Performance of the World Bank Group 2021*. Independent Evaluation Group. World Bank.

World Bank. 2023. *Results and Performance of the World Bank Group 2023*. Independent Evaluation Group. World Bank.

[1] Development outcome is rated on a six-point scale: highly successful, successful, mostly successful, mostly unsuccessful, unsuccessful, and highly unsuccessful. For additionality and work quality, the rating is on a four-point scale: excellent, satisfactory, partly unsatisfactory, and unsatisfactory.

[2] The data set of CLRV ratings is publicly available at https://ieg.worldbankgroup.org/ieg-data-wb-country-partnership-framework-ratings.

[3] More information on the COS is available at https://www.worldbank.org/en/programs/world-bank-country-opinion-surveys.

Appendix B. World Bank Dashboards

This appendix aims to support the learning purposes of *Results and Performance of the World Bank Group 2024* by providing access to online dashboards that underlie the report's findings. Table B.1 presents a variety of links for World Bank and country program dashboards. These present data that combine publicly available data with some additional Independent Evaluation Group–generated elements. The dashboards reinforce the learning purposes of *Results and Performance of the World Bank Group* by allowing different audiences to explore the data according to their specific interests. Appendix A outlines the process for developing the underlying data sets.

The links provided in table B.1 are for interactive dashboards from which charts in this report are derived. Each dashboard offers options for users to explore and visualize the data in various ways. Some explanatory notes are included within each dashboard to guide users. This approach is being tested in *Results and Performance of the World Bank Group 2024* as an alternative to static descriptive appendixes, which often struggle to meet the diverse data needs of different stakeholders. The limitations to these dashboards include the following:

- Varying data compilation dates. The dashboards contain data compiled at different times rather than providing a uniform snapshot. Users should note the specific time periods for each data set when conducting analyses, as comparisons across data sets may not always reflect the same point in time.
- Limited analytic depth. The dashboards provide descriptive visualizations of the data, offering users interactive ways to explore trends and patterns. However, they do not include the more in-depth statistical analyses found in the main report. Users seeking to replicate or expand on the report's full range of analyses would need to perform these additional statistical tests independently.
- Data quality assurance. Rigorous measures have been implemented to ensure data reliability. These measures include the use of standard reputable data sources and the implementation of comprehensive data quality review.

Table B.1. Available Dashboards Compiled for *Results and Performance of the World Bank Group*

Data Set and Link	Data Sources	Coverage
World Bank performance ratings and decomposition	IEG data on ICRR or Project Performance Assessment Report ratings	2,982 World Bank lending projects closed during FY13–23 and evaluated by IEG as of June 30, 2024
World Bank project outcome types	IEG ICRR rating data on project objectives, project data from World Bank Data Explorer, and manually coded outcome types of objectives	1,336 World Bank investment project financing projects closed during FY17–23 and evaluated by IEG as of December 2023
Factors linked to performance of World Bank projects	World Bank project ICR documents, manually coded and machine learning–predicted factors	1,118 World Bank investment project financing projects closed during FY18–23 and evaluated by IEG as of December 2023
Country program ratings	CLRV documents	All 209 CLRVs completed by IEG during FY13–24 (from 118 countries)
Factors linked to performance of country programs	Text of CLRV documents	162 CLRVs from 81 countries, representing the 2 most recent CLRVs for countries with at least 2 completed by IEG during FY13–24

Source: Independent Evaluation Group.

Note: CLRV = Completion and Learning Review Validation; ICR = Implementation Completion and Results Report; ICRR = Implementation Completion and Results Report Review; IEG = Independent Evaluation Group.

Appendix C. Decomposition Analysis of International Finance Corporation Project Rating Trends

This is a technical appendix on decomposition analysis for International Finance Corporation (IFC) investment and advisory services projects. The development outcome ratings for IFC investment projects and the development effectiveness ratings for IFC advisory services projects used in the decomposition analysis are as of June 30, 2024.

Decomposition Analysis of International Finance Corporation Investment Projects

Decomposition analysis shows the net contribution from each subgroup to the changes in IFC's overall development outcome ratings over the long term. This section describes the decomposition analysis for IFC investment projects across various subgroups. The subgroups in the decomposition analysis for IFC investment projects refer to region, International Development Association (IDA) and blend countries, countries classified as fragile and conflict-affected situations (FCS), and industry group. The long term in the decomposition analysis refers to IFC investment projects that have been evaluated and validated by the Independent Evaluation Group (IEG) during two time periods: calendar year (CY)13–15 and CY21–23. The changes in IFC's overall development outcome ratings between these two time periods can be due to (i) changes in development outcome ratings of subgroups without any changes in their respective shares in the number of projects evaluated and validated by IEG in the two time periods; (ii) changes in the respective shares of subgroups in the number of projects evaluated and validated by IEG without any changes in their development outcome ratings in the two time periods; or (iii) both (i) and (ii). The following formula computes the net contribution of subgroups to changes in IFC's overall development outcome ratings between two time periods t_0 (CY13–15) and t_1 (CY21–23) using decomposition analysis:

$$\bar{R}_{t1} - \bar{R}_{t0} = \sum_i s^i_{t1} \times (r^i_{t1} - r^i_{t0}) + \sum_i r^i_{t0} \times (s^i_{t1} - s^i_{t0})$$

where

t_0 = CY13–15, t_1 = CY21–23

i = {regions}, {lending status}, {FCS status}, {others}

r^i_t = development outcome ratings for subgroup i in period t

s^i_t = share of projects for subgroup i evaluated and validated by IEG in period t

The decomposition analysis is purely descriptive, and causality with development outcomes may not be possible because there may be factors beyond the subgroups that could affect IFC's overall development outcome ratings.

Decomposition Analysis by Regions

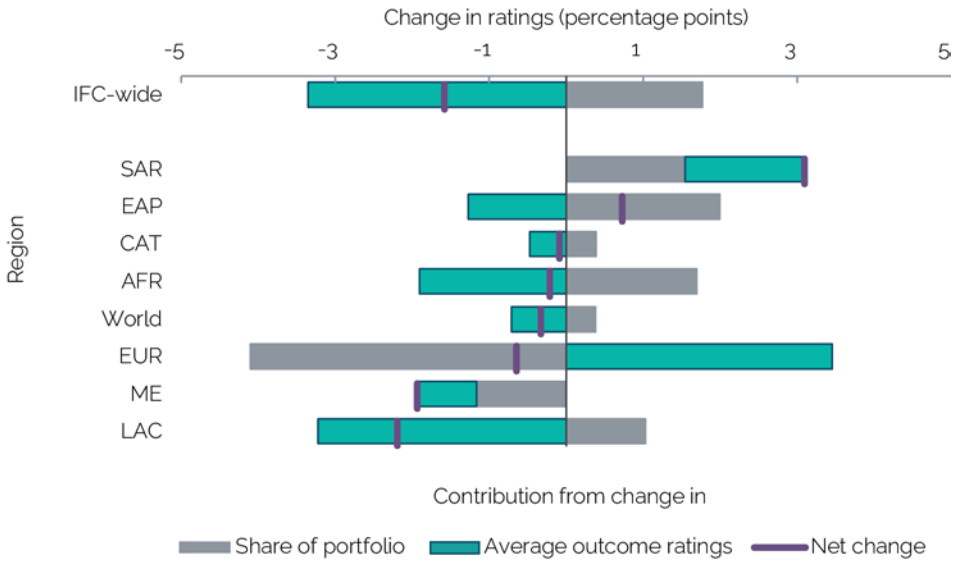
Decomposition analysis shows that both South Asia and East Asia and the Pacific contributed positively to IFC's overall development outcomes over the long term (figure C.1). South Asia's net contribution to IFC's overall development outcomes was the strongest among the regions. This was due to a positive contribution from increase both in ratings (from 58 percent to 69 percent) and in the share of projects (from 11 percent to 14 percent). East Asia and the Pacific also contributed positively to IFC's overall development outcomes, mainly due to a positive change in its share of projects (from 12 percent to 16 percent), while its ratings declined (from 54 percent to 45 percent).

Latin America and the Caribbean was followed by the Middle East as the regions that made the most negative contributions to IFC's overall development outcomes over the long term. The negative contribution from Latin America and the Caribbean to IFC's overall development outcomes was mainly due to a substantial decline in ratings (from 60 percent to 47 percent), while its share of projects increased slightly (from 23 percent to 25 percent)—making it the region with the largest share. The Middle East

was the only region with a negative contribution both from the decline in ratings (from 56 percent to 40 percent) and in the share of projects (from 7 percent to 5 percent).

Other regions such as Europe, Africa, and Central Asia and Türkiye also contributed negatively to IFC's overall development outcomes over the long term. Although Europe's performance improved (from 38 percent to 87 percent), its share of projects declined (from 18 percent to 7 percent). However, Africa and Central Asia and Türkiye showed an increase in the share of projects (Africa: 21 percent to 24 percent; Central Asia and Türkiye: 5.5 percent to 6 percent), while their ratings declined (Africa: 51 percent to 43 percent; Central Asia and Türkiye: 62 percent to 54 percent).

Figure C.1. Decomposition Analysis by Region, Calendar Years 2013–15 Versus Calendar Years 2021–23



Source: Independent Evaluation Group, Expanded Project Supervision Report database.

Note: AFR = Africa; CAT = Central Asia and Türkiye; EAP = East Asia and the Pacific; EUR = Europe; IFC = International Finance Corporation; LAC = Latin America and the Caribbean; ME = Middle East; SAR = South Asia; World = multiregional.

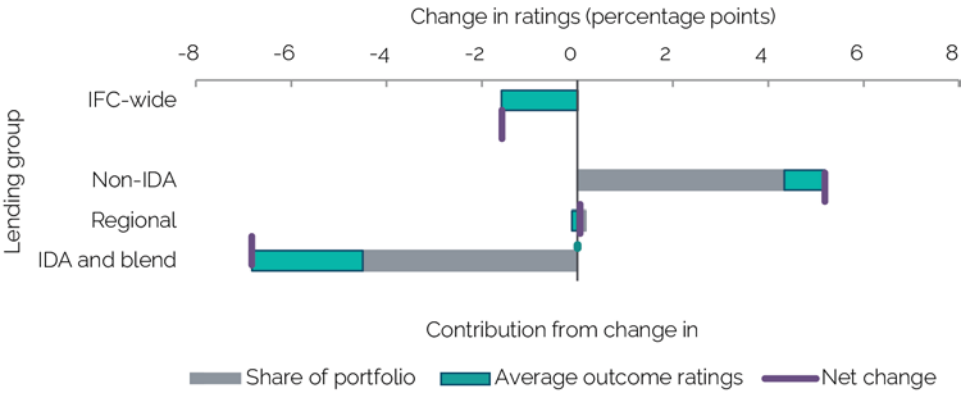
Decomposition Analysis by International Development Association and Blend Countries

Among all subgroups, IDA and blend was the single largest negative contributor to IFC's overall development outcomes over the long term (figure C.2). This was due to a decline in both ratings (from 54 percent to 46 percent) and

in the share of projects (from 35 percent to 27 percent). This decline in the share of projects was mainly due to India's graduation from IDA in FY14, which partially drove the 8 percentage point decline.

Non-IDA countries contributed positively to IFC's overall development outcomes over the long term. Projects in non-IDA countries historically outperformed those in IDA and blend countries. The decomposition analysis shows that non-IDA countries contributed positively to IFC's overall development outcomes because of increases in both ratings and share of projects.

Figure C.2. Decomposition Analysis by International Development Association and Blend Countries, Calendar Years 2013–15 Versus Calendar Years 2021–23



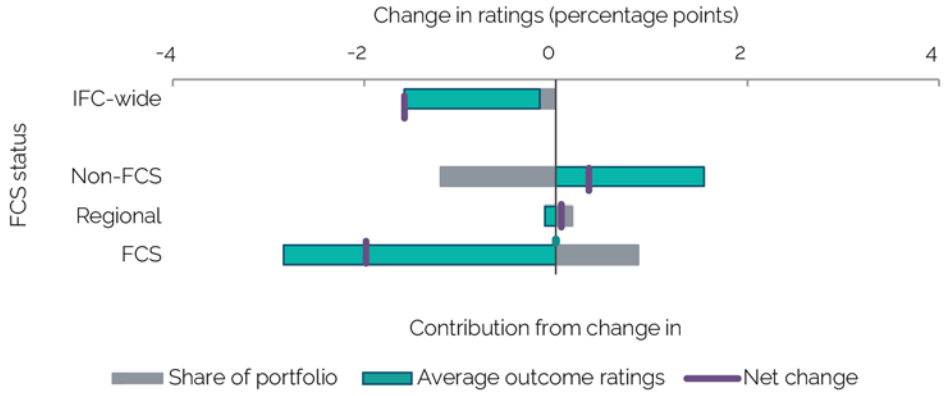
Source: Independent Evaluation Group, Expanded Project Supervision Report database.

Note: Regional projects are projects in more than one country and are a separate category to the IDA and blend and non-IDA groups. IDA = International Development Association; IFC = International Finance Corporation.

Decomposition Analysis by Fragile and Conflict-Affected Situations

FCS contributed negatively to IFC's overall development outcomes over the long term, while non-FCS contributed positively (figure C.3). The negative contribution of FCS was due to the substantial decline in ratings (from 50 percent to 17 percent), while its share of projects increased (from 7 percent to 9 percent). Conversely, projects in non-FCS contributed positively to IFC's overall development outcomes because of an increase in ratings (from 55 percent to 57 percent), while the share of projects declined (from 81 percent to 79 percent).

Figure C.3. Decomposition Analysis by Fragile and Conflict-Affected Situations Status, Calendar Years 2013–15 Versus Calendar Years 2021–23



Source: Independent Evaluation Group, Expanded Project Supervision Report database.

Note: Regional projects are projects in more than one country and are a separate category to the IDA and blend and non-IDA groups. FCS = fragile and conflict-affected situations; IDA = International Development Association; IFC = International Finance Corporation.

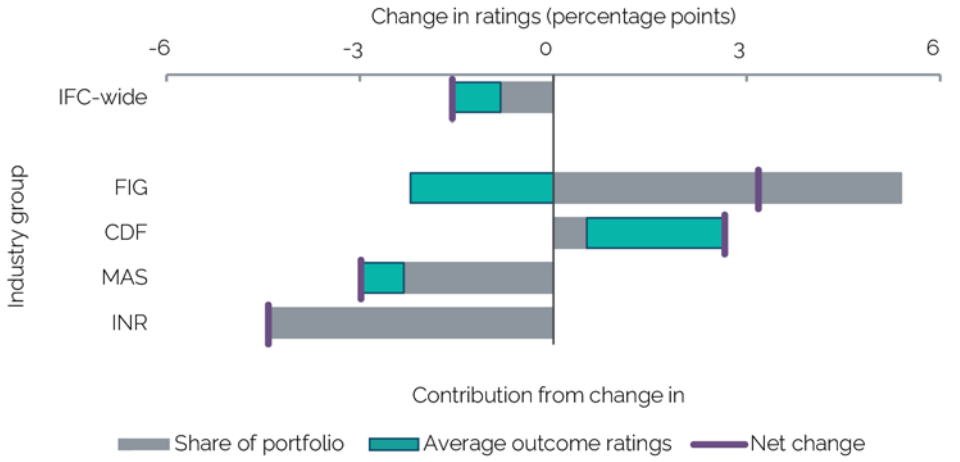
Decomposition Analysis by Industry Group

The Financial Institutions Group, followed by Disruptive Technologies and Funds, contributed positively to IFC's overall development outcomes over the long term. The positive contribution of the Financial Institutions Group was based solely on an increase in the share of projects (from 32 percent to 42 percent), while its ratings declined (from 55 percent to 50 percent). It is noteworthy that Financial Institutions Group ratings are in parallel to the average IFC ratings since the CY13–15 period. Disruptive Technologies and Funds was the only industry group to contribute positively to IFC-wide development outcomes because of an increase in both ratings (from 26 percent to 48 percent) and share of projects (from 8 percent to 10 percent). Despite the recent improvements, Disruptive Technologies and Funds ratings are below the IFC average in CY21–23.

Manufacturing, Agribusiness, and Services and Infrastructure and Natural Resources contributed negatively to IFC's overall development outcomes over the long term (figure C.4). Manufacturing, Agribusiness, and Services contributed negatively because of a decline in both ratings (from 51 percent to 48 percent) and share of projects (from 36 percent to 32 percent). While

the ratings of Infrastructure and Natural Resources projects remained stable at 62 percent, its share of projects declined from 23 percent to 16 percent over the long term and thus contributed negatively to IFC's overall development outcomes.

Figure C.4. Decomposition Analysis by Industry Group, Calendar Years 2013–15 Versus Calendar Years 2021–23



Source: Independent Evaluation Group, Expanded Project Supervision Report database.

Note: CDF = Disruptive Technologies and Funds; FIG = Financial Institutions Group; IFC = International Finance Corporation; INR = Infrastructure and Natural Resources; MAS = Manufacturing, Agribusiness, and Services.

Advisory Services Projects

This section describes the decomposition analysis for IFC advisory services projects across various subgroups.

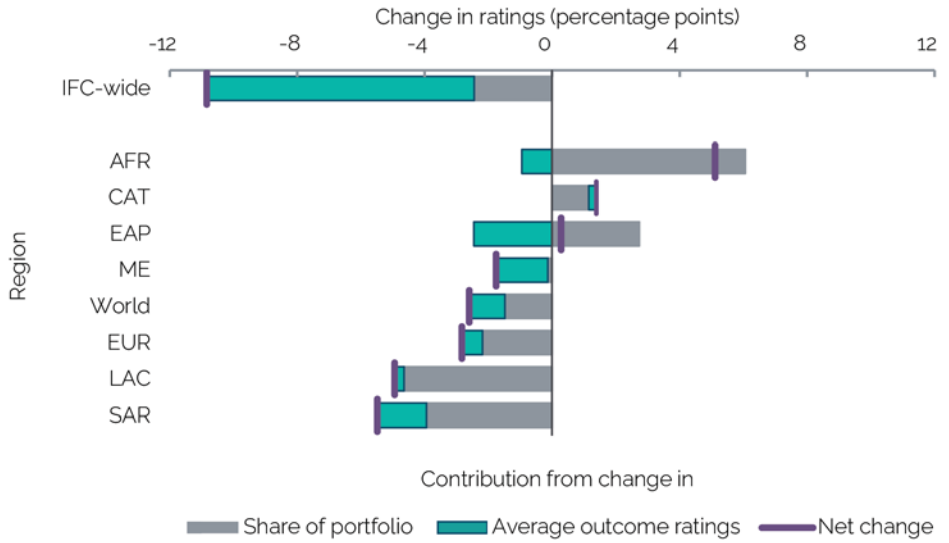
Decomposition Analysis by Region

Three regions contributed positively to IFC's overall development effectiveness over the long term: Africa, Central Asia and Türkiye, and East Asia and the Pacific (figure C.5). Africa contributed to IFC's overall development effectiveness ratings mainly due to the increase in the share of projects (from 23 percent to 36 percent). Its development effectiveness is the weakest among the regions and declined slightly (from 45 percent to 43 percent). Central Asia and Türkiye was the only region to contribute positively to IFC's overall development effectiveness because of an increase in both ratings (from 80 percent to 86 percent) and share of projects (from 3 percent to

4 percent). East Asia and the Pacific contributed positively to IFC's overall development effectiveness as its share of projects increased (from 15 percent to 19 percent), while its ratings declined (from 63 percent to 50 percent).

The remaining regions contributed negatively to IFC's overall development effectiveness. The Middle East, Europe, Latin America and the Caribbean, and South Asia all contributed negatively because of a decline in both ratings and share of projects.

Figure C.5. Decomposition Analysis by Region, Calendar Years 2013–15 Versus Calendar Years 2021–23



Source: Independent Evaluation Group, Project Completion Report database.

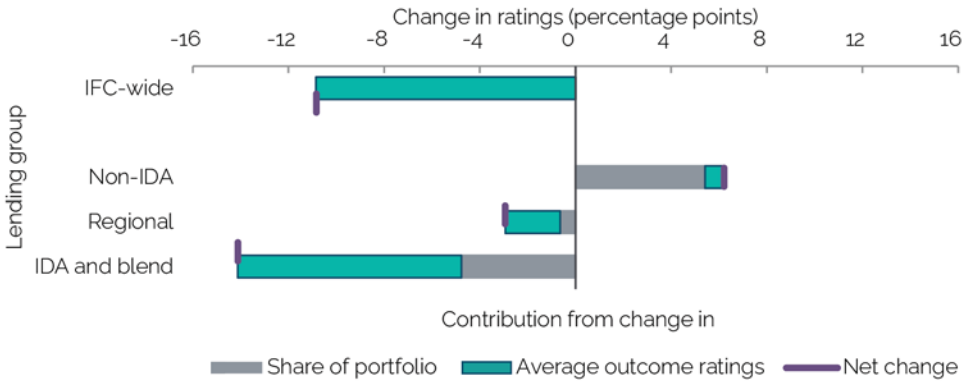
Note: AFR = Africa; CAT = Central Asia and Türkiye; EAP = East Asia and the Pacific; EUR = Europe; IFC = International Finance Corporation; LAC = Latin America and the Caribbean; ME = Middle East; SAR = South Asia; World = multiregional.

Decomposition Analysis of Trends in Challenging Environments: IDA and Blend Countries and Fragile and Conflict-Affected Situations

Among all subgroups, IDA and blend was the single largest negative contributor to IFC's overall development effectiveness over the long term, while non-IDA countries contributed positively (figure C.6). IDA and blend's negative contribution was due to a decline in both ratings (from 59 percent to 43 percent) and share of projects (from 65 percent to 57 percent).

Conversely, projects in non-IDA countries contributed positively both due to an improvement in ratings (from 60 percent to 63 percent) and an increase in share of projects (from 26 percent to 35 percent).

Figure C.6. Decomposition Analysis by Lending Group, Calendar Years 2013–15 Versus Calendar Years 2021–23

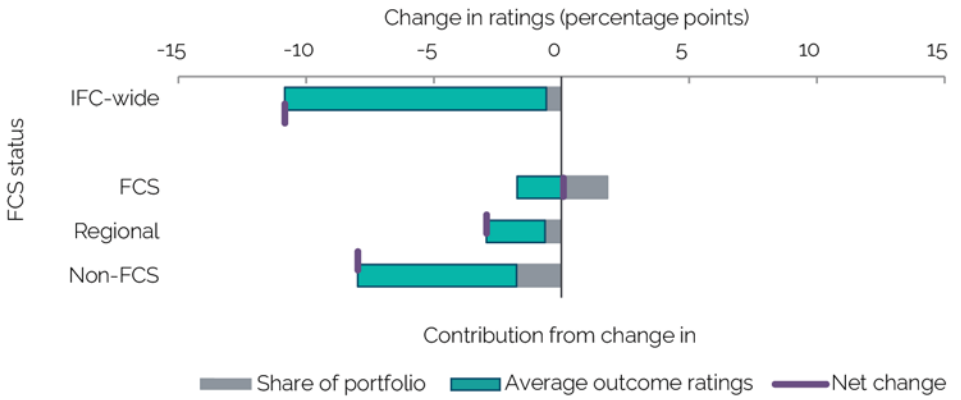


Source: Independent Evaluation Group, Project Completion Report database.

Note: IDA = International Development Association; IFC = International Finance Corporation.

The contribution of FCS was slightly positive to IFC's overall development effectiveness over the long term, while non-FCS contributions were negative (figure C.7). The positive contributions of FCS were due to an increase in the share of projects (from 15 percent to 18 percent), while its ratings declined (from 48 percent to 39 percent). Conversely, the negative contributions of non-FCS were mainly due to a substantial decline in both its ratings (from 60 percent to 53 percent) and share of projects (from 76 percent to 73 percent).

Figure C.7. Decomposition Analysis of Trends by Fragile and Conflict-Affected Situations Status, Calendar Years 2013–15 Versus Calendar Years 2021–23



Source: Independent Evaluation Group, Project Completion Report database.

Note: FCS = fragile and conflict-affected situations; IFC = International Finance Corporation.

Appendix D. International Finance Corporation Investment Outcome Types

The goal of the outcome-type analysis is to describe the intended development outcomes of International Finance Corporation (IFC) investment projects and to assess their achievement rates and their relationship with performance ratings. This analysis adds value because IFC investment projects are not rated on outcome types both in IFC's Expanded Project Supervision Reports and in Independent Evaluation Group (IEG) validations of Expanded Project Supervision Reports or IEG evaluations of projects that closed prematurely. This analysis expands on the outcome-type analysis conducted in the *Results and Performance of the World Bank Group (RAP) 2021* and *2023* reports. In *RAP 2024*, IEG conducted a deep dive of 173 IFC investment projects evaluated and validated during the calendar years 2021–23 with a cutoff date of December 31, 2023. The definitions of all outcome types are provided in appendix A.

International Finance Corporation Outcome Types

IFC's investment project outcomes fall into two broad categories: project-level outcomes and market-level outcomes. Project-level outcomes are defined as a project's direct and indirect effects on stakeholders, the economy, and the environment. Market-level outcomes are defined as a project's ability to catalyze systemic changes beyond those effects brought about by the project itself. Based on the statement of the intended development objective at approval, an individual project's objective could be classified into several outcome types. *RAP 2024* uses the same outcome-type classification that was used in *RAP 2023*, which identified 33 outcome types (28 project-level and 5 market-level outcomes) that were based on IFC's Anticipated Impact Measurement and Monitoring (AIMM) sector frameworks.

For the 173 IFC investment projects covered in this deep-dive analysis, we identified 842 individual outcomes. Of these individual outcomes, 676 were

project-level outcomes and 166 were market-level outcomes. All projects reviewed pursued project-level outcomes, and 77 percent pursued market-level outcomes. On average, projects pursued 5 different outcomes, consisting of 4 project-level outcomes and 1 market-level outcome.

Top Three Prevalent Outcomes

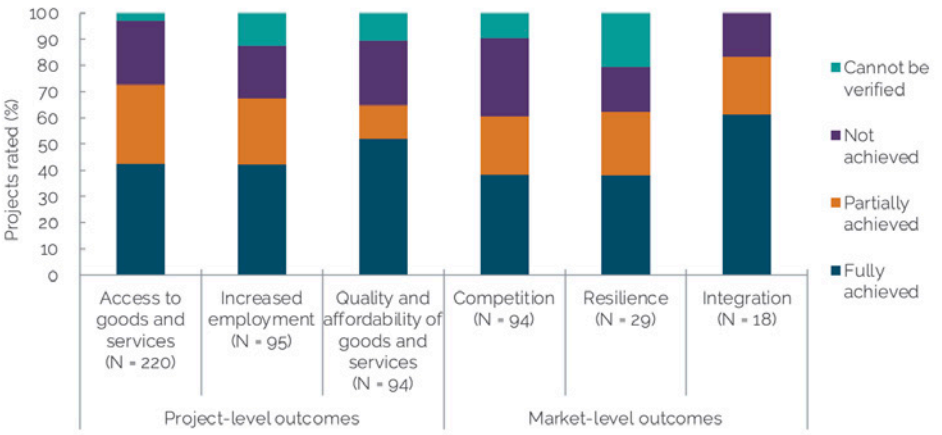
Improved access to goods and services and market competition were the most prevalent project-level and market-level outcomes. Among the project-level outcomes, access to goods and services was prevalent in 97 percent of the projects, followed by increased employment (55 percent) and quality and affordability of goods and services (54 percent). Among the market-level outcomes, competition in the market was prevalent in 54 percent of the projects, followed by resilience in the market (17 percent) and integration in the market (10 percent).

Achievement Rates of Prevalent Outcomes

Outcome achievement rates were substantially lower than what was expected at approval. IFC investment projects had 842 outcomes, of which 365 were fully achieved (43 percent), 194 were partially achieved (23 percent), 187 were not achieved (22 percent), and the remaining 96 outcomes could not be verified (11 percent). The most common reasons why an outcome could not be verified were as follows: (i) the project did not have an indicator to track the outcome, (ii) the client did not report relevant information, (iii) there was insufficient evidence to measure achievement, (iv) there was no clarity in how to measure the outcome, (v) the result could not be attributed to the project, or (vi) it was too early to tell.

Project-level outcomes. IFC investment projects had 676 project-level outcomes, of which 44 percent were fully achieved. Among the top three most prevalent project-level outcomes, access to goods and services and increased employment both had achievement rates of 42 percent (fully achieved), followed by quality and affordability of goods and services (52 percent; figure D.1).

Figure D.1. Achievement Rates of Outcomes



Source: Independent Evaluation Group.

Market-level outcomes. IFC investment projects pursued 166 market-level outcomes, of which 40 percent were fully achieved. Among the top three most prevalent market-level outcomes, competition in the market and resilience in the market both had achievement rates of 38 percent, followed by integration in the market (61 percent).

Verification of Outcome Achievement in Anticipated Impact Measurement and Monitoring

The above outcome analysis showed that 96 outcomes could not be verified, including 76 project-level outcomes and 20 market-level outcomes. IFC management claimed that the introduction of the AIMM system in 2017 overcomes the issue of verifying outcomes. To test this hypothesis, IEG conducted an analysis of a universe of 21 projects evaluated and validated by IEG with "live" AIMM scores—that is, projects that were assigned ex ante AIMM scores at Board approval and that were evaluated and validated by IEG as of December 31, 2023. Given the small cohort, this is a preliminary analysis, with the intent of setting the stage for future analysis.

This analysis found continuing challenges in identifying and tracking outcome indicators. Twenty-two percent of outcomes did not have an indicator in the tracking system (17 percent of project-level outcomes and 43 percent

of market-level outcomes).[1] Eight out of 21 projects were prepaid. After excluding 62 outcomes associated with these 8 projects, there were 108 outcomes (94 project-level and 14 market-level outcomes). Of these, 26 percent were never tracked or could not be tracked because they did not have an indicator either in the tracking system or in the Board paper (18 percent for project-level outcomes and 79 percent for market-level outcomes). Therefore, identification and tracking of outcome indicators, particularly for market outcomes, remain a challenge. This is within IFC's influence.

Association Between Outcomes and Performance Ratings

IFC investment projects with high outcome achievement rates have higher development outcome ratings. For IFC investment projects, development outcome ratings are assigned at the project level and the subdimension level (project business performance, economic sustainability, environmental and social effects, and private sector development) but not at the outcome level. The outcome analysis showed that projects that achieve their specific project-level and market-level outcomes also achieve higher development outcome ratings (table D.1). For example, projects rated successful on development outcomes achieved 84 percent of their project-level outcomes and 89 percent of their market-level outcomes. Conversely, projects rated unsuccessful achieved only 28 percent of their project-level outcomes and 17 percent of their market-level outcomes. Outcome achievement rates also tend to move in line with ratings for economic sustainability and private sector development.

Table D.1. Association Between Development Outcome Ratings and Outcome Achievement Rates

Development Outcome Rating	Total Projects (no.)	Total Outcomes (no.)	Overall Weighted Achievement Rate (%)	Project-Level Outcome Weighted Achievement Rate (%)	Market-Level Outcome Weighted Achievement Rate (%)	Economic Sustainability Average Rating	Private Sector Development Average Rating
Highly successful	0	0	—	—	—	—	—
Successful	37	168	85	84	89	3.1	3.1
Mostly successful	54	274	72	74	64	2.8	3.0
Mostly unsuccessful	37	177	44	46	34	2.1	2.4
Unsuccessful	32	142	26	28	17	1.4	1.9
Highly unsuccessful	13	81	7	7	7	1.0	1.2

Source: Independent Evaluation Group.
Note: — = not applicable.

[1] According to IFC, indicators for environmental, social, and governance outcomes are recorded and monitored separately in a different system (Sustainability Rating Tool, previously Environmental and Social Review Document) other than the AIMM system and the Development Outcome Tracking System. Therefore, 10 environmental, social, and governance outcomes (8 on environment and social, 1 on greenhouse gas emissions, and 1 on improved living standards) in this analysis are considered to have indicators and are being tracked by IFC (although in a different system other than the AIMM system and the Development Outcome Tracking System).

Appendix E. Association Among Development Outcomes, Work Quality, and Additionality in International Finance Corporation Investment Projects

This *Results and Performance of the World Bank Group (RAP)* identifies factors specific to work quality and additionality that are associated with development outcomes.[1] This analysis is particularly important not only given the strong correlation among work quality, additionality, and development outcomes but also because of the slight decline in these three performance indicators as reported in *RAP 2023*. We caution that while the empirical associations among the International Finance Corporation (IFC) work quality, additionality, and development outcome ratings are well established through correlation analysis, causality among IFC work quality, additionality, and development outcome ratings may not be possible in *RAP 2024* because both positive and negative factors beyond IFC work quality and additionality can contribute to the development outcome ratings of IFC investment projects. This analysis will help set the stage for future analysis. Based on the above analytic framework, a more balanced approach (in terms of selecting projects on both ends of the spectrum) can be followed in future analysis.

Association Between Work Quality and Development Outcomes

The Independent Evaluation Group conducted a desk-based review of a universe of 19 projects whose work quality was rated unsatisfactory or development outcomes were rated highly unsuccessful in calendar years 2021–23 (with a cutoff date of December 31, 2023). The aim was to help identify the top three work quality factors that are associated with weak development outcomes and provide a synthesis of evidence on these issues. Because there is only one IFC investment project rated excellent on work quality or highly

successful on development outcome, this project was excluded from the analysis. Hence, the cohort consisted of 19 projects with projects rated highly unsuccessful on development outcomes or unsatisfactory on work quality (table E.1). This analysis will help set the stage for future analysis. Based on this analytic framework, a more balanced approach (in terms of selecting projects on both ends of the spectrum) can be followed in future analysis.

Table E.1. Sample for Analysis of Association Between Work Quality and Development Outcomes

		Development Outcome					
		HU	US	MU	MS	SU	HS
Work quality	E	0	0	0	0	1	0
	S	0	8	9	41	25	0
	PU	10	18	25	11	8	0
	U	4	4	1	0	0	0

Source: Independent Evaluation Group, Expanded Project Supervision Report database.

Note: The shaded region indicates International Finance Corporation investment projects rated highly unsatisfactory on development outcomes or rated unsatisfactory on work quality by the Independent Evaluation Group. E = excellent; HS = highly successful; HU = highly unsuccessful; MS = mostly successful; MU = mostly unsuccessful; PU = partly unsatisfactory; S = satisfactory; SU = successful; U = unsatisfactory; US = unsuccessful.

A desk-based review was conducted to identify the most prevalent work quality factors associated with weak development outcomes. For the purposes of this analysis, we used the same taxonomy that was used to identify factors linked to performance. The application of this taxonomy at the time of the desk-based review allowed us to identify the most prevalent work quality factors that are associated with weak development outcomes. These work quality factors were peer-reviewed by the *RAP 2024* team to ensure consistency and accuracy.

This analysis identified 185 work quality issues in the 19 projects under review. The majority of issues identified (119) had a negative influence on work quality, while 46 had a positive influence. The remaining 20 issues were mostly related to environmental and social performance standards and had no (or a neutral) influence on work quality. To identify the top three work quality issues, the net count (negative influence minus positive influence) was calculated. The identified top three work quality issues were market

assessment (15 occurrences, all negative); assumptions, financial models, and project cost (9 occurrences, all negative); and client quality (14 negative occurrences, 4 positive occurrences). All three are controllable by IFC. This finding suggests that it was largely IFC-controllable work quality shortfalls that led to project failure in the 19 projects with inadequate development outcomes or inadequate work quality assessments.

The following sections give a detailed explanation of the three most prevalent work quality issues and also highlight some project examples to give a better sense of where IFC might want to focus efforts to improve the project preparation and approval process.

Market Assessment

Market assessment is defined as work quality issues related to issues in assessing the characteristics of the market to gauge the viability of the project. This includes market size and composition—that is, information about who buys what and how much and at what prices; how prices affect volumes (demand curve analysis); and what product features matter in demand and pricing. It also includes an assessment of competition; main product substitutes; bargaining power of buyers and suppliers; and how regulations, tariffs, imports, and other factors affect market and project risks. Market assessment is almost always done as part of IFC appraisal, and it underpins the projections, sensitivity analysis, and risk assessment for a project and sets the stage for project structuring. For example, a missing competitive analysis led to the overestimation of a client's sales in one project. Equally, in another project, the lack of pricing or demand analysis led to IFC overinvesting before the investment from the strategic investor, which withdrew its commitment because of country economic concerns and thereby left IFC with a nonviable project. In a farming project, the lack of an assessment of alternative fertilizers in the market led to a pricing structure that farmers were not willing to pay.

Client Quality

Client quality is defined as work quality issues related to assessing the ability, technical expertise, and track record of the sponsor or management team.

Out of 19 projects reviewed, 11 were negative and 4 were positive in work quality related to client quality. For example, in an investment fund project, while the fund sponsor missed most targets under its first fund, IFC invested in the follow-on fund, which also missed its targets. The follow-on fund was terminated, and IFC lost money. In the second example, the sponsor lacked experience and could not execute real estate acquisitions required to make the strategy on assisted living facility expansion in Latin America and the Caribbean work, nor could they price services to meet the market requirements, missing all targets.

Assumptions, Financial Models, and Project Costs

Assumptions, financial models, and project costs is defined as work quality issues related to the work quality in the assumptions, financial models, and project costs used in projecting development and financial outcomes presented in Board papers. Note that assumptions should be based on market assessments, and where these are missing, assumptions are often not grounded in market reality. In all cases, the actual results diverged significantly from projections in the Board paper. For example, in four out of nine cases, assumptions were not based on feasibility studies or market assessments. In one case, feasibility studies were not conducted on new production technologies before project rollout in a copper mine in Africa. In the remaining three cases, expected market prices were not tested, and demand was not carefully assessed with surveys. In three out of nine cases, projections should have been accompanied with downside scenarios. In these cases, assumptions were based on everything going right in very complex projects without any contingencies.

Association Between Additionality and Development Outcomes

The Independent Evaluation Group conducted a desk-based review of a universe of 18 projects whose additionality was rated unsatisfactory or development outcomes were rated highly unsuccessful in calendar years 2021–23 (with a cutoff date of December 31, 2023).

A review was conducted to analyze the frequency of occurrence of financial and nonfinancial additionality anticipated at Board approval. *RAP* also assesses the achievement rates of additionality—realized (fully or partially) or unrealized—and the frequency of the occurrence of missed additionality, which is additionality not anticipated in the Board paper ("missed") but materialized or not materialized and achieved by IFC. *RAP 2024* focuses only on cases that were missed but materialized. Because there are only 2 IFC investment projects rated excellent on additionality or highly successful on development outcome, these projects were excluded from the analysis. Hence, the cohort consists of the 18 projects with the negative extremes of highly unsuccessful development outcomes or unsatisfactory assessments for additionality (table E.2).

Table E.2. Sample for Analysis of Association Between Additionality and Development Outcomes

		Development Outcome					
		HU	US	MU	MS	SU	HS
Additionality	E	0	0	0	0	0	0
	S	2	0	0	0	0	0
	PU	6	0	0	0	0	0
	U	6	2	1	1	0	0

Source: Independent Evaluation Group, Expanded Project Supervision Report database.

Note: The shaded region indicates International Finance Corporation investment projects rated highly unsatisfactory on development outcomes or rated unsatisfactory on additionality by the Independent Evaluation Group. E = excellent; HS = highly successful; HU = highly unsuccessful; MS = mostly successful; MU = mostly unsuccessful; PU = partly unsatisfactory; S = satisfactory; SU = successful; U = unsatisfactory; US = unsuccessful.

To identify key issues concerning work quality, a desk-based review of Independent Evaluation Group evaluations of these 18 projects was conducted to code additionality (financial, nonfinancial, and missed). These results were peer-reviewed to ensure consistency and accuracy.

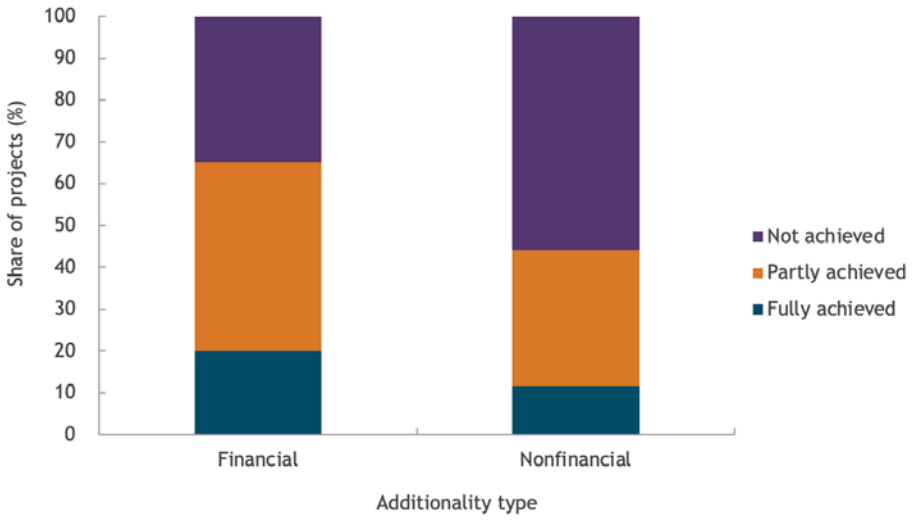
This analysis will help set the stage for future analysis. Based on the above analytic framework, a more balanced approach (in terms of selecting projects on both ends of the spectrum) can be used in future analysis.

Prevalence and Achievement Rates of Additionality

A total of 69 additionalities were identified in the 18 projects, 20 of which were financial additionalities, 46 nonfinancial, and 3 missed.

The combined achievement rate for fully or partially achieved IFC additionality (financial and nonfinancial combined) was 46 percent (32 out of 69 additionalities). However, the achievement rate varies by additionality type (figure E.1). Nonfinancial additionality occurred 46 times, with an achievement rate of 41 percent (19 out of 46 nonfinancial additionalities), while financial additionality occurred 20 times, with a success rate of 65 percent (13 out of 20 financial additionalities). Unforeseen additionality occurred only 3 times and does not, by definition, have an achievement rate.

Figure E.1. Achievement Rate of Additionality Types



Source: Independent Evaluation Group, Expanded Project Supervision Report database.

Among the financial additionalities, financing was the most frequent, occurring 13 times. However, financing had the lowest achievement rate (54 percent of occurrences were fully or partially achieved) among the top financial additionalities. Resource mobilization was the second most prevalent financial additionality and occurred 4 times, with an achievement rate of 75 percent. Risk mitigation occurred twice, with a 100 percent achievement rate.

Among the nonfinancial additionalities, improving environmental and social standards was the most frequent additionality and occurred 15 times, with an achievement rate of 60 percent. Stamp of approval was the second most prevalent, with 8 occurrences and a 50 percent achievement rate. Network sharing was the third most prevalent nonfinancial additionality, as it occurred 5 times, with an achievement rate of 20 percent.

[1] Development outcome is rated on a six-point scale: highly successful, successful, mostly successful, mostly unsuccessful, unsuccessful, and highly unsuccessful. For additionality and work quality, the rating is on a four-point scale: excellent, satisfactory, partly unsatisfactory, and unsatisfactory.

Appendix F. Project Examples of Factors Linked to International Finance Corporation Investment Project Performance

This appendix provides additional information on the top prevalent factors linked to performance as detailed in the report: (i) ability, technical expertise, and track record (referred to as "client quality"); (ii) business risk; (iii) economic issues; (iv) asset quality; and (v) civil unrest. Project examples of what worked and what did not work are given for each factor.

Client Quality

The most prevalent factor is client quality. Client quality is the ability, technical expertise, or track record of the International Finance Corporation (IFC) client. It refers to the quality of the management team and their skills, contractor competency, familiarity, and acumen. This factor occurred in 41 percent of projects reviewed. It usually has a positive influence on IFC investment development outcomes, except for projects in Africa and in countries classified as fragile and conflict-affected situations.

Client Quality: Project Examples of What Worked

By forming a joint venture with an experienced European development institution as a general partner, the first-time fund managers of a growth equity fund in East Asia and the Pacific reduced the perceived riskiness of the fund, helping catalyze capital commitments. With the right local team and structure, as well as strong sponsor commitment, private equity investing in small and medium enterprises proved to be successful and can serve as a good model to replicate, particularly for first-time fund managers.

In the second example, strategic alignment, sponsor selection, a financing plan, design, appraisal, and risk assessment were all present in putting together a project for an Eastern European bank's acquisition of a rival

institution. The choice of a sponsor with a strong commitment to micro, small, and medium enterprise financing was the single most important factor in the project's success. Consequently, the client achieved much better micro, small, and medium enterprise loan volume growth than was seen in the local banking sector overall. The acquisition allowed the client to become the micro, small, and medium enterprise market leader in the country, with a 20 percent market share.

Client Quality: Project Examples of What Did Not Work

In a health-care project in Africa, an overly risky business strategy was pursued by the client, a pharmaceutical company that attempted to grow by acquiring two sizable companies outside of its region. The debt used to finance these two and other acquisitions resulted in unsustainable financial costs and forced the company to go for business restructuring. While the sponsor had a substantial track record in closing acquisitions across several sectors, it lacked experience outside Africa. IFC did not put any restrictive covenants on the sponsor, despite conflicts of interest among the various investments of the sponsor, resulting in substantial losses for IFC.

In the second example, undisclosed related-party lending and management misconduct negatively affected the performance of a banking project in Africa. While bank management and shareholders embezzled funds for their own use, IFC failed to perform adequate governance review of the bank, despite the fact that it had noted governance weaknesses two years before project approval. The client bank was placed under receivership 14 months after the first IFC tranche was disbursed, and its assets shrank by 40–50 percent because of significant loan losses and business runoff. IFC ultimately wrote off 100 percent of its loan in the project.

Business Risk

Business risk refers to risks related to business model, cyclical business, or the operating environment. Business risk usually has a negative influence on IFC investment development outcomes, and it occurred in 25 percent of projects reviewed.

Business Risk: Project Examples of What Worked

Thorough preparation of the project by IFC helped mitigate most business risks in a railway infrastructure project in Africa. IFC adequately identified mining projects with the potential to increase demand for railway services by the client and helped the client renegotiate and amend a concession agreement with the government. The amendment established a nondiscriminatory access tariff regime designed to boost multiuser access, including passenger services, and allowed full cost recovery. In sum, the project's strong preparation translated into good performance.

In the second example, a project in Latin America and the Caribbean supported a local financial institution in expanding its small and medium enterprise lending to farmers. Being professionally managed, the client successfully navigated a difficult business environment. Harsh conditions compelled the client to reduce operating costs during an economic downturn to cope with lower revenues, and such reductions ultimately created efficiency gains later on. The company's credit rating system, developed in-house, allowed it to revise its financing strategy by extending maturities that helped its clients. Ultimately, the company survived the downturn, whereas one of its biggest competitors went out of business because of liquidity problems.

Business Risk: Project Examples of What Did Not Work

The untested business model and wrong market trends assumptions led to the failure of a fertilizer distribution company in the East Asia and the Pacific region. The company had overinvested in creating service centers, which was a costly exercise that did not result in the expected efficiencies. At the same time, its decision to cancel construction of own production facilities proved to be the wrong strategy, as the company missed out on potential cost advantages. The sponsor had not fully grasped the megatrends in fertilizer use in agriculture—that is, lower use and more customization of client solutions.

In the second example, IFC invested in a logistics service provider in the East Asia and the Pacific region that operated warehouses and distribution hubs, delivering parcels and freight within the e-commerce sector. The company tried to increase its market share in a complex, concentrated, and highly competitive e-commerce

logistics sector. The competitors started using a full franchise model, whereas the company's capital expenditure on logistics was higher and it only used franchisees for the last-mile delivery, a less profitable model. High fixed costs and fierce pricing competition depressed the company's earnings.

Economic Issues

Economic issues are defined as risks related to the macroeconomic environment, inflation, monetary policy, or austerity measures. This factor almost always has a negative influence on IFC investment development outcomes. Economic issues occurred in 24 percent of projects reviewed.

Economic Issues: A Project Example of What Worked

Strong economic growth, stable currency, and good management supported the performance of an IFC investment in a microfinance company in the East Asia and the Pacific region. The company exhibited significant loan growth while maintaining the quality of its loan portfolio and attaining adequate profitability levels. This was achieved partly due to the country's remarkable economic growth of 7 percent or more per year between 2011 and 2018, coupled with a relatively stable exchange rate and adequate loan portfolio management by the company.

Economic Issues: Project Examples of What Did Not Work

IFC provided a loan to a leading dairy and food producer in the Middle East and North Africa region to fund its expansion into another country in the region. The project was only partially implemented due to adverse macroeconomic conditions that forced the company to reduce its capital expenditures. IFC had to put on hold its disbursement of the second tranche under the loan agreement. Currency depreciation was the strongest negative factor affecting the company's performance.

In the second project, IFC invested in a company to expand its production and distribution capacity across a country in the Latin America and the Caribbean region. While the project was implemented as planned, the company lagged behind on margins and cash generation projections. This was mainly due to external headwinds, especially currency depreciation and

input price hikes. The currency depreciated by more than 30 percent since the investment was disbursed, with a material impact on the financials—that is, about 80 percent of the company debt was denominated in US dollars. Similarly, key inputs were linked to US prices.

Asset Quality

Low asset quality, such as a rise in the nonperforming loans of a client company, always has a negative influence on IFC investment development outcomes. This challenge occurred in 25 percent of International Development Association and blend projects in countries classified as fragile and conflict-affected situations that were reviewed.

Asset Quality: A Project Example of What Did Not Work

An IFC client's nonperforming loans increased because of the deteriorating performance in a sector that was a key driver of growth for the country. Performance in this sector dropped as a result of the liquidation of one of the largest traders, which had systemic exposure to all top-tier banks in the country. This led to contagion across the banking sector. Ultimately, the contagion significantly increased the number of nonperforming loans for the IFC client. As a result, the client did not achieve its target of increasing its small and medium enterprise loan portfolio.

Civil Unrest

This factor almost always has a negative influence on IFC investment development outcomes. Civil unrest occurred in 21 percent of International Development Association and blend projects in countries classified as fragile and conflict-affected situations.

Civil Unrest: A Project Example of What Did Not Work

An IFC client, a microfinance institution in South Asia, was affected by the challenging macroeconomic situation created by the double impact of civil unrest and COVID-19. Consequently, the local currency depreciated by about one-third, with inflation soaring. As a result of the macroeconomic downturn, the loans outstanding declined significantly.

Appendix G. Decomposition Analysis of Multilateral Investment Guarantee Agency Project Rating Trends

This is a technical appendix on decomposition analysis for Multilateral Investment Guarantee Agency (MIGA) guarantee projects. The development outcome ratings used in the decomposition analysis are as of June 30, 2024.

Decomposition analysis shows the net contribution from each subgroup to the changes in MIGA's overall development outcome ratings over the long term. This section describes the decomposition analysis for MIGA guarantee projects across various subgroups. The subgroups in the decomposition analysis refer to region, International Development Association (IDA) and blend countries, countries classified as fragile and conflict-affected situations (FCS), and industry group. The long term in the decomposition analysis refers to MIGA guarantee projects evaluated and validated by the Independent Evaluation Group in two time periods: FY 2013–18 and FY18–23.

Decomposition Analysis by Region

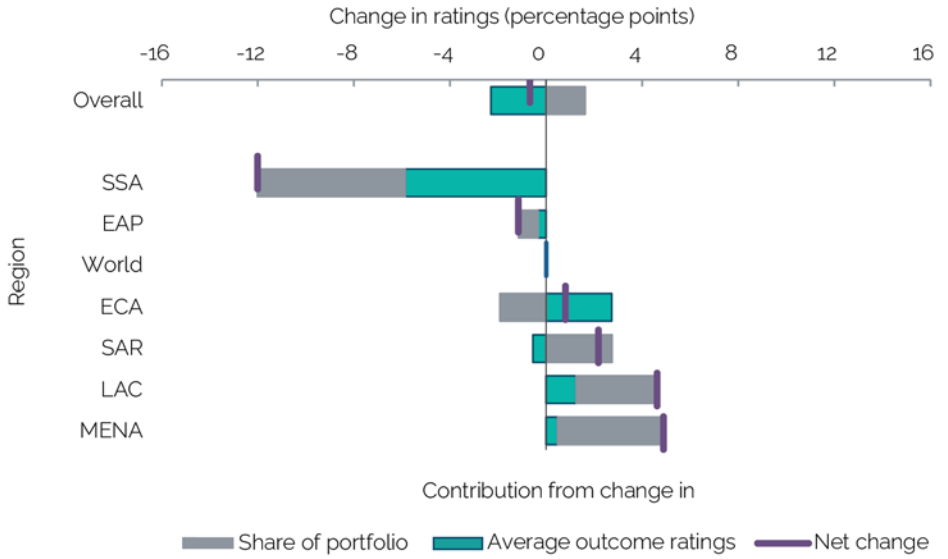
Decomposition analysis shows that two regions contributed positively to MIGA-wide development outcomes both due to a positive contribution from ratings and the increase in share of projects: Middle East and North Africa and Latin America and the Caribbean. Middle East and North Africa's net contribution to MIGA's overall development outcomes was the strongest among regions. This was due to a positive contribution from increases both in ratings (from 86 percent to 89 percent) and in the share of projects (from 10 percent to 15 percent). Latin America and the Caribbean also contributed positively because of an increase both in ratings (from 63 percent to 70 percent) and in the share of projects (from 11 percent to 17 percent).

While contributing positively on a net basis, South Asia's ratings declined, but its share of projects increased; the share of Europe and Central Asia declined, while its ratings increased. The positive contribution from South

Asia was mainly due to the increase in the share of projects (from 4 percent to 8 percent), while its ratings declined (from 67 percent to 60 percent). The positive contribution from Europe and Central Asia was due to the increase in ratings (from 73 percent to 82 percent), while the share of projects declined (from 31 percent to 28 percent).

Sub-Saharan Africa, the largest of MIGA's regions, and East Asia and Pacific, the smallest of MIGA's regions, contributed negatively to MIGA-wide development outcomes (figure G.1). The negative contribution from Sub-Saharan Africa was due to a decline in both ratings (from 72 percent to 50 percent) and the share of projects (from 35 percent to 27 percent). East Asia and Pacific contributed negatively to MIGA's overall development outcomes because of a decline in both ratings (from 40 percent to 33 percent) and the share of projects (from 7 percent to 5 percent).

Figure G.1. Decomposition Analysis by Region, FY13–18 Versus FY18–23



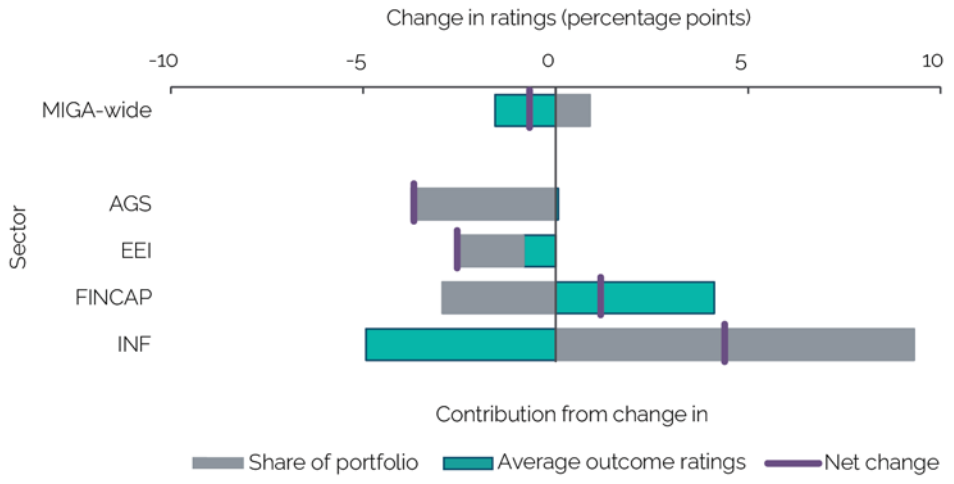
Source: Independent Evaluation Group, Project Evaluation Report database.

Note: EAP = East Asia and Pacific; ECA = Europe and Central Asia; LAC = Latin America and the Caribbean; MENA = Middle East and North Africa; SAR = South Asia; SSA = Sub-Saharan Africa; World = multiregional.

Decomposition Analysis by Multilateral Investment Guarantee Agency Sector

The Infrastructure and Finance and Capital Markets sectors contributed positively to MIGA's overall development outcomes, while Energy and Extractive Industries and Agribusiness and General Services contributed negatively (figure G.2). At the MIGA sector level, Infrastructure contributed the most to MIGA-wide development outcomes over the long term. This positive contribution was due to the increase in the share (from 32 percent to 45 percent), while ratings declined (from 74 percent to 63 percent). Finance and Capital Markets also contributed positively because of an increase in the ratings (from 58 percent to 77 percent), while the share of projects declined (from 27 percent to 22 percent). Conversely, Energy and Extractive Industries contributed negatively to MIGA-wide development outcomes because of a decline in both ratings (from 75 percent to 50 percent) and the share of projects (from 6 percent to 3 percent). Similarly, Agribusiness and General Services contributed negatively, as its share of projects declined (from 35 percent to 30 percent), while its ratings remained stable at 72 percent.

Figure G.2. Decomposition Analysis by Multilateral Investment Guarantee Agency Sector, FY13–18 Versus FY18–23



Source: Independent Evaluation Group, Project Evaluation Report database.

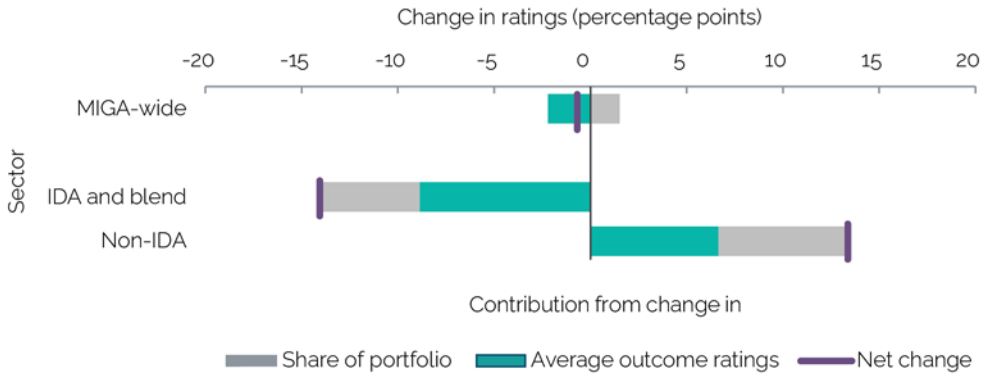
Note: AGS = Agribusiness and General Services; EEI = Energy and Extractive Industries; FINCAP = Finance and Capital Markets; INF = Infrastructure; MIGA = Multilateral Investment Guarantee Agency.

Decomposition Analysis by International Development Association and Blend Countries

Among all subgroups, IDA and blend was the single largest negative contributor to MIGA's overall development outcomes over the long term. This was due to a decline in both ratings (from 74 percent to 50 percent) and the share of projects (from 45 percent to 37 percent).

Non-IDA countries contributed positively to MIGA's overall development outcomes over the long term (figure G.3). Projects in non-IDA countries have performed below those in IDA and blend countries in the long term. However, in recent years, performance of non-IDA has improved and is above IDA and blend. The decomposition analysis shows that non-IDA contributed positively to MIGA's overall development outcomes because of an increase in both ratings (from 68 percent to 79 percent) and the share of projects (from 55 percent to 63 percent).

Figure G.3. Decomposition Analysis by Lending Group, FY13–18 Versus FY18–23



Source: Independent Evaluation Group, Project Evaluation Report database.

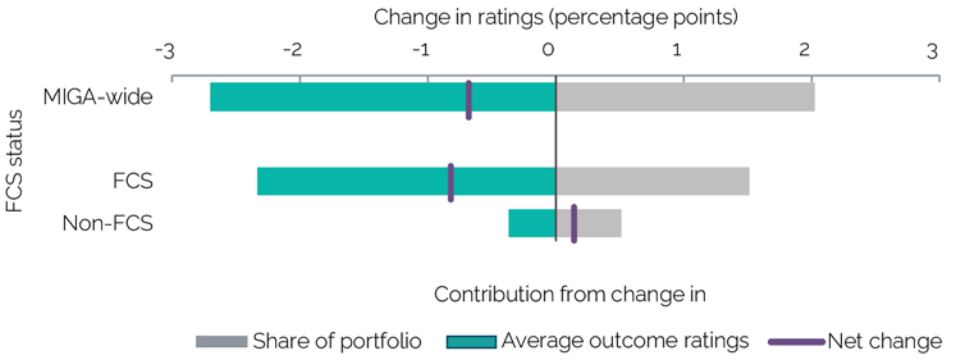
Note: IDA = International Development Association; MIGA = Multilateral Investment Guarantee Agency.

Decomposition Analysis by Fragile and Conflict-Affected Situations

FCS contributed negatively to MIGA's overall development outcomes, while non-FCS contributed positively (figure G.4). FCS contributed negatively to MIGA's overall development outcomes because of a decline in ratings

(from 78 percent to 75 percent), while the share of projects increased (from 12 percent to 13 percent). Projects in non-FCS contributed positively to MIGA's overall development outcomes because of an increase in the share of projects (from 85 percent to 87 percent), while ratings declined (from 70 percent to 67 percent).

Figure G.4. Decomposition Analysis by Fragile and Conflict-Affected Situations Status, FY13–18 Versus FY18–23



Source: Independent Evaluation Group, Project Evaluation Report database.

Note: FCS = fragile and conflict-affected situations; MIGA = Multilateral Investment Guarantee Agency.

Appendix H. Multilateral Investment Guarantee Agency Guarantee Outcome Types

The goal of the outcome-type analysis is to describe the intended development outcomes of Multilateral Investment Guarantee Agency (MIGA) projects and to assess their achievement rates. This analysis adds value because MIGA projects are not rated on outcome types both in MIGA's Project Evaluation Reports and in the Independent Evaluation Group's validations of Project Evaluation Reports. This analysis expands on the outcome-type analysis conducted in the *Results and Performance of the World Bank Group* (*RAP*) *2021* and *2023* reports. In *RAP 2024*, we conducted a deep dive of 15 MIGA guarantee projects evaluated and validated during FY 2021–23, with a cutoff date of December 31, 2023. The results from this analysis should be treated with caution given the small number of guarantee projects evaluated by MIGA and validated by the Independent Evaluation Group.

Multilateral Investment Guarantee Agency Outcome Types

MIGA's project outcomes fall into two broad categories: project-level outcomes and foreign investment–level outcomes. Project-level outcomes are defined as a project's direct and indirect effects on stakeholders, the economy, and the environment. Foreign investment–level outcomes are defined as a project's ability to catalyze systemic changes beyond those effects brought about by the project itself. Based on the statement of the intended development objective at approval, an individual project's objective could be classified into several outcome types. *RAP 2024* uses the same outcome-type classification as was used in *RAP 2023*, which identified 30 outcome types (25 at the project level and 5 at the foreign investment level) for MIGA projects that were based on the Impact Measurement and Project Assessment Comparison Tool system.

For the 15 MIGA guarantee projects covered in the deep-dive analysis, this *RAP* identified 76 individual outcomes, 58 of which were project-level outcomes and 18 of which were foreign investment–level outcomes. There were, on average, 5 outcomes per project, consisting of 4 project-level outcomes and 1 foreign investment–level outcome.

Top Three Prevalent Outcomes

Increased employment and market development were the most prevalent project-level and foreign investment–level outcomes. Among the project-level outcomes, increased employment was prevalent in 73 percent of MIGA guarantee projects, followed by access to goods and services for customers (60 percent) and increased transfers to the government (53 percent). Among the foreign investment–level outcomes, development reach and signaling effects were both prevalent in 33 percent of MIGA guarantee projects, followed by market development (27 percent).

Achievement Rates of Prevalent Outcomes

Outcome achievement rates were substantially lower than what was expected at approval. MIGA guarantee projects had 76 outcomes, 37 of which were fully achieved (49 percent), 22 partially achieved (29 percent), and 4 (5 percent) not achieved, while the remaining 13 outcomes (17 percent) could not be verified.

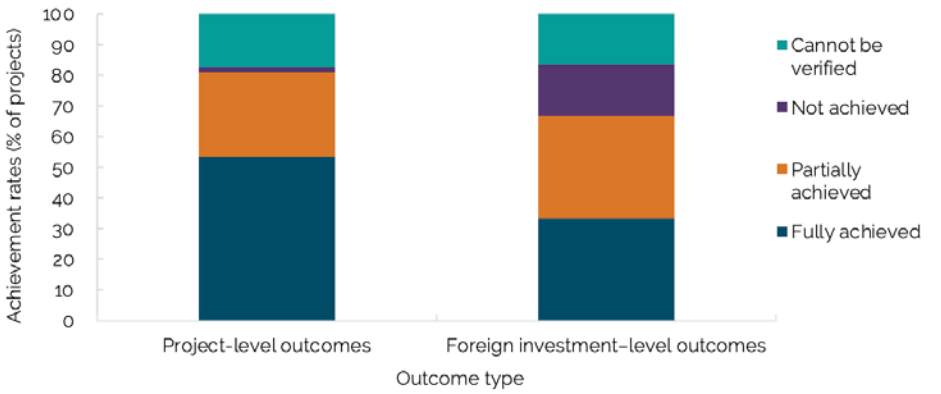
Project-level outcomes. MIGA guarantee projects had 58 project-level outcomes, of which 53 percent were fully achieved. Among the top three prevalent project-level outcomes, increased employment had the highest achievement rate of 64 percent (fully achieved), followed by access to goods and services for customers (56 percent) and increased transfers to the government (25 percent).

Foreign investment–level outcomes. MIGA guarantee projects pursued 18 foreign investment–level outcomes, of which 33 percent were fully achieved (figure H.1).

Among the top three prevalent foreign investment–level outcomes, signaling effect had an achievement rate of 60 percent, followed by development

reach (40 percent) and market development (25 percent). Foreign investment–level outcomes are achieved less often than project-level outcomes for MIGA guarantee projects, which is a concern given MIGA's business model. Given that MIGA's business model involves guaranteeing foreign direct investments, one would expect that it would be at least as good at achieving foreign investment–level outcomes as it is at achieving project-level outcomes. The difference in the achievement rates cannot be due to the breadth of the foreign investment–level outcomes because the share of project-level outcomes (10 out of 58 outcomes, or 17 percent) that cannot be verified is the same as the share of foreign investment–level outcomes (3 out of 18 outcomes, or 17 percent) that cannot be verified.

Figure H.1. Achievement Rate of Outcomes



Source: Independent Evaluation Group.

Appendix I. Project Examples of Factors Linked to Multilateral Investment Guarantee Agency Guarantee Project Performance

This appendix provides additional information on the top prevalent factors linked to performance. The direct client of the Multilateral Investment Guarantee Agency (MIGA) is its guarantee holder, which is the company benefiting directly from the MIGA guarantee. The guarantee holder has direct influence over a project company, which owns and implements the project. Project company quality refers to the ability, technical expertise, and track record of the project company.

Results and Performance of the World Bank Group 2024 conducted a deep-dive analysis of 26 MIGA guarantee projects evaluated and validated by the Independent Evaluation Group for FY 2020–23 using the *Results and Performance of the World Bank Group 2023* taxonomy to identify the top factors linked to development outcomes. A factor can have both positive and negative links to development outcomes. The most prevalent factors linked to development outcomes of MIGA are related to (i) cost overruns and construction delays; (ii) ability, technical expertise, and track record (referred to as "project company quality"); and (iii) legal and regulatory risk. This appendix provides project examples of what worked and what did not work for each of the three most prevalent factors.

Cost Overruns and Construction Delays

The most prevalent factor is cost overruns and construction delays. Cost overruns and construction delays are generally outside of MIGA's direct influence. This is a negative factor in two-thirds of the projects, but due to a limitation in the taxonomy, it can also be a positive factor when the project is executed under budget. MIGA's job stops at underwriting, and its business model includes only limited supervision (involving only environmental and

social assessment) of projects. In that context, MIGA does not directly influence cost overruns or construction delays.

Cost Overruns and Construction Delays: Project Examples of What Worked

MIGA issued a guarantee against the risks of transfer restriction, expropriation, war and civil disturbance, and breach of contract for the construction of a power plant. The project cost at completion was 10 percent less than at approval due to the strong construction and procurement processes of the project company. The project established a successful track record and helped the country continue attracting local and global private investors.

In the second example, MIGA supported the construction of a natural gas pipeline in a country by providing a guarantee for bank loans to the sponsor. The almost 2,000-kilometer-long pipeline was constructed at 45 percent lower cost than the original budget. The pipeline has been transporting gas, reaching a level of 118 percent of the projected amounts of natural gas.

Cost Overruns and Construction Delays: Project Examples of What Did Not Work

MIGA issued a guarantee to cover the equity investment in the development of a hydrocracking and coking facility. The facility would upgrade locally available fuels into higher-value lighter-fuel products for the domestic market. The project was completed with a significant delay of three years and cost overruns of 27 percent above the budget. Apparently, the engineering, procurement, and construction contractor's capacity was substandard for such a large-scale and complex project. The completion delay led to lower financial and economic returns.

In the second example, MIGA issued a breach of contract guarantee for loans to fund a project to build a power plant. The project was delivered with a 10-month delay that was mainly attributed to the engineering, procurement, and construction contractor's failure to complete the construction within the agreed 14 months and the replacement of the incompetent initial operations and maintenance contractor. The capacity of the completed plant was

lower than planned, and during the construction delay, other more efficient power plants came online, reducing the use of the project.

In the third example of cost overruns and construction delays, MIGA issued a guarantee covering the risk of Non-Honoring of Financial Obligations to a financial institution that had extended a loan for the acquisition of rolling stock for a metro line. The construction of the metro line was completed with a delay of one year, and the project had 30 percent lower-than-expected passenger volumes, which resulted in the project's weak viability and lower prospects for growth.

Project Company Quality

MIGA can indirectly influence project company quality by selecting the right guarantee holder. Project company quality is defined as the quality and experience of the management team implementing the project, as well as their technical skills, track record, contractor competency, familiarity, and acumen. MIGA uses factors such as management experience, sector experience, and commitment to select a guarantee holder, which, in turn, selects a project company to implement a project financed by the guarantee holder. Project company quality has a positive influence on MIGA's development outcomes in 75 percent of projects.

Project Company Quality: Project Examples of What Worked

MIGA issued a guarantee covering the risk of transfer restriction, expropriation, and war and civil disturbance for the guarantee holder's equity investment in the construction and operation of a steel dust recycling facility. Steel mills in the country used to dispose of steel dust without proper treatment, exacerbating the risks of pollution and contamination. The project company chose the best available and most energy-efficient steel dust recycling technology, reducing the pollution generated by the steel sector.

In the second example, MIGA issued a guarantee to cover against the risks of transfer restriction, expropriation, and breach of contract of the equity investment into the design, construction, financing, and maintenance of a greenfield integrated health-care facility. Construction was completed well

within budget (about 3 percent below the original cost) and six months earlier than planned, which reflects the experience acquired by the sponsor from implementing previous projects in the same country. The objective of establishing a modern hospital with specialized facilities was seamlessly achieved.

Project Company Quality: A Project Example of What Did Not Work

In a refinery construction project, MIGA issued a guarantee to cover a sponsor's equity investment. The sponsor had no experience with building refineries or negotiating construction contracts. This lack of experience led to project cost estimates exceeding industry benchmarks. The contractor also took advantage of the sponsor's lack of sector expertise during the execution phase of the project, which added to the project's cost overruns.

Legal and Regulatory Risk

Legal and regulatory risk is outside of MIGA's influence. Legal and regulatory risk is defined as risk related to regulatory policies, government, legislation, and bureaucratic mechanisms. This factor has a positive influence on development outcomes in about 85 percent of MIGA projects. However, it has a negative influence in about 15 percent of projects.

Legal and Regulatory Risk: A Project Example of What Worked

MIGA provided a guarantee against the risks of transfer restriction, expropriation, and breach of contract to an investor for the construction of a greenfield hospital. Construction was completed on time and on budget. The most important element in ensuring the project's success was the revised payment methodology that had been agreed to with the government. The renegotiation helped ensure timely debt service payments and a reasonable return to the investors while making payments to the project more predictable for the government. This revised payment mechanism improved the financial stability of the project and ensured a more sustainable payment plan, a critical goal given the country's high inflation environment and the volatility of its currency.

Legal and Regulatory Risk: A Project Example of What Did Not Work

MIGA issued a guarantee to provide cover for the risks of transfer restriction, expropriation, war and civil disturbance, and breach of contract to the sponsor of a project involving the construction, operation, and maintenance of a greenfield bulk water treatment facility. Even though the project was able to increase the volume of treated water, the sustained high level of water loss meant that the project's development outcome was not achieved. The cost of the project's water was also higher than end-user tariffs, resulting in subsidies. These represent economic costs that weakened the project's economic return and sustainability.

Appendix J. One World Bank Group Collaboration

This appendix provides an analysis of the ways in which the World Bank Group collaborates based on Independent Evaluation Group evidence. Breaking down how the World Bank Group collaborates is important, as attempts to increase collaboration among the World Bank, the International Finance Corporation (IFC), and the Multilateral Investment Guarantee Agency have been made since at least 1989.

Methodology

For this analysis, we systematically analyzed completed reports that provide direct evidence of One World Bank Group collaboration. The method employed three steps: identification, analysis, and synthesis.

The identification step consisted of downloading all documents that met specific criteria. Independent Evaluation Group reports were included that provided either a country-level analysis or a focused assessment of Bank Group collaboration that covered the FY 2013–23 period. This led to the inclusion of 147 Completion and Learning Review Validations, all reports from the *Results and Performance of the World Bank Group* series, 22 Country Program Evaluations, one evaluation of joint projects, one evaluation of state-owned enterprises, and two internal Independent Evaluation Group notes.

These documents were analyzed in two different ways: a structured content analysis of the Completion and Learning Review Validations was undertaken, while a semistructured review was undertaken of all other documents. To undertake the review of the Completion and Learning Review Validations, the text linked to Bank Group collaboration was extracted manually. The text was then reviewed to identify instances of collaboration and the type of collaboration. These data were captured in Microsoft Excel and aggregated. A quality check was undertaken on the data entry and coding. The semistructured review of all other documents was undertaken by two reviewers. Sections of documents that were defined as relating to Bank Group

collaboration were reviewed. For Country Program Evaluations, all sections of text that mentioned IFC or the Multilateral Investment Guarantee Agency were individually reviewed for relevance. Text extracts related to Bank Group collaboration were then captured. Each reviewer interpreted the text extracts for patterns individually, and then the reviewers met to discuss interpretations until they formed a consensus analysis of clusters of issues.

Our synthesis drew together findings from quantitative and qualitative analyses. We reread text clusters and developed descriptions to identify shared issues that enable and hinder World Bank collaboration. The findings from the qualitative and quantitative analyses were compared and findings developed. The findings in this appendix complement those presented in chapter 5 of this report.

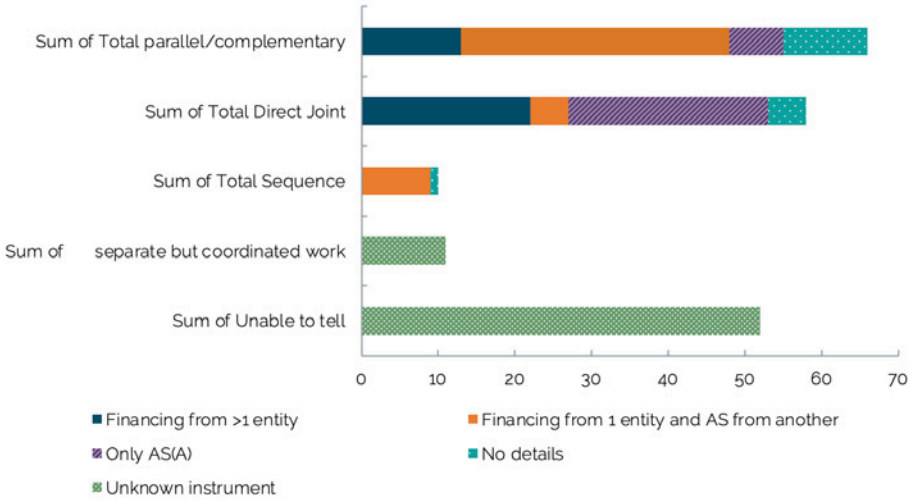
Findings

One World Bank collaboration can be undertaken in several different ways. Collaboration among Bank Group entities takes different forms: direct joint implementation, parallel and complementary work, sequenced work, and separate but coordinated efforts. Direct joint implementation entails the use of instruments within the same intervention. For example, IFC mobilized capital for renewable energy generation, and the Multilateral Investment Guarantee Agency provided political risk guarantees for those investments. Parallel and complementary work is when Bank Group institutions work in the same sector toward shared outcomes, such as by using different instruments to support selected agricultural value chains that show high export potential. Sequenced work is where Bank Group instruments are applied one after the other (for example, the World Bank supports upstream renewable energy reforms, and IFC undertakes investment projects). Separate but coordinated efforts entail general coordination in different sectors.

The most common form of collaboration is parallel and complementary, followed by direct joint implementation and then a different combination of instruments (figure J.1). Just over one-third of all collaboration was identified as parallel and complementary. Where the instruments can be identified, collaboration most often entails the use of advisory services and analytics (ASA) by one or both institutions. Direct joint implementation is

evident in 29 percent of instances of collaboration and has the most instances of financing from more than one institution. Sequenced and separate and coordinated efforts are identified in 5 percent of instances of collaboration. Where available, evidence shows that 30 percent of collaboration countries involved financing from more than one Bank Group entities; 42 percent involved financing from one Bank Group entity and advisory services (advisory services from IFC or ASA from the World Bank) from another; and 17 percent purely used ASA. No details on either the type of collaboration or the instruments used were reported in just over one-quarter of the instances.

Figure J.1. Types of Collaboration by Instrument



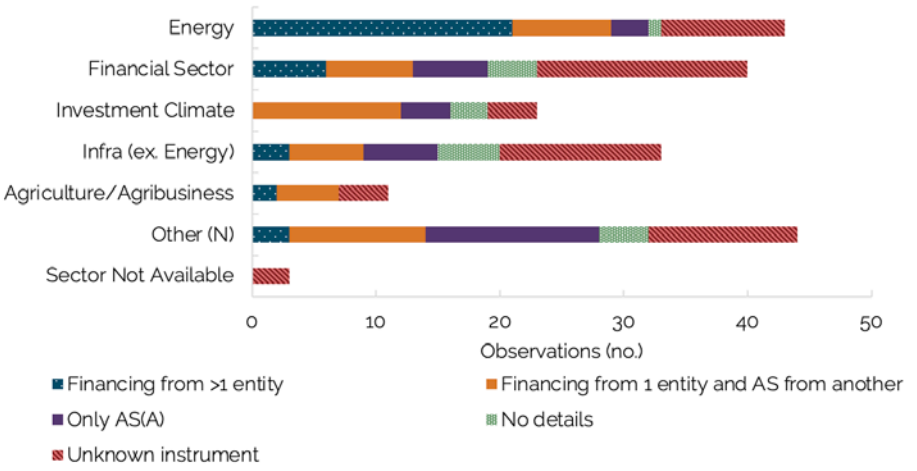
Source: Independent Evaluation Group.

Note: *N* = 197. AS = advisory services; ASA = advisory services and analytics.

Most collaboration occurs in a few sectors, with some notable smaller sectors (figure J.2). Collaboration in the energy, financial, infrastructure, investment, and agriculture sectors accounts for just over three-quarters of the instances of collaboration. In the energy sector, when details on instruments were available, 9 out of 10 of the instances of collaboration involved financing from one or more Bank Group entities. This is the largest proportion for any sector. In the financial sector, there were nearly equal contributions between collaboration financing from more than one Bank Group entity, financing from one Bank Group entity and ASA from another, and purely ASA financing. In the

investment climate, two-thirds of collaboration involved financing from one entity and AS(A) from another, while one-third was only AS(A).

Figure J.2. Collaboration by Sector and Type of Investment



Source: Independent Evaluation Group.

Note: N = 197. Each bar represents the number of observations across countries of collaboration in a sector. AS = advisory services; ASA = advisory services and analytics; Infra = infrastructure.

Based on these observations, the following considerations arose for the country engagement model. In addition to the observations shared in chapter 5, a core issue is to identify and prioritize major constraints to development and enable country teams to remain engaged in addressing them:

- A sound understanding of the major development constraints in a document that is appropriately timed to inform the Country Partnership Framework and is not negotiated with the government (as per the current Systematic Country Diagnostic) appears to be important. This underscores the importance of the Systematic Country Diagnostic that builds on core diagnostics such as the Country Economic Memorandum,[1] Country Climate and Development Report, and Country Private Sector Diagnostic.
- A sound analytic basis for the sectors and topics on which the Bank Group will engage is needed. This analytic basis often serves as a foundation for Bank Group collaboration.
- In areas that are preconditions for the achievement of broader and higher-level development objectives, even when the government does not have the

appetite to borrow or reform, the Bank Group can remain engaged, including by remaining current on issues through analytic work.

- Close coordination between staff from different Bank Group entities has been a factor in successful collaboration. Professional relationships between World Bank and IFC staff facilitate knowledge exchange and readiness to work together. This has happened through several channels:
  - The structural integration of World Bank and IFC implemented through the pilot joint Global Practices strengthened collaboration and helped achieve development outcomes.
  - Colocation of World Bank and IFC offices has been a positive factor in collaboration.
  - To an extent, informal networks and personal connections have substituted for structural solutions (including after the joint Global Practices were dismantled).

[1] We understand there are plans to rename this the Growth and Jobs Diagnostic.

IEG
Independent Evaluation Group
WORLD BANK GROUP
World Bank • IFC • MIGA

The World Bank
1818 H Street NW
Washington, DC 20433