

The World Bank Group Outcome Orientation at the Country Level

An Independent Evaluation



IEG
INDEPENDENT
EVALUATION GROUP

WORLD BANK GROUP
World Bank • IFC • MIGA



© 2020 International Bank for Reconstruction and Development / The World Bank
1818 H Street NW
Washington, DC 20433
Telephone: 202-473-1000
Internet: www.worldbank.org

ATTRIBUTION

Please cite the report as: World Bank. 2020. *The World Bank Group Outcome Orientation at the Country Level*. Independent Evaluation Group. Washington, DC: World Bank.

COVER PHOTO

shutterstock / Kogorash

EDITING AND PRODUCTION

Amanda O'Brien

GRAPHIC DESIGN

Luisa Ulhoa

This work is a product of the staff of The World Bank with external contributions. The findings, interpretations, and conclusions expressed in this work do not necessarily reflect the views of The World Bank, its Board of Executive Directors, or the governments they represent. The World Bank does not guarantee the accuracy of the data included in this work. The boundaries, colors, denominations, and other information shown on any map in this work do not imply any judgment on the part of The World Bank concerning the legal status of any territory or the endorsement or acceptance of such boundaries.

RIGHTS AND PERMISSIONS

The material in this work is subject to copyright. Because The World Bank encourages dissemination of its knowledge, this work may be reproduced, in whole or in part, for noncommercial purposes as long as full attribution to this work is given.

Any queries on rights and licenses, including subsidiary rights, should be addressed to World Bank Publications, The World Bank Group, 1818 H Street NW, Washington, DC 20433, USA; fax: 202-522-2625; e-mail: pubrights@worldbank.org.

The World Bank Group Outcome Orientation at the Country Level

An Independent Evaluation

October 16, 2020

Contents

Abbreviations	v
Acknowledgments	vi
Glossary	viii
Overview	x
Management Response	xvii
Management Action Record	xxv
Report to the Board from the Committee on Development Effectiveness	xxvi
1. Introduction	1
Methods	4
2. Aiming for Outcomes	7
Objectives for Country Engagements	8
Pathways to Achieving Development Objectives	12
Aiming for Development Objectives across the Bank Group Institutions	14
3. Capturing Outcomes	17
Results Frameworks	18
Completion and Learning Reviews	29
Other Evaluative Exercises	33
4. Managing for Outcomes	35
Types of Adaptation	36
How Teams Practice Adaptive Management	41
5. Toward Stronger Outcome Orientation	49
Rethinking Accountability	50
Rethinking the Tool Kit	52
Rethinking Incentives	57
Recommendation	62
 Bibliography	64

Boxes

Box 2.1. Objective and Indicator Classification	9
Box 3.1. Illustrating the Mismatch: Nicaragua Early Childhood	23
Box 3.2. Evaluating Advisory Services in Romania	25
Box 3.3. The International Finance Corporation's Contributions to Country Outcomes in Guyana	27
Box 3.4. A Partial View of Bank Group Progress on Long-Term Reform Agendas: Examples of State-Owned Enterprise Reforms in Serbia	30
Box 3.5. Examples of Lessons from the Completion and Learning Review	32
Box 5.1. Trade-Offs in Adopting a New Country-Level Results System	60
Box 5.2. How Could Monitoring, Evaluation, and Learning Plans Work?	62

Figures

Figure 1.1. Conceptual Framework	3
Figure 2.1. Country Partnership Framework Objectives by Outcome Level	10
Figure 3.1. Indicator Outcome Level by Type of Objective	19
Figure 3.2. Discrepancy between Objective and Indicator Level of Outcomes	21
Figure 5.1. Alternative Causal Reasoning at the Country Level	55

Tables

Table 2.1. Morocco's Adaptive Country Partnership Framework in Contrast with More Conventional Approaches	11
Table 4.1. Portfolio Adaptation and Evidence Needs	37
Table 5.1. How a Monitoring, Evaluation, and Learning Plan Could Address Current Limitations	54

Appendixes

Appendix A. Methods	79
Appendix B. Structured Document Review	97
Appendix C. Summary of Findings by Country Engagement Type	150
Appendix D. Options for a Prototype Monitoring, Evaluation, and Learning Plan	160

Abbreviations

ASA	advisory services and analytics
CLR	Completion and Learning Review
CLRR	Completion and Learning Report Review
CMU	Country Management Unit
CPF	Country Partnership Framework
CPPR	Country Portfolio Performance Review
CS	Country Strategy
EU	European Union
FCV	fragility, conflict, and violence
GP	Global Practice
IBRD	International Bank for Reconstruction and Development
IDA	International Development Association
IEG	Independent Evaluation Group
IFC	International Finance Corporation
MEL	monitoring, evaluation, and learning
MIGA	Multilateral Investment Guarantee Agency
PLR	Performance and Learning Review
RAP	Results and Performance of the World Bank Group
RAS	reimbursable advisory services
ROC	Regional Operations Committee
SCD	Systematic Country Diagnostic

All dollar amounts are US dollars unless otherwise indicated.

Acknowledgments

This evaluation was prepared by an Independent Evaluation Group (IEG) team led by Estelle Raimondo, senior evaluation officer, under the overall direction of Alison Evans, Director-general, Evaluation, and with the guidance and supervision of Galina Sotirova, manager, Corporate and Human Development, Oscar Calvo-Gonzalez, director, Human Development and Economic Management, and Sophie Sirtaine, director, Strategy and Operations.

The team comprised Stephen Hutton, senior evaluation officer, who coauthored the report, led the design of case studies, and conducted several; Ana Belen Barbeito, senior evaluation officer, who led the International Finance Corporation (IFC) and Multilateral Investment Guarantee Agency analysis and conducted case studies; April Connelly, senior natural resource management specialist, who conducted case studies; Eduardo Fernandez Maldonado, evaluation officer, who managed the structured document review and conducted a number of case studies; Gaby Loibl, program assistant, who coordinated missions, meetings, interviews, and report preparation; Juan Carlos Mendez De La Torre, consultant, who led the literature review and the adaptive management content analysis; and Xiaoxiao Peng, consultant, who led the structured document analysis, quantitative analysis, and sentiment analysis. Maximillian Ashwill was the report's primary editor.

External reviewers for this evaluation were Ms. Ellen Goldstein, president of EAG Global Leadership and former Country Director at the World Bank; Dr. Ruth Levine, an expert in global health, education, and evaluation who is currently chief executive officer of IDinsight and was previously program director, Global Development and Population at the Hewlett Foundation, a senior fellow and vice president for programs and operations at the Center for Global Development, and deputy assistant administrator in the Bureau of Policy, Planning and Learning at the United States Agency for International Development; Professor Frans Leeuw, emeritus professor at Maastricht University and an international expert in monitoring and evaluation systems; and Dr. Ben Ramalingam, international expert on adaptive management in international development, executive director of the Global Learning Alli-

ance on Adaptive Management, and senior fellow at the Overseas Development Institute.

The team is grateful for the very useful feedback provided throughout the evaluation process by the Operations Policy and Country Services teams, the IFC Corporate Strategy and Partnership Unit, and Country Economics Team, as well as regional vice presidential units. Finally, the team is grateful to the many World Bank Group country teams who engaged with us and gave generously of their time for interviews. Thanks are also due to the country offices of Peru, Romania, Serbia, and the Solomon Islands, as well as Hubs in Vienna and Sydney, for their support during our missions.

Glossary

Adaptive management An iterative approach to decision-making where approaches are adjusted based on evidence and evolving context.

Attribution The expectation that results considered for accountability purposes are directly caused by World Bank Group interventions.

Country engagement cycle The Bank Group's country engagement cycle is made up of four country reports: the Systematic Country Diagnostic, the Country Partnership Framework, the Performance and Learning Review, and the Completion and Learning Review.

Country-level results system The World Bank Group's results system for the country engagement cycle, above the project-level results system that covers individual engagements and below corporate-level results systems that cover whole institution(s).

Country outcomes Development outcomes that take place in client countries, and are at a higher level than a single intervention.

Country Partnership Framework (CPF) The main strategy document for the Bank Group, which lays out the main country development goals that the Bank Group aims to help the country achieve and proposes a selective program of indicative Bank Group interventions for this purpose. It also contains a results framework, which includes specific objectives and indicators against which the program is monitored and evaluated.

Country team The core staff working on strategy and operational management for a World Bank Group country engagement. It includes the Country Management Unit (in particular country director, country manager or operations manager, country officers, country program coordinator, and resident representative, where applicable) but also key sector staff who drive strategic direction above the level of an individual engagement, such as program leads or sometimes senior/lead technical staff who play similar roles. It includes International Finance Corporation (IFC) staff, particularly the IFC country manager and country officers.

Completion and Learning Review (CLR) A self-assessment carried out at the end of the country engagement cycle that assesses the lessons emerging from the cycle and whether the intended objectives were achieved. The CLR is then validated by the Independent Evaluation Group in a Completion and Learning Report Review (CLRR).

IFC country products A new set of internal documents used by IFC to support its contribution to the country engagement cycle. They include the joint IFC–World Bank Country Private Sector Diagnostic to inform and complement the Systematic Country Diagnostic (SCD), the IFC Country Strategy to lay out strategic priorities and reinforce IFC’s position in the CPF, and a country-level business plan and its subsequent reviews.

IFC Country Strategies A new internal document used by IFC to improve IFC’s inputs to the CPF and better guide its investment decisions within countries, and to help target World Bank programs aiming to support private sector development. The strategy lays out IFC’s strategic vision in a country by elaborating IFC’s delivery capacity and strategic priorities under different policy reform scenarios.

Monitoring, evaluation, and learning (MEL) plan An alternative country-level results system proposed by this evaluation.

Outcome orientation The organization’s ability to generate feedback on what works, what does not, and why, use this feedback to engage clients and adapt country programs, and ultimately bolster its contribution to country development outcomes.

Performance and Learning Review (PLR) The midterm review instrument for country strategies, through which the Bank Group takes stock of progress and decides on needed changes and course correction.

Results framework CPF results frameworks, also referred to as results matrices, assign a set of indicators and targets to CPF objectives. Results frameworks are introduced in the CPF at the beginning of the country engagement cycle, can be revised midway through the cycle in the PLR, and are used to rate the Bank Group’s results and performance at the end of the cycle in the CLR.

Overview

Highlights

||| This evaluation assesses whether the World Bank Group's country-level results system adequately supports the organization's monitoring, evaluation, and learning needs to effectively manage country engagements. The evaluation does not assess the project-level results system.

||| The model of how the Bank Group aims for outcomes in its client countries is sound, and, for the most part, teams practice this well.

||| However, the country-level results system does not capture the Bank Group's contribution to country outcomes well, because it relies on results frameworks premised on metrics, attribution, and time-boundness that do not fit the nature of country programs well.

||| Country teams practice adaptive management, but the country-level results system does not effectively support them in doing so. Instead they rely on tacit knowledge, professional experience, and professional networks when making adaptive decisions. This has some disadvantages.

||| This evaluation provides a new vision of how the country-level results system's accountability principles, tools, and incentives could function.

The World Bank Group's success rests on its ability to help its clients overcome challenges to achieve the development outcomes they desire. This demands a strong outcome orientation at the country level, defined as the organization's ability to generate feedback on what works, what does not, and why, use this feedback to engage clients and adapt country programs, and ultimately bolster its contribution to country development outcomes. The Bank Group uses the country engagement cycle and its underlying results system to support the organization's monitoring, evaluation, and learning (MEL) needs to effectively manage country engagement. As part of this cycle, country teams are expected to aim to make meaningful contributions to country development outcomes, to capture those contributions in a results system, and to use information from the results system to inform adaptive management.

This learning-focused evaluation aims to help country teams and management strengthen the Bank Group's country-level outcome orientation. It does this by focusing on how country teams carry out four key practices of outcome orientation. These practices include the following: (i) aim for outcomes in the Country Partnership Framework (CPF) and CPF results framework; (ii) monitor and evaluate the Bank Group's contribution to country outcomes and learn from this evaluative evidence; (iii) adapt country engagements; and (iv) engage clients to measure and manage outcomes. The evaluation does not assess the project-level results system. The evaluation is based on a sample of 39 country engagements selected to represent the diversity of Bank Group country clients and portfolios, and draws on a combination of content analysis of country documents, interviews with over 185 Bank Group staff and clients, and a review of practices and literature on evaluation systems at the Bank Group and other organizations.

Aiming for Outcomes

The evaluation finds that the model of how the Bank Group aims for outcomes at the country level is sound and, for the most part, country teams practice this model well. The Bank Group's country strategies set out to influence country outcomes and frame their objectives in terms of outcomes, while facing difficulties in being selective because of internal competition for program space and heterogeneous client demand. The Bank Group

pursues its objectives through both direct pathways (such as the outputs and services provided by projects) and indirect pathways (including institutional development, capacity building, knowledge transfer, demonstration effects, and market creation). About one-third of CPF objectives coded for this evaluation explicitly aimed at building institutional capacity within a sector. However, these indirect pathways are not well articulated in country documents, and the relatively short CPF cycles do not match the long-term vision needed to sustain indirect pathways. There has been progress in integrating International Finance Corporation (IFC) and Multilateral Investment Guarantee Agency teams into the CPF design, but CPFs do not serve IFC's strategic needs well. This motivated IFC to launch its own internal country strategies and other country products. Early experience with these IFC country strategies shows that they offer potential for increasing IFC's focus on country outcomes, but close alignment with the World Bank and Bank Group country engagement is a condition of success and to avoid risks of duplicated efforts.

Capturing Outcomes

The Bank Group falls short of capturing its contribution to country outcomes because of the limitations of the country-level results system. The results system relies on the use of CPF results frameworks as its primary tool for assessment predicated on three principles: metrics, attribution, and time-boundedness. These three principles present critical challenges when applied to country programs.

Metrics. CPF results frameworks are built predominantly on quantitative indicators, but these indicators do not capture the CPF objectives being pursued. Newly collected data show that about half of the time these indicators are framed below the level of the objectives they seek to track. Some outcomes, such as improved governance, are more intangible and inherently more difficult to measure. Even for outcomes where meaningful metrics exist, data gaps in countries pose serious challenges.

Attribution. Teams are reluctant to—and are advised not to—include indicators that are not directly attributable to Bank Group activities, making it difficult to capture the Bank Group contribution to country outcomes and outcomes pursued through indirect pathways.

Time bound. The country-level results system is time bound and does not capture the medium- to long-term effects of Bank Group interventions. It also largely captures results from operations launched in the previous strategy period, rather than the new forward-looking strategy.

Most indicators come from projects because teams can have confidence that the data will exist and will be attributable to the Bank Group's activities, and they can collect the data without additional evidence-gathering expenses.

There are direct and opportunity costs associated with compiling, processing, controlling quality, and reporting on indicators that are not used.

The Bank Group's country evaluation product, the Completion and Learning Review (CLR), provides a partial picture of Bank Group contributions. The picture is partial because of its overemphasis on those results that can be measured and on results from lending projects. The CLR rarely captures complementarities across instruments or institutions and so is not able to establish whether the Bank Group's contribution to country outcomes amounts to more than the sum of its parts. Because of their timing and limited content, the reviews are little used within the institution. Staff almost invariably argue that the accountability elements of the CLR lack any actual bite, in part because they are rarely discussed or used by senior management or the Board to create incentives and make decisions.

Although some country teams carry out other evaluative exercises, they are rare and overshadowed by the results system's emphasis on indicators and targets. There are many examples of activities that seek to assess the outcomes of interventions embedded in sector work, but these are not well captured in country reports. The Bank Group's monitoring and evaluation staff are consumed by corporate reporting requirements with little time to support country teams' outcome measurement.

Managing for Outcomes

The country-level results system does not provide adequate support to country teams engaging in adaptive management. Country teams practice adaptive management by closely monitoring the health of the country portfolios

focused on disbursements, addressing delivery issues, and navigating changes in country contexts. However, teams rarely use the country-level results system to make adaptive management decisions because they do not find the information to be useful for showing whether the country is progressing or not, and because the information is not up to date and has to be gathered manually. Instead, country teams rely on tacit knowledge, professional experience, and advice from professional networks when making adaptive decisions. This reliance has some disadvantages.

Country strategies' midterm review instrument, the Performance and Learning Review (PLR), is used largely to document decisions that were already made and adjust the results frameworks. More than 80 percent of CPF indicators reviewed in this evaluation were revised or dropped at the PLR stage. However, the extent to which teams use the PLR for collective reflection and adaptive management depends on whether country team leaders make it a priority and choose to go beyond the requirements of the system.

Clients do not engage much in the Bank Group's country-level results system, and the World Bank seldom helps countries develop their own monitoring, evaluation, and results management systems. Moreover, there is little coordination of results management activities among development partners.

Toward a Stronger Outcome Orientation

This evaluation provides a new vision of how the country-level results system's accountability principles, tools, and incentives could function. It will be important not to repeat the mistakes of the past, which have doubled down on attribution and added more performance measures and indicators that have not been useful. In renewing the system, trade-offs will inevitably emerge between external reporting versus internal use, standardization versus customization, centralized versus decentralized models, and comprehensive versus selective approaches.

The current country-level results system relies on specific principles of accountability that do not fit well with the way the Bank Group pursues outcomes at the country level. The system is premised on a narrow view of accountability that equates it with counting results that can be translated

into metrics and ratings, yet staff do not find that the system substantially drives their incentives.

A renewed country-level results system could conceive accountability differently, based on evidence of achievement and failures and descriptions of learning and adaptation. It could acknowledge that the Bank Group can influence but not control country outcomes, recognizing that country teams cannot decide all targets and objectives at design but must adapt during implementation, and realizing that capturing contributions to country outcomes and assessing cumulative effects from multiple interventions requires dedicated evaluation inquiries, not just measurement of indicators.

The country-level results system is too reliant on results frameworks and could benefit from introducing new approaches such as MEL plans. Under these plans, approaches would change in the following ways:

- » Monitoring approaches would focus on selectively tracking critical country outcomes to assess progress on objectives and to support adaptive management.
- » Evaluation approaches would focus on the Bank Group's contribution to country outcomes, scale back ratings, and emphasize selectivity over coverage.
- » Learning approaches would focus on creating safe spaces for teams to engage and discuss evaluative evidence.
- » Changes in the results system would need to be supplemented by changes in signals and incentives to center on results achievement and management rather than overemphasizing approvals, commitments, and output delivery. Incentives could better match staff's intrinsic motivations.

This evaluation also makes the following concrete recommendation:

Recommendation

The Bank Group should reform the country-level results system to ensure that it accurately captures the Bank Group's contribution to country outcomes and usefully informs decision-making on country engagements. This would entail keeping certain practices, discontinuing others, and introducing new ones.

What to keep. The country engagement should continue to articulate clear outcome-level objectives and lay out the various pathways that will be pursued to achieve them. Periodic reviews to take stock of progress, achievements, and failures should be maintained, as well as an end-of-cycle review of evaluative evidence and learning informed by this evidence.

What to discontinue. The current requirements of developing a CPF results framework with a complete set of measurable indicators and targets at the design stage, making adjustments at midterm, and rating country programs on how far achieved metrics departed from initial targets should be discontinued.

What to introduce. Instead of the country results framework, adopt a new approach and enable country teams to develop a MEL plan at the CPF design stage that can be revised throughout the country cycle. A monitoring plan would describe how the country team will monitor performance, changes in key country outcomes that the Bank Group seeks to influence, and relevant contextual risks and conditions. An evaluation plan would define the criteria for selecting and using evaluations and would synthesize existing evaluative evidence. A learning plan would identify knowledge gaps to be filled and plan activities for reflection.

Management Response

Management of the World Bank Group institutions would like to thank the Independent Evaluation Group (IEG) for its report, *The World Bank Group Outcome Orientation at the Country Level*. Management welcomes IEG's positive findings that the Bank Group's country-level objectives are indeed outcome oriented and articulate well the Bank Group's contributions to achieving targeted outcomes. Management also welcomes opportunities to explore how it may improve the links between Bank Group interventions and higher-level outcomes and development goals.

World Bank Management Response

Management is pleased with the report's conclusion that the model of how the Bank Group aims for outcomes in its client countries is sound and that teams, for the most part, practice it well. Management is encouraged by the assessment that in the Country Partnership Frameworks (CPFs) reviewed by this evaluation, “nearly all objectives were framed in terms of outcomes,” and “country teams understand well the country outcomes they are working toward” (8). Management also notes that “country teams articulate well how the combination of Bank Group operations aim to influence certain outcomes” (14). These practices have proactively been encouraged by staff training, learning activities and guidance on results and the country engagement model.

Management shares the report's view that continuous effort is needed to better articulate indirect pathways to country outcomes in country strategy documents and is taking steps in that direction. These pathways are the Bank Group's most distinct contribution to development results in most client segments. Management is aware that in an ever-changing development landscape, the Bank Group must clearly articulate how it combines financing, technical, and country expertise to tackle protracted challenges and respond to international commitments. In a paper presented to the Committee on Development Effectiveness in August 2020, management presented four development pathways to enhance the World Bank's outcome orientation,

one of which comprises concrete steps to enhancing the connection of World Bank-supported results to higher-level goals. Specifically, the World Bank is expanding the use of a theory of change (and links to the results framework) to improve its focus on achieving outcomes at the country level while also connecting those outcomes to higher-level goals articulated in the 2030 Agenda for Sustainable Development. Germane findings of this evaluation will be used to reinforce the proposed approach.

Management acknowledges the report's view that, as the center of the country-level results architecture, the results framework has shown limitations, and management is willing to explore ways to strengthen the framework's usefulness and outcome focus. While sharing the report's view of the results frameworks' limitations in terms of quantitative metrics, focus on attribution, and arbitrary time-boundedness anchored to World Bank processes, management believes that they have still played an instrumental role in instilling a more robust culture of results in both country teams and client counterparts. As a signatory of the Paris Declaration on Aid Effectiveness, the Accra Agenda for Action, and the Busan Partnership for Effective Development Cooperation, the World Bank has endeavored to develop results frameworks that are aligned with countries' own results frameworks to promote mutual accountability and enhanced management for results. Nevertheless, recognizing the need to avoid overreliance on quantitative metrics to assess progress toward development outcomes, the World Bank is piloting an effort to assess how multiple Bank Group activities, in successive CPFs, contribute to higher-level outcomes and how such learning can be integrated into the CPF process to inform future country programming.

Management is of the view that any new approach, such as the one proposed by the monitoring, evaluation and learning plan, should complement rather than replace the current results system. Management appreciates that in the spirit of the Management Action Record reform, IEG has offered a general prototype (in appendix D) of what the new proposal "might look like," presenting a vision that "is not intended to be prescriptive" (160). In the same spirit, management is thus willing to discuss with IEG the details and practicalities of such vision to ensure a good "fit" with Bank Group operations. Once a common vision is developed, management proposes to pilot such approaches in a select number of client countries that represent different client

segments. These pilots could, for example, be rolled into the above-mentioned pilots currently underway to better document how interventions over multiple CPF cycles collectively contribute to country-level outcomes against specific themes. Management notes that any potential changes made to the current results system would require IEG to make appropriate changes to the way it reviews country-level performance.

International Finance Corporation Management Comments

International Finance Corporation (IFC) management welcomes IEG's report, *The World Bank Group Outcome Orientation at the Country Level*. IFC supports the finding that the current country-level results system does not fully capture the Bank Group's contribution to country outcomes and that reliance on results frameworks premised on metrics, attribution, and time-boundedness can limit a broader assessment of the complex package of support that the Bank Group delivers. IFC agrees that the current results system does not completely answer questions about how well the Bank Group's interventions build institutions, create markets, or shape policies and under what conditions these institutions contribute to sustained positive changes to well-being.

Although the current system has weaknesses, IFC management believes that results frameworks have played an instrumental role in instilling a culture of results in country teams and client counterparts and that this should not be lost in any new approach. In this spirit, IFC management supports partnering with the World Bank and Multilateral Investment Guarantee Agency (MIGA) to pilot a new approach, in select countries, that would better assess how interventions over multiple CPF cycles collectively contribute to country-level outcomes against specific themes. IFC management appreciates the additional ideas provided in appendix D of the report, which was added too late for management to fully digest. However, the ideas presented in the new appendix, while useful, are far from what is required to determine a new approach and developing an effective system will require a significant effort. In doing this, it is important to work closely with IEG, whose approach to CPF evaluations drives much of the current process. IFC would like to reiter-

ate the importance of streamlining any new approach so as not to add to the burden of already stretched operational teams.

IFC management welcomes the recognition in the evaluation of IFC's efforts to deepen IFC's inputs into the CPF and better guide investment decisions within countries through the introduction of several new tools that enrich the Bank Group's country engagement cycle—which IEG recognizes represents a significant departure from business as usual. As highlighted in chapter 2, these include the joint IFC and World Bank Country Private Sector Diagnostics to inform and complement the Systematic Country Diagnostic, the IFC Country Strategies to lay out strategic priorities and reinforce IFC's position in the CPF, and a country-level business plan and its subsequent reviews. The Country Strategy articulates clear investment and development impact scenarios captured by “if-then” matrixes, which lay out critical policy barriers where reform (and potentially World Bank engagement) could unlock investment opportunities.

IFC management is pleased to note the recognition provided in the evaluation that IFC Country Strategies support increasing focus on country outcomes. It is IFC's ongoing experience that the process of combining an assessment of country development gaps and IFC's comparative advantage with market data and analytics and a detailed understanding of regulatory and capacity barriers enables IFC to develop a common road map of priority sectors for engagement, improving the internal coherence of IFC interventions and depth of impact by influencing which sectors IFC is active in. IFC has found that it has been an especially useful tool in shaping business development efforts with strong development impact, and discouraging project proposals that are not well aligned with the Bank Group's strategic priorities.

IFC management particularly appreciates the finding that aspects of IFC's actions, are particularly advisory services, not yet fully reflected in the current CPF process. IFC is expecting to increase its efforts toward creating projects that help solve a development problem through IFC upstream activities, and inherent in the design of this approach is the expectation that some things tried might need to be set aside. In that context, IEG's proposal that more relevant and eclectic evidence on assessment of achievements and failure is of interest and worth further consideration.

Multilateral Investment Guarantee Agency

Management Comments

MIGA welcomes IEG's report, *The World Bank Group Outcome Orientation at the Country Level*, and finds it useful and important. The report presents many significant findings, and MIGA appreciates IEG's observations. MIGA thanks IEG for the productive engagement during the drafting of the report.

MIGA's integration into Bank Group country engagement. MIGA welcomes the report's finding that there has been significant progress in integrating MIGA teams into the country engagement process, including the preparation of Systematic Country Diagnostic and CPFs as well as Performance Learning Review and Completion and Learning Reviews. MIGA's Risk Management Officers leverage the in-depth country knowledge acquired from the country engagement process for determining country risk ratings, a key input for MIGA operational work. MIGA values highly the knowledge sharing and network building across the Bank Group that emanates from the country engagement process and has recently added additional staff to contribute to the effort. MIGA agrees with the report's finding that the Agency prioritizes its involvement in countries that are of strategic importance, primarily International Development Association and fragility, conflict, and violence countries.

MIGA notes that the World Bank, IFC and MIGA are working closely together in several areas where synergies and complementarities are being actively pursued, leveraging each other's strengths. In operations, the institutions are complementing and integrating their work in creating markets and fostering cross-border investment, and in overcoming obstacles that block private sector solutions. As part of this effort, the three institutions are better coordinating analytical and advisory deliverables, especially for upstream work that can greatly enhance MIGA's ability to play a greater role downstream in support of private sector clients.

MIGA is also expanding its efforts to participate in IFC's Country Private Sector Diagnostic, which informs and complements the Systematic Country Diagnostic, as well as IFC Country Strategies, which will also influence MIGA's strategic choices at the country level given the often close collaboration between IFC and MIGA at the sectoral level within individual countries.

MIGA's deepening partnership with the World Bank and IFC are complementary and will promote stronger joint program delivery with the ultimate goal of delivering greater development impact than would otherwise have been the case.

MIGA also uses sector frameworks such as the Financial Sector Assessment Programs and Infrastructure Sector Assessment Programs to collaborate with the Bank Group (and the International Monetary Fund when relevant) and identify program and project opportunities. Overall, MIGA is actively engaged in all of the countries for which a Bank Group country engagement process is undertaken. MIGA works to both learn and actively contribute across the Bank Group country engagement process, as MIGA's mandate, like that of the Bank Group, is to serve all clients—a mandate MIGA takes seriously.

Client demand and country program design. MIGA agrees with the report's finding that client demand, including that of MIGA clients, has been a key factor in country program design. The report found that in certain contexts, the heterogeneity of client demand significantly constrains country teams' ability for selective and sustained contribution to country outcomes. MIGA provides guarantees to private sector clients against noncommercial risks. Therefore, MIGA is typically not in a position to know in detail the specific clients or projects that it will be supporting over the medium term. Nonetheless, the role that MIGA can play in supporting development outcomes, as articulated in the various country engagement deliverables, has been clarified in recent years by adoption of the Cascade approach across the Bank Group. The Cascade approach inscribes a key and pivotal role for MIGA in signaling that the Bank Group should first seek to mobilize private sector solutions, applying official and public resources only where market solutions are not possible, including through the application of sector reforms and risk mitigations instruments, including MIGA guarantees.

MIGA contributions to country outcomes. The report finds that MIGA's contributions to outcomes are not well captured in CPF results frameworks, as evidenced by the disconnect between stated MIGA contributions to CPF objectives and actual results indicators that capture MIGA contributions. The report recognizes that the disconnect is explained in part by the fact that MIGA's project pipeline is largely unknown at the CPF design stage,

which makes it difficult for task teams to establish indicators and targets that incorporate MIGA projects. MIGA agrees with the findings and notes the inherent challenges for the CPF results framework in capturing the distinct and joint contributions of the three Bank Group institutions—World Bank IFC, and MIGA—to country outcomes, given the unique mandates and purposes of the Bank Group institutions and the attendant business models.

Development impacts and financial returns. The report notes that IMPACT (Impact Measurement and Project Assessment Comparison Tool), MIGA's ex ante development impact assessment framework, assesses the ex ante development effects of MIGA guarantee projects, including their ability to attract and mobilize foreign direct investments. The Report also notes that IMPACT has been designed to further encourage staff to prioritize potential development impacts as well as financial returns. In addition, MIGA notes that IMPACT was developed to provide a standardized methodology for assessing and communicating the potential development impact of proposed MIGA guarantee projects. The objectives include the following: (i) inform project prioritization; (ii) compare expected development impacts across the portfolio; (iii) standardize communications to relevant stakeholders; and (iv) serve as an input for MIGA's project assessment along with other criteria. Regarding the aspect of designing IMPACT to further encourage staff to prioritize development impact as well as financial returns, the approach is consistent with MIGA's mandate to serve all client countries and remains vital for its own financial sustainability as a development institution. MIGA will continue to support projects across the country-income spectrum, while continuing to focus on the poorest and most fragile countries, consistent with the vision outlined in its FY21–23 Strategy.

Country-level results systems and Bank Group contribution to country outcomes. The report finds that the Bank Group approach for country outcomes is sound, with country teams practicing it well for the most part. However, the country-level results system does not capture the Bank Group's contribution to country outcomes as well as it might because it relies on results frameworks premised on metrics, attribution, and time-boundedness that do not fit well the nature of country programs. While agreeing broadly with these conclusions, MIGA underscores the fact that country-level results systems are primarily focused on accountability. Results measurement sys-

tems and approaches have progressed steadily over time, across the three Bank Group institutions. Even so, the Bank Group has recognized the inherent limitations of existing results measurement systems from the standpoint of measuring higher-order results—outcomes and impacts—that cannot be achieved through a single activity. Further, country-level results depend on contributions from multiple interventions and a broad spectrum of partners and sources of finance or advice. With respect to the specific IEG recommendation of developing an alternative country-level results system through the monitoring, evaluation, and learning plan, MIGA is of the view that the plan could be piloted for understanding better MIGA’s contributions to country outcomes.

Management Action Record

IEG Findings and Conclusions. The Bank Group should keep a country engagement model that articulates clear outcome-level objectives and lays out the pathways that will be pursued to achieve them, conducts periodic reviews to take stock of progress, and includes an end-of-cycle review of evidence and learning. These reviews need not be a Board deliverable but should be carried out in time to meaningfully inform course correction and the next strategy. It should discontinue the reliance on results frameworks in their country strategies and midterm and terminal reviews. And it should adopt monitoring, evaluation, and learning plans for its country engagements.

IEG Recommendations. Reform the country-level results system to ensure that it accurately captures the Bank Group's contribution to country outcomes and usefully informs decision-making on country engagements.

Acceptance by Management. Yes.

Management Response. Management appreciates that in the spirit of the Management Action Record reform, IEG has offered a general prototype of the new monitoring, evaluation, and learning proposal, the details of which would be determined jointly with management to ensure a good “fit” with Bank Group operations.

In the same spirit, management is thus willing to pilot such approaches in a selected number of client countries, representing different client segments. These pilots could be rolled into ongoing efforts to assess how multiple Bank Group activities, in successive Country Partnership Frameworks, contribute to higher-level outcomes, and how such learning can be integrated into the Country Partnership Frameworks process to inform future country programming.

Report to the Board from the Committee on Development Effectiveness

The Committee on Development Effectiveness met to consider the report entitled *The World Bank Group Outcome Orientation at the Country Level* and draft management response.

The committee welcomed the report's findings and recommendations for improving the outcome orientation of country programs and commended management and the Independent Evaluation Group (IEG) for their collaborative and constructive engagement and for having reached agreement on the best way to improve the current system for measuring country outcomes. Members were pleased to learn that the Bank Group's model for gauging outcomes in its client countries is sound; that, for the most part, teams practice it well; and that it has played an important role in establishing a culture of results measurements in the Bank Group's operational teams and the clients it supports. Nonetheless, they agreed with IEG and management that there was room for improvement, noting that the Bank Group country-level results system did not effectively capture the Country Partnership Framework's objectives or the Bank Group's contribution to changes in country outcomes, and it did not enable teams to respond to changing circumstances or drive incentives or behaviors to uphold accountability for results at the level of the country program. The committee was encouraged to learn that management agreed with the need to better evaluate the pathways to country outcomes in Country Partnership Frameworks, including better integrating the contributions of the International Finance Corporation (IFC)'s and the Multilateral Investment Guarantee Agency, exploring ways to capture how the Bank Group contributes to higher development outcomes through institutional change and capacity building and moving away from the false dichotomy of attribution versus contribution.

In calling for these reforms, IEG recognized the value of a results measurement system that accurately captures country outcomes while allowing for adaptive management and agreed that the country results system needed to

be adjusted to find a better balance between contribution and attribution, strengthening and capturing links with higher-level objectives such as the corporate goals and Sustainable Development Goals, and strengthening the integration of contributions by IFC and the Multilateral Investment Guarantee Agency. Members noted that the new model should build on the work IFC is doing as part of IFC 3.0 and that IEG will need to make appropriate changes to the way it validates country-level performance. The committee noted testing and adjusting the monitoring, evaluation and learning plans would be an iterative process that would take time and require wide consultation and thus, the committee appreciated management's commitment to define and test IEG's suggested approach in a selected number of pilot countries and to keep the Board of Executive Directors apprised of these pilots and of the possibilities of mainstreaming the approach. IEG emphasized that for a renewed system to be useful, reforms should not layer more demands on top of the existing system; some practices have to be stopped so that new ones can take root. Management noted that, for the new approach to be useful, it would need to be practical and grounded in field and operational realities and client capabilities.

Members underscored the value of strong collaboration across the different actors, including important roles for client-driven ownership of country programs and understanding of the Bank Group's contributions to country outcomes jointly with other development partners. Members encouraged management to consider incentive systems that would enable staff to focus on achieving higher development objectives, particularly through funding and leadership support. Management noted that reforming the system to allow for better reflection of achievement of higher development outcomes may require further flexibility in adjusting design and course correction, leniency in data collection, adjustments to the way IEG evaluates results, and the Board's understanding of measurement challenges.

Members noted the need to link the exercise to the outcome classification proposed in IEG's *Results and Performance of the World Bank Group 2020* and to the upcoming knowledge paper. They encouraged management to address the difference in time frames between country cycle documents and higher-level outcomes; have a client and partnership focus; and take advisory services, policy dialogue, and the Bank Group's convening power in the Country Partnership Frameworks into account.

1 | Introduction

The World Bank Group's success rests on its ability to help its clients overcome challenges to achieve the development outcomes they desire. The unprecedented health and economic crisis prompted by the coronavirus pandemic (COVID-19) has put into sharper focus the importance of well-functioning and flexible country engagements to the Bank Group's success. In this context, the Bank Group's range of financing and knowledge instruments are called on to propose solutions to the crisis's multidimensional nature and to help country clients' transition to recovery. These solutions must have a long-term vision that taps into the Bank Group's global knowledge while being adapted to the specific context of each country. Responding effectively to these types of crises and other development challenges demands a strong country-level outcome orientation, defined as the organization's ability to generate feedback on what works, what does not, and why; use this feedback to engage clients and adapt country programs; and ultimately bolster the Bank Group's contribution to country development outcomes. A country-level results system that adequately supports adaptive decision-making is a key element for achieving this outcome orientation.

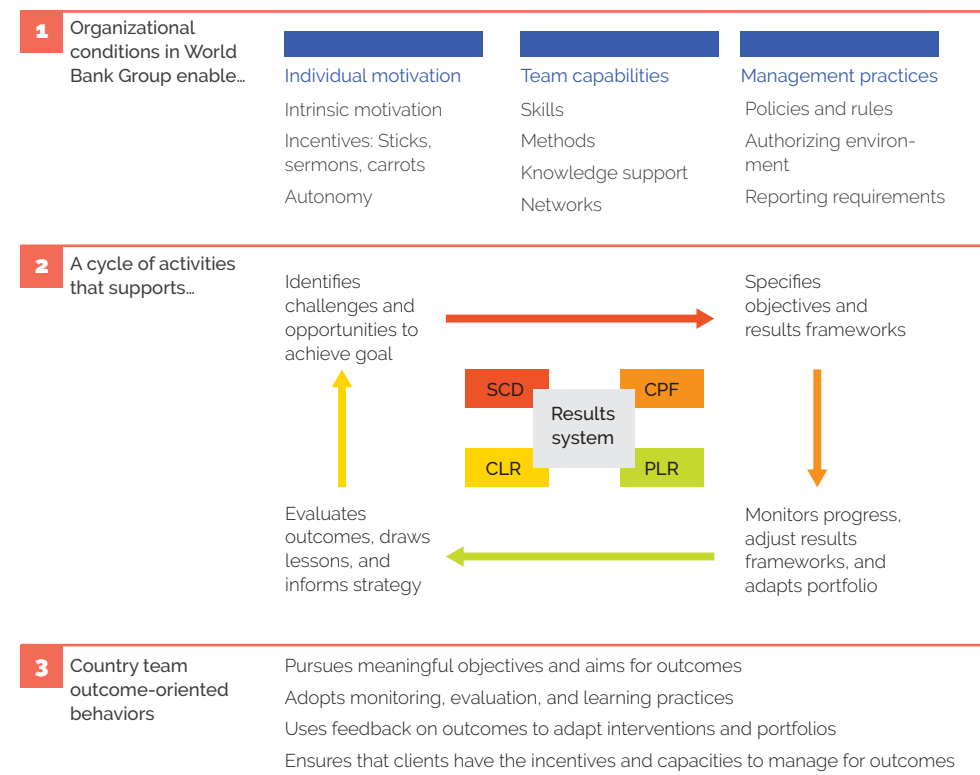
The Bank Group uses the country engagement cycle to organize how it manages and measures its contribution to country outcomes. The country engagement cycle begins when the Bank Group conducts a Systematic Country Diagnostic (SCD), which it uses to define priorities and objectives jointly with its clients in Country Partnership Frameworks (CPFs). The World Bank's country team adapts the CPF throughout the country engagement cycle and documents these adaptations in the Performance and Learning Review (PLR). The country team assesses the results of the country engagement in Completion and Learning Reviews (CLRs), which is intended to help determine the Bank Group's future position in the country. This process is shown in the central part of figure 1.1. It is through the country engagement cycle that country teams organize, manage, and adapt their interventions and client dialogues for best results. At the country level, the Bank Group's

value lies in its capacity to deploy, combine, or sequence a wide range of instruments from each of its component institutions to help client countries eradicate poverty, increase shared prosperity, and achieve the Sustainable Development Goals. The Bank Group's contribution to country outcomes comprises all these instruments plus informal advice, policy dialogues, stakeholder convening, country teams' networks and know-how, and the day-to-day interactions between Bank Group staff and clients. As such, country engagement outcomes are meant to be more than the sum of their project inputs. It is at the country level that the clearest picture of the Bank Group's development impact should emerge and coalesce.

The Independent Evaluation Group (IEG) conducted this evaluation to help ensure that country-level results systems adequately support the outcome orientation of country engagements. IEG did not conduct this evaluation to assess the project-level results system or the Bank Group's achievement of country-level outcomes, which other IEG instruments, such as the Country Program Evaluations and the *Results and Performance of the World Bank Group (RAP)*, do. This evaluation is designed to enhance the Bank Group's results accountability, informed risk taking, and organizational learning. The evaluation provides insight into whether country-level results systems work as intended, which is to provide useful, timely, and credible information on the Bank Group's performance and contribution to country outcomes. The country-level results system is meant to serve two purposes: (i) upholding results accountability to clients, the Board, and senior management, and (ii) informing country teams' decision-making, learning, and adapting. There can be tensions between these dual purposes, and this evaluation examines how country teams navigate these tensions. The country-level results system is made up of tools (such as the CPF results framework and the CLR, and its validation by IEG in the Completion and Learning Report Review [CLRR]), the data and evidence that compose these tools, and the rules, norms, routines, behaviors, and organizational processes that shape how the system works in practice (Højlund 2014; Leeuw and Furubo 2008; Rist and Stame 2006). The country-level results system sits between project-level reporting—which is embodied by results and performance information for investments, guarantees, advisory services, and discrete lending operations—and Bank Group-wide aggregate reporting, which is embodied by Corporate

Scorecards and International Development Association (IDA) results frameworks. Both project- and corporate-level reporting are outside the scope of this evaluation.

Figure 1.1. Conceptual Framework



Source: Independent Evaluation Group.

Note: The country engagement is meant to work as follows: First the World Bank Group sets out to diagnose the biggest development constraints faced by their clients through a Systematic Country Diagnostic (SCD). Second, on the basis of the SCD findings, the Bank Group circumscribes specific contributions it will make to country development outcomes, at the intersection of the Bank Group's comparative advantage, what others are doing, and the clients' demands and priorities. The Country Partnership Framework (CPF) establishes these priorities, makes explicit the Bank Group's objectives and the theory of change that links the Bank Group's activities to the desired outcomes, and translates them into a results framework. Third, through a Performance and Learning Review (PLR), the Bank Group takes stock of progress and decides on needed changes and course correction. Fourth, at the end of the cycle, the Bank Group undertakes a Completion and Learning Review (CLR), a self-assessment of whether the intended objectives were achieved and of the lessons emerging from the cycle. The CLR is then validated by the Independent Evaluation Group.

The evaluation focuses on how country teams carry out four key practices of outcome orientation. These four practices include how country teams (i) aim for outcomes in the CPF and CPF results framework; (ii) monitor and evaluate the Bank Group's contribution to country outcomes and learn from this evaluative evidence; (iii) adapt country engagements; and (iv) engage clients to measure and manage outcomes. The evaluation deliberately focused on country teams as the “frontline” agents tasked with making the country engagement model work in practice. A country team refers to the core staff working on strategy and operations for a Bank Group country engagement. This team includes the Country Management Unit (CMU), particularly the country director, country manager or resident representative, operations manager, country officers, country program coordinators, and key sector staff, including practice managers and program leaders who drive sectoral strategic direction. The country team also includes International Finance Corporation (IFC) staff, particularly the IFC country manager and IFC country officers.

Methods

The evaluation's conceptual framework focused on the organizational conditions that enable country teams to manage for outcomes. The scientific literature identifies three broad categories through which country teams can use well-functioning results systems to inform adaptive decisions and make strategic course corrections (for example, ADB 2012; Dahler-Larsen 2011; Leeuw and Furubo 2008; Højlund 2014; OECD 2019; OECD and World Bank 2014; and World Bank 2016a, 2017d.) These categories include individual motivation, team capabilities, and management practices, all of which are featured at the bottom of figure 1.1. Individual motivation and incentives matter significantly. In international development, the staff's intrinsic motivation to achieve results must be reinforced by extrinsic incentives, such as “carrots, sticks, and sermons” (Bemelmans-Videc, Rist, and Vedung 1998; Levine and Savedoff 2006). The staff's autonomy to use outcome information and make decisions to adapt interventions and portfolios is also essential (Honig 2019; Ramalingam, Wild, and Buffarid 2019). Staff incentives and autonomy are largely a consequence of management practices, which are influenced by organizational rules, signals received from the authorizing environment,

and managers' leadership. Team capabilities, including staff skills, methods, knowledge, learning support, and networks, are also central to building outcome-oriented behaviors (OECD 2019; Vähämäki and Verger 2019).

The evaluation carefully selected a sample of 39 country engagements to generate evidence from the Bank Group's diversity of clients and portfolios. These were enough cases to enable comparisons within categories and to identify challenges or opportunities that cut across country contexts and engagements. Among the 39 countries, the evaluation collected evidence from interviews for 10 countries, from desk reviews for 15 countries, and from both for 14 countries. IEG also planned to visit six countries to interview additional clients and partners, but two of these trips were canceled because of the COVID-19 outbreak. The evaluation's detailed methodology is presented in appendix A.

The evaluation team undertook three types of data collection and analysis activities. These activities included the following:

- » In-depth interviews: First, the evaluation team completed in-depth interviews with country teams. These interviews were essential for understanding how country teams interact with country-level results systems, what challenges they face in so doing, how teams make adaptation decisions, and how incentives and management signals influence staff behavior. IEG interviewed clients and partners in four countries, supplemented by interviews with Bank Group management in Regions, Operations Policy and Country Services, IFC, and the Multilateral Investment Guarantee Agency (MIGA). In total, the evaluation team conducted 185 interviews. Interviews focused on the sampled countries but included the interviewed staff's experiences from other countries as well.
- » Country product reviews: Second, the evaluation team conducted a structured review of 65 country products for 29 countries to understand what is being reported externally in the country-level results system. The team used a classification that was developed for the RAP 2020 to identify how country teams frame and measure objectives in results frameworks. IEG also analyzed PLRs and CLRs to identify types of country engagement adaptations and what evidence and justifications country teams use for these adaptations in formal

documents. To help do this, IEG adapted a typology of portfolio-level adaptive management practices developed by the Global Learning for Adaptive Management initiative.

- » Literature reviews: Third, the evaluation team conducted literature reviews to identify commonalities and idiosyncrasies of the Bank Group's country-level results system and to identify ways to improve on it. The team reviewed 16 evaluation reports from IEG and the World Bank and 80 external reports from other development partners on results-based management. IEG used the external reports to understand adaptive management practices carried out by governments, the private sector, and philanthropic groups over a range of organizational contexts.

The report presents evidence and findings on how the Bank Group aims for, captures, and manages country-level outcomes. Chapter 2 reviews how country teams set objectives, define development pathways, and generally aim for outcomes in their country engagements. Chapter 2 finds that the Bank Group's country-level outcome model is sound and country teams use this model well, with some room to be more selective and work better across Bank Group institutions. Chapter 3 examines how country teams measure the Bank Group's influence on countries' development outcomes. It finds that the country-level results system does not properly represent the Bank Group's influence because of the system's overreliance on results frameworks, which are unfit for this purpose. Chapter 4 reviews how country teams practice adaptive management and whether country-level results systems support them in making course corrections. It finds that country teams practice adaptive management in several ways: they closely monitor the health of country portfolios, they take action to address delivery issues, and they navigate changes in country contexts. However, country teams barely use country-level results systems to inform their adaptation decisions. Chapter 5 lays out a way forward for country programs to improve their outcome orientation. It makes specific suggestions for reforming the country-level results system so that it is better fit for this purpose. The report's appendixes present the evaluation team's underlying analyses.

2 | Aiming for Outcomes

Highlights

The model of how the World Bank Group aims for outcomes at the country level is sound and, for the most part, country teams practice this model well.

The Bank Group's country strategies set out to influence country outcomes and frame their objectives in terms of outcomes, while facing difficulties in being selective because of internal competition for program space and heterogeneous client demand.

The Bank Group pursues its objectives through both direct and indirect pathways, but indirect pathways are not well articulated in country documents and the relatively short country engagement cycle does not promote the long-term vision needed to sustain indirect pathways.

There has been progress in integrating teams from the International Finance Corporation (IFC) and the Multilateral Investment Guarantee Agency into Country Partnership Framework design, but Country Partnership Frameworks do not serve IFC's strategic needs well. This motivated IFC to launch its own internal country strategies and other country products. Early experience with these IFC country strategies shows that they offer potential for increasing IFC's focus on country outcomes; close alignment with the World Bank and Bank Group country engagement is a condition of success and of avoiding risks of duplicated efforts.

This chapter reviews how country teams aim for outcomes in their country engagement. The country engagement is not meant to be the simple aggregation of the Bank Group's discrete projects taking place in a country. Its value proposition is in providing a strategic platform for articulating and coherently managing the Bank Group's response to country needs. It is intended to pursue meaningful and significant objectives, to combine instruments in the pursuit of short- and long-term development trajectories, and to combine the strengths of the three Bank Group institutions.

Objectives for Country Engagements

The Bank Group's country strategies set out to influence country outcomes and frame their objectives in terms of outcomes. The CPF is the institutional platform created to identify the strategic direction of the Bank Group in each country. Country teams use this platform well to articulate how the Bank Group will contribute to the longer-term development trajectory. Aided by the SCD and in-depth consultation, the CPF narratives spell out the country's most critical development challenges and how Bank Group interventions will contribute to address them in explicitly identified outcome areas and sectors, each with its own objectives. The evaluation applied a classification plan developed for the RAP 2020 to identify the range of objectives pursued in CPF and indicators used to capture them (box 2.1). In the CPF and outcome areas reviewed for this evaluation, nearly all objectives were framed in terms of outcomes. A quarter sought to influence long-term country outcomes, seeking a major change in service delivery or governance systems or a sustained benefit to citizens (level 4); 15 percent sought to change the behavior or capability of key government agents or ultimate beneficiaries (level 3); and while the majority pursued intermediate outcomes (level 2)—whereas virtually none were defined in terms of outputs (level 1). There is significant variation across outcome areas, however, with strategies more likely to frame intangible goals such as governance in terms of high-level outcomes (figure 2.1). Interviews also made clear that country teams understand well the country outcomes they are working toward.

Box 2.1. Objective and Indicator Classification

Level 1—Output. Product or service provided within the control of the client.

- » For example, creating a national performance management framework (Botswana CPF FY16–20).

Level 2—Immediate outcome. Development of the capability of a group or organization, an initial benefit to people, infrastructure delivered.

- » For example, improving access to quality education and health services in targeted rural areas (Indonesia CPF FY16–20).

Level 3—Intermediate outcome. Stakeholders apply a new capability to solve an issue; change in the lives of ultimate beneficiaries.

- » For example, enhanced educational outcomes for better employability (Bulgaria CPF FY17–22)

Level 4—Long-term outcome. A sustained change in delivery or governance or a sustained benefit to an ultimate beneficiary.

- » For example, improving living standards in the lagging areas (Sri Lanka CPF FY17–20).

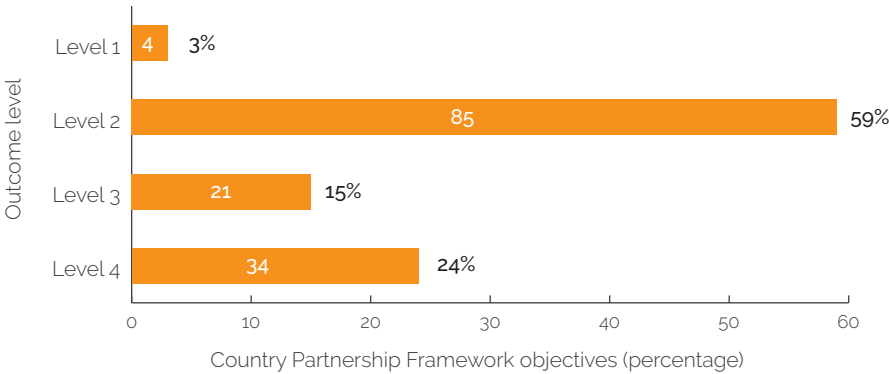
Source: World Bank 2015b, 2015f, 2016b, 2016e, 2020b.

Note: CPF = Country Partnership Framework.

In selecting objectives for country engagements, the most commonly observed challenge is the difficulty of being selective because of internal competition for program space. Among the 29 CPFs reviewed for this evaluation, CPFs with more objectives were less likely to have good-quality indicators that adequately measured the objective. In interviews, staff argued that the Bank Group “spreads itself too thin in CPFs,” which reduces the chances of significantly influencing country outcomes. Interviewees believe the lack of selectivity is driven by the World Bank operational model whereby Global Practices jockey for space in CPF envelopes, in part for the associated work programs and budget but in part because of the institutional incentives for

them to maximize lending portfolios. As one staff member put it, “I know there is no space in the portfolio for my sector, but my practice manager still pushes me to try to get a project in.”

Figure 2.1. Country Partnership Framework Objectives by Outcome Level



Source: Independent Evaluation Group.

Note: n = 144 Country Partnership Framework objectives. Level 1: Outputs; level 2: Immediate outcome; level 3: Intermediate outcome; level 4: Long-term outcome.

Challenges in selectivity are also driven by heterogeneous client demand. Client demand is a key factor in country program design. As a financial institution, the Bank Group needs to balance strategic and opportunistic responses to development challenges to generate the lending on which its business model relies. This is particularly true of IFC and MIGA investments and guarantees. In certain contexts, the heterogeneity of client demand significantly constrains country teams’ ability to aim for a selective and sustained contribution to country outcomes. Especially in countries where Bank Group financing volumes are small compared with overall investment, it can be unrealistic to expect to achieve significant effects on country outcomes. In countries where the Bank Group faces significant competition for lending, the imperative to lend can make it difficult to be selective, so CPF objectives are deliberately written to be very broad to ensure that any demand that arises can be classified as relevant to the strategy. Both effects are prevalent in large upper-middle-income countries, which tend to articulate broad high-level objectives. Similarly, in countries with very fast-changing government priorities, such as Peru or Sri Lanka, of necessity, the CPF is written in broad terms to give country teams room to maneuver during the cycle with fast turnaround.

Country leaders use a range of practices to curb competition and enhance selectivity to achieve country objectives. Country managers and directors play a pivotal role in catalyzing the Global Practice (GP) energy and maintaining the focus of the portfolio. Some are more effective than others. Interviewees consistently see leadership skills as the most critical factor, but some tools and filters have also proven useful. For example, drawing lessons from the CLR, the Zambia CPF used spatial alignment to tighten the selectivity of the use of IDA funds. It identified three spatial circles as the poorest areas of the country and put in place mechanisms to prioritize investments in these high-need areas. For the preparation of the Morocco CPF, the country leadership organized multi-GP results teams around focused outcomes. This took away the competition for securing a place in the CPF for a particular GP. The Morocco CPF was also designed as an adaptive framework in contrast with more conventional practices, as shown in table 2.1. A CMU staff member described it as “a living strategy that incorporates learning-by-doing and constantly adapts to country circumstances and unforeseen or changing priorities.”

Table 2.1. Morocco's Adaptive Country Partnership Framework in Contrast with More Conventional Approaches

From	To
CPF as a document	CPF as a living strategy that needs adjustments
Overly ambitious project design that does not get implemented	Experimental design that learns early and incorporates lessons as it scales
Linear log-frames and causal chains	Recognition that the World Bank Group works on complex problems in complex systems
Measuring what is known	Measuring what is important
Supply-driven, sector-focused interventions	Multisectoral programmatic approaches
Focusing on project design	Focusing on implementation and implementation support
Reputational risk as a key driver for decisions	Learning mind-set to achieve results and generate capacity by doing

Source: World Bank 2019m, 37.

Note: CPF = Country Partnership Framework.

In Romania, the country manager uses CPF filters to provide a strong strategic direction to sector teams developing new business in the country. All new operations and reimbursable advisory services (RAS) have to demon-

strate how they contribute to institutional strengthening, which is the Bank Group's main contribution to Romania's development. In the Western Balkans, the country director is requiring that all proposed advisory services and analytics (ASA) clearly articulate how they contribute to a specific outcome area of the CPF and that they include an outcome indicator. In the Pacific Islands, the CMU has established a ceiling on the number of projects that can be prepared for a country within an IDA cycle to avoid overwhelming thin capacity. This has helped delimit the sectors where the World Bank is active.

Pathways to Achieving Development Objectives

The Bank Group pursues its objectives through direct and indirect pathways to influence clients' development trajectories. Bank Group country engagements have direct effects from investing in assets, services, or other measures that yield short-term benefits to citizens that are tangible, measurable, and attributable to specific interventions, and indirect effects that strengthen institutions, improve the quality of governments' spending, or catalyze private sector development. These indirect effects also come from fostering learning-by-doing in governments and market institutions and can originate from both project and nonproject instruments. By engaging with Bank Group interventions, clients gain capacity in project management, procurement, financial management, safeguards, and so on, ultimately improving how governments function. They are among the most common CPF objectives. About one-third of CPF objectives coded for this evaluation explicitly aimed at building institutional capacity within a sector. Country teams working in International Bank for Reconstruction and Development (IBRD) countries consider that this is the most significant contribution to country outcomes, because direct investment in service delivery can only have marginal effects. For example, in European Union (EU) member or candidate countries, the World Bank is a small funder compared with the EU but is playing a key role in addressing the institutional weaknesses (institutional bottlenecks and weak capacity in managing funds, implementing projects, and carrying out policy changes) that prevent the countries from taking full advantage of the financial support provided by the European Commission. The Bank Group's main contributions in Romania and the Western

Balkans are focused on addressing these challenges, mainly through RAS for the former and a mix of technical assistance and lending in the latter.

However, indirect pathways to country outcomes are not well articulated in country documents. In part, this is because theories of change for indirect pathways are not as well articulated compared with those of direct pathways. Although indirect pathways are one of the most important contributions of the Bank Group to country outcomes, there are still many unknowns on how institution building connects to country outcomes, including identifying the type of support that leads to better institutions and understanding under what conditions stronger institutions lead to better outcomes for citizens (World Bank 2019f). The results measurement system is also not designed to encourage teams to emphasize indirect pathway effects. The effects of these pathways for citizens do not materialize immediately, causal chains are long, and the influence of specific actors is harder to trace and the effects harder to quantify. For example, in the Solomon Islands, Bank Group projects on rural development and youth employment have probably played a contributing role in peace and stability through project activities that build social dispute resolution mechanisms, increasing rural incomes, improving service delivery, and providing jobs in some areas. The country team was dissuaded from emphasizing those types of changes in their country evaluation products, because they are hard to quantify and to attribute to specific interventions.

The relatively short CPF timeline does not match the long-term vision needed to sustain indirect pathways. Interviews with country teams show that though indirect pathways and long-term development trajectories are part of their thinking, the platforms created by the country products do not promote a long-term vision. There is limited discussion in CPFs or other country documents about how each cycle builds on past achievements and contributes to long-term development trajectories. Although CPFs use the active portfolio as a baseline and the closing portfolio as the end line to assess results, there is no line of sight on the long-term development trajectory. In our sample of CPFs, there were only four country engagements reviewed that provided a sense of how the Bank Group will seek to maintain and strengthen outcomes over multiple country engagement cycles (Afghanistan, Romania, North Macedonia, and Vietnam).

Aiming for Development Objectives across the Bank Group Institutions

The CPF is intended to provide an overarching strategy incorporating not only the World Bank but also IFC and MIGA. The Bank Group tries to pursue objectives through mixing and sequencing multiple interventions to achieve outcomes greater than the sum of their parts; however, this is not always possible. At the CPF stage, country teams articulate well how the combination of Bank Group operations aims to influence certain outcomes. When such packages exist, the intervention logic laid out in the CPF narrative and in the dedicated portion of results matrixes explains quite well how multiple interventions, including ASA, private sector investment, and guarantees, contribute to each objective. However, where portfolios of activities within sectors are small, articulating a results chain that conveys cumulative effects across operations and contributions to higher-level objectives can be artificial. For instance, in the Western Balkans or the Pacific Islands, portfolios tend to have a single project in any given sector that can be the sole vehicle for influencing outcomes. In these contexts, project objectives and country objectives are largely the same, and so the country-level results system adds little value to the project-level results system.

There has been progress in integrating IFC and MIGA teams into the CPF design, but CPFs do not serve IFC's strategic needs well. According to interviews with IFC, MIGA, and World Bank staff, the inclusion of IFC and MIGA teams in the preparation of the SCD and CPF has improved over time. In the sample of CPF objectives reviewed for this evaluation, IFC and MIGA were integrated into 40 percent of the CPF objectives, with 39 percent as contributors to Bank Group-wide objectives, and 1 percent as a stand-alone IFC/MIGA objective. MIGA's risk management officers need to have an in-depth understanding of the country to arrive at a risk rating, which is at the core of MIGA's business model. Thus, MIGA values highly the knowledge sharing and network building across the Bank Group that comes with the country engagement model. MIGA participates in the CPF process in all the countries for which the Bank Group is engaged but prioritizes its involvement in countries that are of strategic priority, primarily IDA and fragility, conflict, and violence (FCV) countries. Conversely, more than two-thirds of IFC staff interviewed contin-

ue to consider the CPF as a World Bank-centric document geared toward an external audience and with limited influence on IFC's strategic outlook. The content is considered too upstream and detached from IFC's business model, which engages with specific corporate clients rather than governments. It is not used by IFC as a road map for action within countries. Staff consider it inherently difficult for the CPF to represent IFC well and to serve its strategic needs; the strategy discussions and document concentrate on the larger World Bank engagement, and the consultations prioritize the single government client rather than numerous private sector clients.

IFC launched its own Country Strategies (CSs) in 2016 to improve IFC's inputs to the CPF and to better guide its investment decisions within countries. IFC has introduced a number of internal country products with the intention of informing the Bank Group country cycle: the joint IFC-World Bank Country Private Sector Diagnostics to inform and complement the SCD, the IFC CSs to lay out strategic priorities and reinforce IFC's position in the CPF, and a country-level business plan and its subsequent reviews. The CS articulates clear investment and sometimes development impact scenarios captured by "if-then" matrixes, which lay out critical policy barriers where reform (and potentially World Bank engagement) could unlock investment opportunities. The process of developing CSs has created a useful platform to bridge the gap between country and industry teams. In interviews, staff report largely positive feedback on the consultations, team brainstorming, and leadership involvement in the preparation of the CS. Anchoring advisory and investment decisions in a country framework marks a departure from business as usual at IFC. Acceptance by industry teams of this country lens is driven by the IFC 3.0 strategy, the commitment of the senior management team (regions and industries), and the ability of the CS to add value to business development activities.

Early experience with IFC CSs suggests they offer potential for increasing IFC's focus on country outcomes. The process of combining an assessment of country development gaps and IFC's comparative advantage with market data and analytics and a detailed understanding of regulatory and capacity barriers enables IFC to develop a common road map of priority sectors for engagement. This has the potential to enhance the internal coherence of IFC interventions and depth of impact by influencing which sectors IFC is active

in. When there is strong support from both regional and industry directors, CSs can influence project decisions by guiding business development efforts and discouraging project proposals that are not well aligned with CS priorities. Links to investment scenarios and expected developmental impact are made more explicit in if-then matrixes. IFC CSs contain investment scenarios that are articulated based on the likelihood of specific reforms taking place. In scenarios where few reforms take place, the matrix identifies the IFC 1.0 and 2.0 investments that could take place by industry. The matrix also identifies IFC 3.0 projects that could materialize in high-reform cases. These matrixes make explicit specific IFC requests from the World Bank or the government to work on reform processes that could unlock IFC investments. All if-then matrixes articulate an estimated IFC investment target for each scenario. However, only a few lay out a developmental impact objective linked to the scenarios. This practice is a good marker of enhanced country outcome orientation and should be encouraged but need not include quantified targets that could create false expectations of precision.

Close alignment with the World Bank and Bank Group country engagement is a condition of success. IFC's country products are supposed to feed into the Bank Group country engagement cycle. The first set of IFC Country Private Sector Diagnostics and CSs were not always timed to be integrated in the SCD and CPF, and so the CPFs did not capture their contents, leading the two to function as parallel and duplicative systems. However, sequencing of IFC and Bank Group country products has been improving over time. Moreover, the investment scenarios and necessary reforms are not always discussed with the government counterparts, which means that some of the assumptions on which IFC's CSs rest can become "wishful thinking," as one interviewee put it. The degree of alignment and joint commitment between the Bank Group and IFC on the if-then matrix is highly dependent on the relationships between World Bank country director and IFC regional leadership. It was clear from interviews with IFC staff and management that the proliferation of strategy-related documents puts a strain on IFC staff. Interviewees highlighted that close coordination with World Bank colleagues is necessary to avoid duplicative efforts and ensure alignment between IFC and Bank Group country products, but this comes with transaction costs.

3 | Capturing Outcomes

Highlights

The World Bank Group's country-level results system does not capture the Bank Group's contributions to country-level outcomes because it relies on results frameworks predicated on three principles that do not fit well the nature of country programs: metrics, attribution, and time-boundedness.

Country Partnership Framework results frameworks do not capture well the country strategies' objectives or the Bank Group's contributions to country outcomes, which are difficult to measure because they are often not quantifiable and because of data gaps in many countries.

Country strategies lay out future objectives, but Country Partnership Framework results frameworks primarily capture past operations and do not capture medium- to long-term effects. Most country strategy indicators come from lending projects and do not properly capture the contribution of advisory services, analytics, policy dialogues, or stakeholders' convening. The contributions from the International Finance Corporation and the Multilateral Investment Guarantee Agency are also not well captured.

There are costs associated with generating indicators that are not used.

The Bank Group's country self-evaluation product, the Completion and Learning Review, is little used within the institution and provides a partial picture of Bank Group contributions.

Although some country teams carry out other evaluative exercises to assess the achievement of outcomes, these are rare and overshadowed by the results system's emphasis on targets and indicators, an emphasis that also consumes the time of monitoring and evaluation staff.

This chapter reviews how country teams measure the Bank Group's influence over country outcomes. There are several tools used in the country-level results system: CPF results frameworks, which are developed at the CPF design stage; the CLR, which is a mandatory self-evaluation of the country engagement carried out by the country team at the end of the cycle and is subsequently validated by IEG in the CLRR; and other nonmandatory evaluative activities. This chapter reviews how country teams use these three sets of tools. The chapter shows that country-level results systems provide an incomplete representation of the Bank Group's contribution to country outcomes. The cause of this does not lie in country teams' practices—they strive to faithfully follow the guidelines—but rather in the system's measurement principles and tools and their institutional execution.

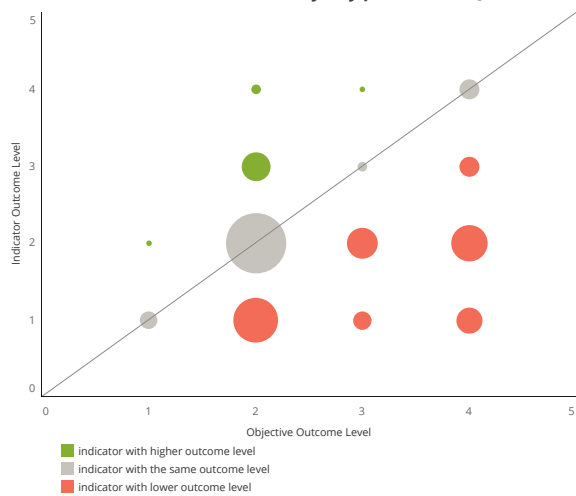
Results Frameworks

The CPF results framework is the cornerstone of the Bank Group's country-level results system. CPF results frameworks, also referred to as “results matrixes,” assign a set of indicators and targets to CPF objectives. CPF results frameworks are at the beginning of the country engagement cycle, can be revised midway through the cycle in the PLR, and are used to rate the Bank Group's results and performance at the end of the cycle in the CLR. As the principle of results-based management expanded from the project level to the country level, the main features of the project results measurement system were transposed to country programs (World Bank 2016f). The system measures objectives through time-bounded quantifiable metrics that are directly attributable to Bank Group interventions. Like projects, country programs are rated, through an objective-based assessment, on whether they have achieved their stated objectives during the CPF cycle. The ratings methodology adopted by the Bank Group and IEG assesses the extent to which the target indicators for each objective have been met and whether the indicators sufficiently reflect the objective. The purpose of this approach is to promote accountability for delivering results by judging success or failure by how far results ex post deviate from targets that were forecast ex ante, through indicators that can be verified by external validators and are time bound, as only achievements made during the CPF cycle are counted. The rest of this chapter shows how these principles shape country teams'

practice and inadequately capture the Bank Group’s contributions to country-level outcomes.

CPF results frameworks’ measurements do not capture well the objectives pursued and about half the time are framed below the level of the objectives they seek to capture, particularly because of the expectation of attribution. A comparison between CPF objectives and indicators shows a sizable mismatch between the outcome focus of objectives and their associated indicators. More than half (52 percent) of CPF objectives are measured by indicators that are below the objective’s level. For example, figure 3.1 shows that when the Bank Group pursues intermediate or long-term outcomes, it tends to measure them with lower-level indicators. In interviews, staff suggested that this is because the guidance on CPF results frameworks limits them to selecting indicators that are mostly attributable to Bank Group efforts and can be verified by IEG. Country staff are also understandably reluctant to be held accountable for achieving a target indicator that is beyond their control, as illustrated in figure 3.1. They also find it challenging to identify quantifiable, time-bound, attributable, and verifiable indicators that can capture the CPF’s objectives. As a result, the indicators they select often do not accurately reflect the CPF’s objectives or fundamentally capture the country engagement’s most important development results.

Figure 3.1. Indicator Outcome Level by Type of Objective



Source: Independent Evaluation Group.

Note: n = 361 Country Partnership Framework indicators. Level 1: Outputs; level 2: Immediate outcome; level 3: Intermediate outcome; level 4: Long-term outcome. The size of the circles is proportional to the share of indicators at each level.

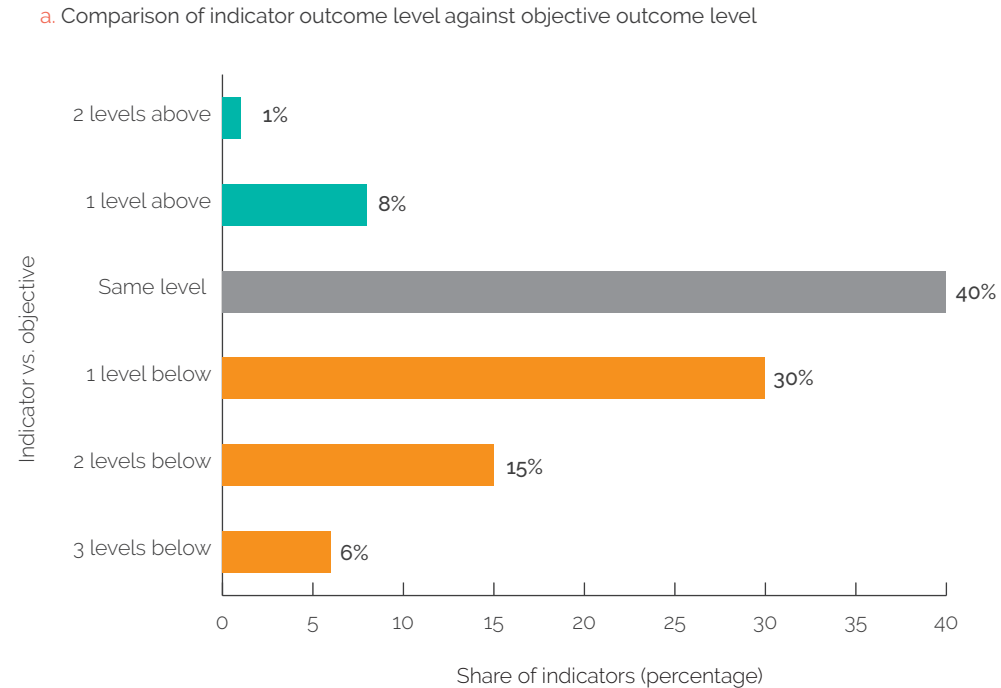
Although CPFs lay out objectives for the future, their results frameworks primarily capture past operations. There is a built-in timing mismatch between CPF objectives and their measurement that country teams cannot reconcile. Country teams are asked to lay out objectives for the future that support high-level country outcomes; however, they are also tasked with finding result indicators that can be achieved and measured within a typical three-to-five-year CPF cycle (World Bank 2019u, 20). Because it often takes many years for a Bank Group intervention to move from conception to impact, in practice this means that when country teams are designing a CPF, they must select indicators from interventions that were designed long before the CPF objectives are formulated. This is because these result indicators will materialize and be measurable by the end of the CPF, even though they may not align well with the new CPF's objectives. Alternatively, teams must speculate on the measurable results of a forward-looking pipeline, which can be uncertain, especially for private sector investments or in volatile country contexts. Country teams' widely shared sentiment is that this approach to developing results frameworks is artificial and detracts from outcome orientation. There are cases where this mismatch is particularly salient. An extreme example is Vietnam, where new lending dropped dramatically after the country's graduation from the IDA, from roughly \$10 billion in lending to \$67 million in the first year after graduation (though this is expected to rise). Because of this, roughly 70 percent of Vietnam's CPF results framework was based on the legacy portfolio.

The country-level results system does not adequately capture the medium- to long-term effects of Bank Group interventions. A consequence of the reliance on metrics and on project-level monitoring systems is that results frameworks capture only the results that occur by the time a project closes. Yet many effects, especially indirect effects, take longer to manifest. For example, some of the most important benefits from slum upgrading and off-grid power projects in Vietnam were the way in which they influenced the design of much larger government programs. Similarly, IFC's support for mobile phone providers in the Pacific Islands has been transformative for telecommunications, but it took time for these effects to be fully realized. These types of effects are not well captured by the country-level results system.

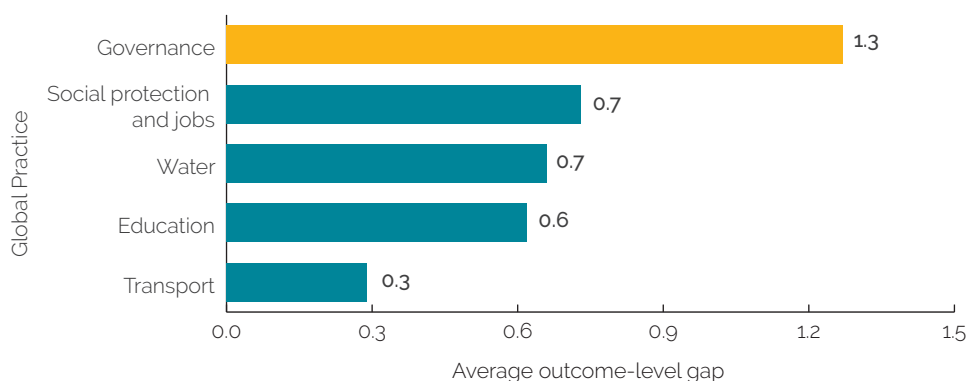
The country-level results system relies on quantitative metrics, but these do not capture many of the Bank Group's outcomes at the country level. There

are several elements in the CLR methodology that privilege metrics and numbers. For instance, the guidance emphasizes that indicators should be observable and externally verifiable by IEG. Yet, as one staff member noted, “By focusing on what is measurable, we are not capturing what is meaningful or useful, especially in terms of the big picture.” Most interviewed Bank Group staff perceived metrics as a challenge, or as a distortion, and shared the concern that the current results measurement system does not measure the most valuable Bank Group contributions to country outcomes, especially in sectors that are inherently harder to measure than others. To illustrate, among the five sectors reviewed by IEG, the discrepancy between objectives and indicators is the sharpest in the governance outcome area, which includes many of the Bank Group’s institutional capacity-building efforts and is the hardest to adequately capture with simple metrics (figure 3.2). In other sectors—where the Bank Group can deliver countable, verifiable outputs such as building a road network—the discrepancy is smaller.

Figure 3.2. Discrepancy between Objective and Indicator Level of Outcomes



b. Average outcome-level gap between indicators and objectives for sampled sectors



Source: Independent Evaluation Group.

Note: n = 361 Country Partnership Framework indicators. The Independent Evaluation Group calculated the average difference between the outcome level of objectives and their attached indicators and found that, on average, indicators are formulated at an outcome level that is below the outcome level of their measured objectives. This measurement gap is higher for the Governance Global Practice, where the formulated indicators are, on average, 1.3 levels below the outcome level of their measured objectives.

Data gaps in countries also make it difficult for country teams to select outcome indicators. Many countries, particularly IDA and FCV-affected countries, lack strong statistical and monitoring and evaluation systems; some countries are also reluctant to share available data, and therefore have less data on country outcomes, even for sectors that are easier to measure (World Bank 2017a). Some sectors, such as governance, that seek system changes are inherently difficult to measure. In other sectors, like health, energy, or social protection, outcome data can be routinely collected through regular sectoral operations (for example, electricity or energy generation and consumption, prices, access, system losses, stunting rate). But even for these, many World Bank clients do not have adequate information systems to track development progress. For instance, the 2018 Global Partnership Monitoring Round report reveals that 64 percent of countries have good-quality national development strategies, but only 35 percent of governments have data and monitoring and evaluation systems to track these strategies (OECD and UNDP 2019). Country teams working in IDA, and especially in FCV countries, emphasized that they struggle with scarce and low-quality data. For example, in the Solomon

Islands, there have only been two major household surveys in the past two decades, and methodological differences between them mean they are not easily comparable, limiting the country's ability to monitor poverty. FCV governments often lack a system to track their own national development strategies and rely on development partners to track even basic outputs. However, partners have their own results measurement and reporting systems, making data from different partners difficult to reconcile. In many FCV contexts, poor security situations have made basic supervision and monitoring tasks difficult to carry out. The Bank Group has made efforts to mitigate these issues by investing in third-party monitoring and other monitoring and evaluation platforms, such as the Geo-enabling Initiative for Monitoring and Supervision, to compile data collection and map portfolios. But these solutions are only short-term fixes and do not address the larger data scarcity issues in FCV countries.

Box 3.1. Illustrating the Mismatch: Nicaragua Early Childhood

Nicaragua preschool education: In Nicaragua, the Country Partnership Strategy (CPS) FY13–17 sought to increase coverage of preschool education (ages three to five; World Bank 2012b). The team included an indicator on access in the CPS results framework, but the indicator was dropped at the Performance and Learning Review stage. The reason was that it was considered too ambitious to expect results within one CPS cycle. The outcome indicator was replaced with an output measure or milestone that was within the control of the World Bank "system to assess early childhood development outcome in place and operational." Once the CPS cycle closed, the Completion and Learning Review failed to take stock of progress toward access. At the beginning of the next cycle, the Systematic Country Diagnostic noted that Nicaragua had the lowest rates of access to preschool of Central America, after Guatemala (World Bank 2017c, 75), thus again signaling the importance of moving toward increased access. Yet the Country Partnership Framework did not measure progress toward this outcome and again included a more output-oriented measure within the control of the World Bank Group: the percentage of preschool teachers with desired teaching practices (World Bank 2018d). In short, two country engagement cycles have passed, and the World Bank Group's documents are still unclear on whether progress in access has been achieved and how Bank Group interventions influenced these changes.

Most CPF results indicators are taken from lending projects. For instance, among the 361 CPF indicators reviewed by IEG, two-thirds were traced to indicators of a single or multiple Bank Group projects. This was hard to calculate because 63 percent of indicators did not list their source in country documents, so IEG had to collect this information through the World Bank's operations portal, the World Bank's operations database, and the Bank Group's project documents. The remaining one-third of indicators that did not come from projects typically measure an action taken by a government with the Bank Group's support—usually through development policy financing—and only rarely measure changes to country outcomes. Interviewees explained that country teams harvest indicators from discrete lending and investment projects because this ensures the data are available and attributable to the World Bank's interventions. Another reason for using project indicators is that monitoring and evaluation is funded through projects but under-resourced at the country level. Country teams try to avoid additional evidence gathering because of the limited resources and expertise available.

CPF results frameworks do not properly capture the contributions of advisory services, analytics, policy dialogues, or convening power. The country-level results system is meant to capture the effects of various types of interventions, lending and nonlending. Interviewed World Bank staff recognized the challenge of quantifying verifiable metrics for analytical and capacity-building work. At the same time, there is emerging empirical evidence—from AidData microlevel surveys with 1,244 public sector officials in 121 developing countries—that the World Bank has substantially greater influence over the direction, design, and implementation of government policies than most of its bilateral and multilateral peers. This policy influence is at least in part because of the World Bank's analytical work, advisory services, and technical assistance. Country teams, especially staff interviewed working on IBRD countries, find it difficult in the current results system to do justice to the World Bank's ASA contributions to institutional strengthening and capacity building. Box 3.2 illustrates an attempt at going beyond what is expected from the results system and carrying out a dedicated evaluation of the CMU's RAS program. IFC has developed a project-level evaluation system for advisory services that seeks to capture the influence of the project on direct recipients of the advice as well as the extent to which project activities have made a difference at the market or sector level.

However, IFC's contribution is underrepresented in CPF results frameworks and a very small number of indicators in the sample reviewed captured the results of IFC's advisory services.

Box 3.2. Evaluating Advisory Services in Romania

The Romania country engagement has a large reimbursable advisory services (RAS) program, and the country office decided to commission two evaluations of its program to better understand its contribution to institutional change in the country. The first covered the RAS program between 2012 and 2015; the second is in process. The first evaluation assessed the 39 RAS agreements conducted with line ministries, government agencies, and local authorities for a total value of approximately \$59 million. The evaluation classified the type of capacity challenges that the RAS program was seeking to address and traced the effect of RAS along an institutional capacity change pathway. The evaluation concluded that the World Bank had strengthened institutional capacity across 15 Romanian agencies at the central, regional, and country levels through three main channels: enhancing the effectiveness of organizational arrangements, enhancing the efficiency of policy instruments, and strengthening the ownership for reforms among various levels of institutional actors. The RAS program had played a key catalytic role in helping Romania leverage European Union funding for achieving development objectives. The Romania country office found the evaluation useful and better adapted to its knowledge and accountability needs than Country Partnership Framework results frameworks. To dig deeper into assessing and understanding its contribution to Romania's institutional strengthening, the country team is partnering with the Governance Global Practice to carry out a second evaluation and test specific parts of its theory of institutional change.

Source: World Bank 2016d.

Staff raised concerns that country-level results systems have a disproportionate focus on lending projects, rather than on other instruments, and on medium-term goals rather than long-term outcomes. Current country-level results systems collect information primarily on lending and investments because they supply measurable indicators of progress. However, this approach misses the results contributions from other instruments, such as

capacity building, analytical work and policy dialogue, and combined interventions; thus, it does not promote synergy across interventions to support objectives. There are signs that complying with the country-level results system has subtle but pernicious effects on teams' behaviors. Staff argue that if the results system disproportionately focuses on lending, it can have the adverse effect of driving attention and resources away from other instruments that in certain contexts are better suited to pursue certain types of outcomes. Many staff also argue that the emphasis on results frameworks and attributable time-bound indicators has other negative effects. It encourages country teams to pay too much attention to outputs and immediate goals rather than the big picture; it misses opportunities to identify synergies and complementarities across instruments; and it encourages risk aversion by favoring safe and achievable projects that might have solid but incremental effects rather than risky endeavors that might have transformative impacts or might miss their targets.

IFC's and MIGA's contributions to outcomes are not well captured in CPF results frameworks. Forty percent of the CPF objectives reviewed for this evaluation mention IFC and MIGA as contributing to the objective, yet only 16 percent of CPF objectives had an indicator that captures these IFC or MIGA contributions. This gap is partially explained by IFC's and MIGA's project pipelines being largely unknown at the CPF design stage, making it difficult for teams to establish clear IFC and MIGA metrics and targets. Interviews with IFC and MIGA staff highlighted the challenges they face in linking their interventions with country outcomes. They find it more reasonable to link IFC and MIGA work to country outcomes in cases where they are supporting projects that are of national significance, such as when they are supporting the first in kind of a particular industry or technology or a major player with high market share. This is more common in certain industries, such as infrastructure. In the five sectors reviewed by IEG, the transport sector had the highest share of IFC or MIGA contributions covered by indicators (76 percent), followed by the water sector (44 percent). IFC investments into sustainable mining in Guyana or renewable energy in Argentina are examples of how IFC can generate impacts beyond a specific client (see box 3.3). The IFC Anticipated Impact Measurement and Monitoring and MIGA IMPACT tools incentivize teams to consider development impact when struc-

turing transactions by more systematically assessing potential impacts. For example, Anticipated Impact Measurement and Monitoring develops ex ante impact claims based on IFC’s anticipated market creation and stakeholder and environmental effects. IMPACT assesses the ex ante development effects of MIGA guarantee projects, including their ability to attract and mobilize foreign direct investment. These tools are designed to encourage staff to prioritize potential development impacts as well as financial returns. This is different from the challenge in the World Bank, which is to create a clearer line of sight between country programs and country-level outcomes.

Box 3.3. The International Finance Corporation's Contributions to Country Outcomes in Guyana

After International Finance Corporation (IFC) initial investments in Guyana Goldfields and Aurora Gold (mining), IFC played an increasingly active role in the country. IFC joined the World Bank’s efforts in playing an important role in the implementation of the of the country’s Low Carbon Emissions Development Strategy through its contribution to the development of sustainable mining, transfer of environmental management standards and practices, and introduction of appropriate resource use and conservation methods. In fact, IFC’s engagement with Guyana Goldfields helped establish a viable model for sustainable mining development in the country. It fostered local economic development, skills transfer, and building environmental management capacity. IFC has made a second equity investment of \$4.15 million in the project as a continuation of a partnership focused on introducing clean energy, preserving the biodiversity in the rain forest, and implementing international standards for sustainable mining in a remote region where artisanal mining practices had a significant adverse impact on the environment. The project contributed to local economic growth by generating more than 300 jobs and \$1.5 million in purchases of local goods and services.

Source: Independent Evaluation Group.

Country teams make great efforts to comply with the results system, though formal rules allow for some flexibility that is not fully used. Teams take the design of the results framework seriously and apply the methodology as intended. World Bank, IFC, and MIGA staff interviewed for this evaluation

say there is a lot of organizational effort and scrutiny that goes into designing CPF results frameworks. Staff tend to identify indicators that meet the system's basic requirements: for example, among the 361 CPF indicators IEG reviewed, most identified baselines (95 percent) and targets (94 percent). In most cases, staff felt there were not better indicators to use, though they recognized that the indicators they chose were often not meaningful. Overall, staff seemed resigned to complying with a system that did not really work for them. The formal guidance for results frameworks, the Shared Approach for Assessing CPF, does allow some flexibility that is not fully used. It does not require direct attribution; it allows for "alternative indicators" to be used to provide evidence when the results framework evidence is weak; and in principle it assigns ratings based on the degree to which objectives, rather than targets, are achieved. However, teams have internalized a stricter interpretation of the guidance because of advice provided by quality assurance processes, peer reviewers, and interactions between teams and Operations Policy and Country Services or IEG, which has established a strong culture of attribution. In particular, teams are understandably resistant to being held accountable for the achievement of objectives that are influenced by many other factors beyond their control. Yet relaxing this stricter interpretation would still not address the inherent challenges of the results system.

There are costs associated with compiling, processing, quality controlling, and reporting on indicators that are not used. The direct costs associated with the country-level results system are hard to track, as described in box B.17 of appendix B. World Bank and IFC country staff, portfolio analysts, and monitoring and evaluation officers have described many transaction and opportunity costs associated with the time and effort they spend on developing CPF results frameworks at the design stage and revising them later, and then updating indicators for the PLR and CLR through a laborious process that involves individual requests for updates on indicators to sector and industry staff. The difficulty of forecasting a country program's eventual portfolio at the CPF stage often means that the initial CPF indicators will not be a good representation of eventual outcomes, which adds to the teams' burden of continually revising indicators. Given the frequent development of country products, some World Bank and IFC staff dedicate a large portion of their work program to feeding results systems, especially in multicountry

CMUs with many products. IFC staff especially struggle to keep up, given the additional IFC country products over and above the Bank Group products. These costs of managing indicators tend to overshadow or discourage country teams' attempts at bringing more relevant and eclectic evidence to any assessment of their achievements and failures. These opportunity costs also crowd out the space afforded by PLRs and CLRs to collectively reflect and learn from experience. The time spent on creating these products and shepherding them through formal Board review processes is time that could be devoted to more useful evaluative or collective thinking tasks.

Completion and Learning Reviews

The CLRs provide a partial picture of the Bank Group's development outcomes at the country level. In CLRs, country teams tend to focus their efforts on the part of the methodology that assesses the deviation of reported results from the CPF's forecasted result targets and indicators. IEG validators often have to conduct additional research, consulting databases and external reports, to find additional and more relevant sources of information on progress toward objectives. Thus, the CLR also falls prey to the limitations of the results frameworks by overly focusing on quantifiable, time-bounded, and attributable indicators from individual lending projects that do not capture country outcomes or consider the contribution of IFC and MIGA or ASA products and other tools. For example, in China, many of the World Bank's most important contributions have come from capacity-building efforts and learning-by-doing facilitated by the World Bank. Government officials gained exposure and experience by applying international best practice in procurement, social and environmental risk management, and project management to their own projects. In Chad, some of the World Bank's most important contributions were from persuading and convening stakeholders to help Chad obtain heavily indebted poor countries' debt relief. Neither of these impressive country outcomes was well captured in those countries' CLRs.

CLRs do not establish whether the Bank Group's contribution to country outcomes is greater than the sum of individual project objectives. At the CPF design stage, country teams articulate how Bank Group interventions will together contribute to CPF objectives. This articulation is thorough

and typically included in both the CPF document’s narrative and the results framework annex. However, at the CLR stage, there is a tendency to engage in “bean counting,” or a mechanical accounting of project results achieved under each CPF objective. For the most part, CLRs do not assess the cumulative effects from multiple projects. This is partially explained by the CLR’s overreliance on results framework indicators as the main source of evidence. IEG found that results frameworks measure over half of CPF objectives, with only one or two indicators pulled directly from a single Bank Group operation, whereas most CPF objectives’ intervention logics rely on the combined effects of multiple activities. For instance, the Bank Group has used various instruments in 30 or more relevant interventions to help Serbia reform its state-owned enterprises. Obviously, these multiple and sustained efforts are likely to have a greater cumulative country impact than a simple sum of individual lending outcomes. Box 3.4 illustrates how country programs’ results frameworks reveal very few cumulative results from Serbia’s state-owned enterprises agenda and shows how difficult it is to assess the agenda’s overall progress toward achieving country outcomes.

Box 3.4. A Partial View of Bank Group Progress on Long-Term Reform Agendas: Examples of State-Owned Enterprise Reforms in Serbia

The most impactful contributions of the World Bank Group to country outcomes in the Western Balkans are not quantifiable—they have to do with strengthening the implementation capacity of the governments in key areas of their European Union accession process, including public administration reforms, state-owned enterprise (SOE) reforms, and railway and energy provision reforms. The Bank Group has been a change agent in SOEs and in building various ministries’ capacity to undertake complex and politically sensitive SOE reforms, including in railways and energy.

However, only a partial view of these meaningful contributions can be gleaned from country documents and their results frameworks. The fiscal years 2012–15 strategy tracked only the introduction of rules for the appointments of managers and the coverage of payments for workers made redundant, and the 2016 indicator on the number of companies still to be privatized (a measure of overall progress on privatization) was replaced by an indicator tracking the resolution of unproductive SOEs, as privatization turned out to be more difficult than anticipated.

(continued)

Box 3.4. A Partial View of Bank Group Progress on Long-Term Reform Agendas: Examples of State-Owned Enterprise Reforms in Serbia (*continued*)

The Completion and Learning Review considered the objective achieved because the targets were met; the parliament adopted a law and allocated enough resources in the budget to provide severance payment to redundant workers. The institution and capacity building, as well as the transfer of know-how from the Bank Group to the client through years of joint work on SOEs, is missing from these assessments.

Sources: World Bank 2015i, 2015j, 2019q.

Because of their timing and limited content, CLRs are little used within the institution. The CLR is intended to be a learning tool for country teams to collectively reflect on achievements, failures, and experiences (World Bank 2018c, 19). The CLR is carried out by the country team concomitantly with the new CPF and SCD and presented to the Board as a CPF appendix. This is too late for the CLR to meaningfully inform the CPF and means that CLRs are overshadowed by the CPF or SCD, both of which receive much more attention from vice presidents, the Board, and other stakeholders. As a result, very few people read the CLR and little learning emerges from it. Country staff working on CLRs expressed the sentiment that they were writing mostly for IEG validators. They viewed the CLR as a compliance exercise because of its emphasis on updating results frameworks and applying ratings, and they had a tendency to outsource it or assign it to junior staff. Country teams find the lessons from the CLR and its validation by IEG, the CLRR, to be generic. Box 3.5 provides examples of the types of lessons found in CLRs. As one country manager summed it up: “Real evaluation—meaning collecting proper evidence, reflecting on what we do, how we do it, and then distilling these lessons learned—is absolutely critical. The devil is in what the system pushes teams to do. In practice we spend too much time on meaningless things, such as revising targets so that they look ‘perfect,’ and on determining the rating. This type of bean-counting mentality is detrimental to learning and innovating.”

Box 3.5. Examples of Lessons from the Completion and Learning Review

- » Open dialogue with a broad spectrum of stakeholders increases the ownership of the program (Albania CLR for CPS FY11–14).
- » It is important to quickly respond to changes in the country context and to monitor the program promptly (Benin CLR for CPS FY13–18).
- » Proactive engagement with the donor community is important in leveraging our support (Arab Republic of Egypt CLR for CPS FY06–14).
- » The World Bank Group's contribution to the country's development agenda is not limited by the size of its portfolio (Honduras CLR for CPS FY11–15).
- » Shifting client demand has required a high degree of flexibility (Sri Lanka CLR for CPS FY13–16).

Sources: World Bank 2015a, 2015b, 2015d, 2016f, 2018b.

Note: CLR = Completion and Learning Review; CPS = Country Partnership Strategy; FY = fiscal year.

Staff almost invariably say that the CLR's accountability elements lack bite. The CLR and the CLRR are tools that are meant to uphold the Bank Group's accountability for achieving results. Although this evaluation finds that these tools enable accounting and transparent reporting, there is little evidence that the CLR and its validation actually support accountability, which requires better staff incentives and signals from management and the Board, as described in chapter 5. Interviews with country staff and management and a review of Regional Memorandums of Understanding, Scorecards, and managerial dashboards show that staff are incentivized to focus on commitments, disbursements, and the preparation of documents, such as the CPF, rather than on results management. The CLR and the CLRR are hardly discussed by the Board when a new CPF is approved. The aggregation of ratings from CLR and their validation into corporate-level scorecards provides some indication of trends, but this is not used to orient regional strategies. Regional briefings to the Board do not discuss outcomes or lessons from country-level results systems, including the CLR. Interviewed staff also highlighted that the frequent staff

rotation reduces country teams' accountability for outcome management. Staff also argue that promotions or appointments to desirable posts are not based on a track record of achieving development results.

Other Evaluative Exercises

There is a lot of evaluative thinking embedded in sector work, but this is rarely captured by country engagement cycle documents. IEG's desk review found 12 out of 29 CPFs with planned activities for assessing intervention outcomes, including through impact evaluations related to CPF objectives. However, these assessments covered only a very small portion of programmed activities because of a lack of impetus and high costs. Other evaluative activities mentioned in PLRs or CLRs include client and beneficiary satisfaction surveys, third-party monitoring and evaluation, census and household survey data reviews, and gender-sensitive analyses with gender-disaggregated data. In interviews, GP staff often pointed to sectoral ASAs as the primary conduit for informing future actions to progress on sector outcomes. However, these ASA products were largely forward-looking and typically did not derive lessons from experience of what worked or what did not. Moreover, the knowledge and evidence stemming from these ASA products and evaluative exercises are not well featured in CLRs and had little weight in the assessment of results. For its part, the CLR rarely synthesizes existing knowledge about what does and does not work and why, including through synthesizing project Implementation Completion and Results Report Reviews.

Some country teams carry out helpful country-level evaluation activities outside of the CLR, but these are rare and overshadowed by the results system's emphasis on indicators and targets. The Bank Group's cultural orientation toward generating reportable metrics crowds out space for mixed-method evaluations of the overall successes and failures of multiple, complex interventions. There are some promising examples of such evaluative exercises taking place, but case studies and the structured document review found these examples to be rare. In Haiti, the World Bank and other major donors carried out an assessment of the first two years of donor responses to the 2010 Port-au-Prince earthquake. The motivation for this assessment was

the vast sums donors spent on the response, the high profile of the interventions, and the recognition that there were significant shortcomings. In Vietnam, the country team is working on a retrospective of the 25 years of partnership between the World Bank and Vietnam that connects the World Bank's long-term engagement to high-level country development outcomes, such as power generation capacity/access, social and health insurance beneficiaries, educational attainment, and poverty. The report also outlines the World Bank's main projects and knowledge work contributing to key policy reforms and outcomes in those sectors.

The Bank Group's monitoring and evaluation staff are consumed by corporate reporting requirements, leaving little time to support country teams' outcome measurements. Country team leadership pointed out that the monitoring and evaluation functions that could play a meaningful role in assessing the Bank Group's contribution to country outcomes are instead focused on reporting corporate requirements and checking compliance with methodologies. For example, country teams believe that Regional development effectiveness units are helpful at the time of Regional Operations Committees (ROCs) and help them navigate corporate requirements. However, as one staff member put it: "If they had to do less 'corporate tick the box' and instead focused on helping us design and carry out meaningful monitoring and evaluation activities, wouldn't it be better?" Similarly, country teams mentioned IEG's important role assessing the Bank Group's country-level contributions through mixed-method evaluations but also feel that "IEG reinforces a broken system." They noted that CLRRs reinforced the practice of validating results frameworks and prioritizing attributable, quantifiable, verifiable, and time-bounded indicators. Country program evaluations were seen as useful but too infrequent and rating focused. In a nutshell, many country staff regretted that "IEG's expertise is spent on judging and rating and not on helping us assess and learn our true contribution to outcomes."

4 | Managing for Outcomes

Highlights

Country teams practice adaptive management by closely monitoring the health of country portfolios focused on disbursements, addressing delivery issues, and navigating changes in country contexts.

However, teams rarely use the country-level results system to inform adaptive management decisions. Instead, they rely on tacit knowledge, professional experiences, and advice from personal networks when making adaptive decisions. This reliance has some disadvantages.

Country strategies' midterm review instrument, the Performance and Learning Review, largely documents decisions that were already made and adjusts results frameworks; however, the extent to which teams use it for collective reflection and adaptive management depends on the impetus of country team leaders.

Clients do not engage much in the World Bank Group's country-level results system and the World Bank seldom helps countries develop their own monitoring, evaluation, and results management systems. Moreover, there is little coordination of results management activities among development partners.

This chapter reviews how country teams practice adaptive management and whether country-level systems support them in making course corrections. This chapter uses a typology to define different adaptive management practices for country portfolios to examine the types of adaptation that are most common in Bank Group country engagements (Buffardi et al. 2019). The chapter then looks at how country teams practice adaptive management. It reviews whether country teams use evidence from the country-level results system to inform their judgment, whether they use the PLR as a key adaptation moment, and how they engage clients throughout the engagement cycle on outcome measurements and adaptive management. In general, this chapter finds that country teams adaptively manage their portfolios to meet evolving client needs, but the country-level results measurement system does not effectively support teams in making course corrections to country programs.

Types of Adaptation

Managing country engagements includes a range of adaptive management decisions. In short, adaptive management is an iterative approach to decision-making, whereby interventions and portfolios are adjusted based on evidence and evolving context. The literature shows that most adaptive management approaches were developed for individual projects (Sweetman and Conboy 2018) and later mimicked with little adjustment at the country portfolio level. At the Bank Group, the types of adaptive portfolio decisions country teams can make include establishing and managing the operations pipeline; approving or not approving project extensions, restructuring, and additional financing; determining whether government reforms are sufficient to proceed with policy financing; deciding which pieces of analytics are needed to inform policy or support the future operations pipeline; and (re)allocating resources across the portfolio. To organize and understand these decisions, IEG adapted a typology developed by the Global Learning for Adaptive Management initiative (2019). As shown in table 4.1, the typology identifies six main categories of portfolio analyses: health checks, context responsiveness, sum of the parts, scale-ups, risk and outcome hedging, and comparative advantages. Not all require the same type of evidence.

Table 4.1. Portfolio Adaptation and Evidence Needs

Purpose of Analysis	Type of Adaptation	Type of Evidence
Health checks: What is happening? How are things going?	Attention of management and clients to specific projects and key actions to resolve problems during the cycle	Performance and delivery indicators, risk information
Context responsiveness: What do all activities need to consider or change in response to shifts in the context?	Shifts in implementation practices or change in composition of the portfolio during the cycle	Understanding of the external operating context
Sum of the parts: What are the combined effects of multiple projects?	Results reporting and organizational learning rather than adaptation	Outcome evidence
Scale up or down: What should be scaled up, scaled down, or discontinued?	Adaptation in implementation of current projects or future engagement, resource allocation shift, client dialogue	Outcome and value-for-money evidence
Hedging: How can the portfolio maintain a pipeline of outcomes over different time frames, and how likely is outcome to be achieved, given the risk?	Adjustment in composition of portfolio at the beginning, throughout, and across the cycle	Expected outcome trajectory, risks, strength of outcome evidence in other contexts
Comparative advantage: How can we maximize our unique contribution relative to others? How should future resources be allocated?	Inform development of future portfolio	Widest range of evidence: outcome evidence, stakeholder mapping, evidence of others' effectiveness

Source: Adapted from Buffardi et al. 2019.

Note: This typology was used as a basis of a content analysis of the Performance and Learning Review and the Completion and Learning Review and as an underlying framework for interviews and case studies on adaptive management practices.

Health Checks

Country teams closely monitor the health of country portfolios. Through various exercises, including the Country Portfolio Performance Review (CPPR), teams monitor the health of the country portfolio, particularly its safeguards, project risks, output delivery, and disbursement indicators. In all sampled PLRs, country teams documented their systematic efforts to identify portfolio delivery concerns early and remedy them with their clients. Port-

folio officers in IFC's regional industry departments conduct regular portfolio reviews that assess financial performance and how evolving macroeconomic situations impact the portfolio's health. MIGA's country ratings committee meets quarterly to review country risk ratings for its noncommercial risk coverages, and how these feed into MIGA's assessments of risks it faces in upcoming projects. In sampled CLRs, country teams reflect on the creative solutions they have applied to bolster delivery. For example, in Haiti the country team established a country implementation strategy drawing lessons from CPPRs. The strategy enshrined specific measures for project preparation, supervision, and performance management, including increasing face time with the client through country-based task team leaders, intensifying scrutiny on portfolio performance with monthly Haiti management team reviews. In Ghana, the country team set up an Implementation Support Team composed of Ministry of Finance and CMU staff to identify underperforming projects and help fiduciary and project teams tackle project issues as they arise rather than solely during supervision missions. The CMU informs the Implementation Support Team's actions by preparing a monthly portfolio analysis.

Country teams consider health checks an essential but insufficient part of managing country engagements. In interviews, staff emphasized that regular health checks of the portfolio are important for achieving results, especially in contexts where delivery is difficult. At the same time, country teams recognized that there is insufficient attention paid to results information and a disproportionate emphasis on disbursement and output delivery for the Bank Group and on financial return for IFC. As several interviewees put it, "If we don't disburse or deliver, we can't expect outcomes." Teams said the limited attention to results information was due to (i) country-level results systems generating inadequate information; (ii) difficulty engaging clients on the topic of results; and (iii) the incentive of country leaders to prioritize delivery over results management. These three factors are explored further in the second part of this chapter and in chapter 5.

Some teams have tried to consider results evidence in routine health checks through portfolio reviews at IFC or result-focused CPPRs at the World Bank. Vietnam is a good example of this type of practice. The country team opted for a "Programmatic CPPR" informed by a set of indicators to track progress toward outcomes, and they felt this practice improved the portfolio's perfor-

mance. It also enabled closer coordination with other partners, notably the Asian Development Bank and the Japan International Cooperation Agency, through joint reviews. Other country staff were more ambivalent about the utility of result-focused CPPRs, saying the practice is ad hoc. Staff also said it was difficult to translate the resulting data into visual, actionable, and understandable information. These CPPRs were deemed useful when the client demonstrates a keen interest in results information, which is rare, or the country managers use the results information as a tool for dialogue with clients. For its part, IFC recently introduced a new type of portfolio review that assesses the progress achieved on the reforms identified in its CSs (the “ifs”) as critical to enable IFC investments in a country’s specific sectors (the “then”). The first set of portfolio reviews gathered industry teams, country teams, and IFC’s senior management, including the chief operations officer and sometimes the chief executive officer. Interviewed IFC staff were positive about these reviews, saying they sent a strong signal from management to use IFC CSs as strategic road maps to understand IFC’s contributions to country development outcomes.

Context Responsiveness

Country teams adapt country engagements to external shocks and evolving government priorities. In interviews, country teams highlighted the importance of adjusting portfolios to evolving client demands while remaining focused on achieving the program’s overall objectives. Country team leaders said these adaptive management decisions require foresight and informed judgment but cannot be entirely data driven. Interviewed clients highlighted that the World Bank’s responsiveness to changing country dynamics was a strength compared with the clients’ other development partners. In PLRs, country teams document how they adapt portfolios to changing country contexts. More than half of the sampled countries with an available PLR (14 out of 25) described adapting portfolios to changing government priorities and government staff turnover and to the presence or absence of key reform champions. Country teams often have to adjust their engagements to external shocks. More than half of sampled countries with an available PLR (13 out of 25) documented adapting to unexpected events such as epidemics, civil unrest, natural disasters, economic shocks, or sudden increases in

migration. These events often lead to short-lived intensification of donor collaboration, rapid impact assessments, and a series of short-term adaptive actions to simplify processes, deploy emergency funds, increase risk mitigation measures, and refocus interventions to achieve immediate impacts. For example, in Colombia, the World Bank's country team conducted evaluative exercises to assess the immediate impact of the Venezuelan migration crisis before engaging in program course corrections.

In FCV countries, country teams constantly adapt to changing circumstances in situations of great uncertainty and low institutional capacity. In these countries, achieving development impact can mean maintaining or expanding development gains or preventing them from slipping backward. Four examples illustrate well the need to constantly adjust to shocks to avert development losses. In Haiti, the World Bank poured in additional resources after a catastrophic earthquake and used existing projects as platforms to address new challenges. When Chad's security situation deteriorated and government resources were diverted to security, the World Bank provided budget support and restructured an education project to sustain teacher and civil servant salaries and associated service delivery. The World Bank also responded to a crisis in neighboring Central African Republic by using an existing agriculture project to provide short-term assistance to refugees and returnees until a larger project could be approved. In Afghanistan, the World Bank has frequently adapted to changing security situations and evolving donor priorities and expectations. In the Solomon Islands, the World Bank used the turnover of senior government officials as an opportunity to deepen policy dialogues with the new government on macroeconomic issues, which helped renew opportunities for joint donor budget support.

Evidence-Informed Adaptation

Other types of adaptive management require evaluative evidence that the country-level results system does not provide. This evidence includes strong evaluative analysis and evidence on the outcomes of decisions. As shown in table 4.1, this evidence requires analyses of program expansions or contractions, the World Bank's comparative advantages, the cumulative effects of multiple interventions, and the risks versus rewards trade-offs to achieve outcomes. IEG found little evidence that the country results system provides

the necessary evidence to support and inform portfolio adaptations. Staff interviews, case studies, and document analyses suggest that decisions to increase Bank Group interventions are based more on the observed need and satisfaction of clients and whether the program's implementation and output delivery have been smooth than on evaluative evidence.

Decisions on balancing risks and rewards are also not well informed by the results system. Country teams pay close attention to evolving country risks across the portfolio. However, country teams reported that their risk analysis is largely separate from the information they gather on results and that the two should be combined to inform decisions on whether and when to support risky interventions with high potential impact. Country engagement documents rarely discuss how to balance risks against payoffs. In only three of 25 countries with available PLRs were there details on how the Bank Group “hedged” against different risks; for example, if engagement in sector X were to fail, how would the Bank Group adapt to achieve at least a minimum level of results within the strategy period?

How Teams Practice Adaptive Management

Country teams rarely use the country-level results system to inform adaptive management. Country teams find that information captured by the country-level results systems is not useful for decision-making for several reasons. First, they feel the information is not helpful in giving them a sense of the country's progress in a sector or outcome area or in grasping whether the Bank Group is hitting key milestones on its results chain. The information does not supplement the project-level evaluation system's blind spots on ASA, convening, or policy dialogue efforts, because these are also not well covered at the country level. Second, country teams observe that information is not up to date because there are only project monitoring systems to track the CPF's progress throughout the CPF cycle. Results information is updated only at the PLR and CLR stages. Third, country teams emphasize that results frameworks are not user friendly. In the absence of a portal that can track country-level indicators, country teams are unable to access, update, and visualize results information. IEG found that only a few country teams attempted to adapt formal CPF results frameworks into a results-gathering

tool for internal decision-making and client engagement. One team that did this was the Western Balkans CMU, which converted the CPF results framework's indicators into a traffic light tool, highlighting good, medium, or poor progress on CPF objectives. Results updates based on this tool were discussed internally on a quarterly basis and externally with clients at the time of the CPPR.

Instead, country teams rely primarily on tacit knowledge, professional experiences, and advice from personal networks when making adaptive decisions. Country teams' adaptive actions are shaped through professional judgment and tacit and informal knowledge rather than formal evidence from country results systems. Teams rely on advice from friends and colleagues, their personal reading of the country's political economy, and their understanding of government power dynamics to assess the likelihood of success of adaptive actions and navigate accordingly. Country teams use "soft information" that they can observe and learn from but cannot verify with metrics. As one country director agreed, "Data, however robust they are, are no substitute for informed judgment." When program objectives and intended results are multifaceted and not well captured through metrics—such as objectives related to systems change or institution building—or when the operating environment is highly uncertain, navigating by judgment can be an effective strategy for adaptive management (Aghion and Tirole 1997; Honig 2018, 2020; Muller 2018).

This reliance on tacit knowledge and personal judgment has some disadvantages. The *World Development Report* highlighted that development professionals, including Bank Group staff, can be susceptible to confirmation bias and fall into decision traps if their choices are influenced by their social environment, mental models, and groupthink (World Bank 2015k). The report shows that "development professionals can make consequential mistakes even when they are diligent, sincere, technically competent and experienced" (189). The IEG report on how the World Bank learns also highlighted several limits to overly relying on tacit knowledge exchange to drive action (World Bank 2015g). Both reports emphasized the importance of measures to counteract the risks of groupthink. They provide examples of "challenge functions," creating deliberative environments where teams are exposed to opposing views and invited to defend their own and where an independent

voice that specializes in challenging assumptions has been shown to effectively counter ingrained biases.

Country teams sometimes use evidence from sources outside of results reporting systems to inform adaptive management. To inform program adaptation, country teams sometimes carry out internal stocktaking exercises to maximize candor and self-reflection. Country teams also find rapid feedback gathering useful when shaping action. Several mentioned the Survey of Well-Being via Instant and Frequent Tracking and Iterative Beneficiary Monitoring as particularly useful. “Listening to Tajikistan” is an example of a frequently recurring survey that allows the country team to monitor Tajikistan’s shocks, fragility, and the life satisfaction of Tajiks during the country’s engagement cycle. The Morocco “portfolio footprint analysis” is another example of targeted information gathering to inform outcome orientation. The Morocco team was concerned that projects focused too much on geographic areas where implementation would be easy rather than areas where needs are high. In collaboration with project task team leaders, the country team overlaid poverty data with the geographical scope of the projects to gauge the poverty-reducing potential of the World Bank’s operations. This exercise equipped the team with visual tools to improve project targeting and led to a poverty targeting index that ranks the CPF priority in geographical areas at the provincial and communal levels. Country teams consider this type of feedback much more useful than information provided by the formal country-level results system.

Frequent turnover of staff and managers makes it difficult to retain knowledge over time and weakens tacit knowledge and accountability. Interviewees suggested that turnover hinders outcome management. Country teams rely on staff’s tacit knowledge about country context and their personal relationships with stakeholders, access to networks, and experience in the field to make informed decisions. As such, the frequent turnover in key staff and managerial positions without a proper handover or succession plan can be detrimental and complicate the Bank Group’s ability to focus on long-term results. The combination of a truncated institutional memory and the lack of results systems that focus on long-term success makes repeating mistakes more likely. Interviewees noted that there are few formal mechanisms to transfer knowledge during staff handovers, and the quality of this transfer depends on the behavior of the

incoming and outgoing staff. Country teams also argued that staff rotation undermines accountability and responsibility. Managers and staff who are working on country issues during the CPF design stage have sometimes moved on by the time of the PLR reality check and certainly by the time results are achieved or evaluated. Several country managers acknowledged this as a weakness of the system, or as one put it: “I have no incentive to invest in a results approach that is useful for managerial or reporting purposes, as I will have moved to another country by the time we have to make adjustments or have to show what we’ve achieved.” Among IEG’s sample of cases, there were only a few instances where the same country managers who designed the CPF stayed on until the PLR stage or beyond. In these cases, there was stronger ownership over the strategy and commitment to achieve results.

Some country teams have improved their adaptive management capacity by finding ways to build and retain institutional knowledge. Interviews and case studies revealed that country teams must find ways to overcome the frequent turnover of staff and management. Methods to pass on knowledge and hand over responsibilities from departing to arriving staff are considered very important by country teams and counterparts alike. However, these methods are too often neglected because there are no formal handover mechanisms and, as a result, handovers occur based on staff’s individual discretion and motivation. To mitigate this challenge, some teams have used past country managers or directors as peer reviewers for current CPFs or PLRs to ensure some continuity in thinking. Another way to overcome this challenge is to empower local staff to take an active role in portfolio and outcome management. Local staff can be powerful conduits of institutional and portfolio memory, can maintain strong networks across a range of stakeholders, and can carry a long-term view of the program’s outcome orientation. In the Western Balkans, for instance, the cross-country portfolio analysis is implemented primarily by local staff who play an increasingly active role in the countries’ engagement cycles and contribute significantly to the results management of country portfolios. Yet another way to maintain institutional knowledge is by establishing regular safe spaces to discuss the successes and failures of the portfolio. In Peru, for instance, the country director frequently convenes the full country team and GP members to brainstorm lessons from past actions and experiences.

Performance and Learning Reviews

Country teams use PLRs to document changes in the country engagement. The *Guidance on Country Engagement* expects that “the WBG [World Bank Group] carries out a continuous process of monitoring and learning from implementation” (World Bank Group 2019u, 16). The PLR is envisaged as a key adaptation moment with feedback loops between evidence and action. In interviews, the majority (83 percent) of staff who expressed an opinion on PLRs found them useful for revising the program’s ambition. Many (40 percent) found them useful for adjusting the CPF results frameworks. IEG’s document review found that they use it well for that purpose. Out of the 60 CPF objectives tracked by IEG, there were 27 changes to CPF objectives made at the PLR stage. Most of these changes (17) were to the CPF objectives scope or expected outcome levels. Another nine of these changes were caused by project delays or failures and changes in the country’s situation. MIGA and IFC also use PLRs to update targets and indicators and review the pipeline.

Much of the time and effort put into PLRs is taken up by revising the results frameworks, and 81 percent of indicators are revised or dropped. It makes sense that program indicators from the CPF results framework are updated or dropped at the PLR stage, given the difficulty of forecasting these results at the CPF stage. IEG reviewed 334 indicators from all sampled CPFs that had a PLR. After the PLR stage, only 19 percent of these CPF indicators were maintained and tracked as originally formulated. Most of the CPF indicators were either dropped (38 percent) or revised (43 percent). Thirty percent of PLR adjustments address shortcomings in the initial program design or align results frameworks to portfolio changes caused by new activities, project cancellations, and changes to the scope of work, among other reasons. Country teams considered these adjustments to be tedious and mechanical, crowding out the space to use PLRs for learning and reflection. Having the PLR as a Board deliverable adds to the transaction costs of the exercise, with a high level of scrutiny and multiple drafts going back and forth. Country teams also emphasized that the process was time-consuming, especially without online portals that link CPF results frameworks to underlying indicator data sets. As a result, country teams described having to “run after project TTLs [task team leaders] and IOs [Investment Officers]” for indica-

tors. Overall, teams did not see the value of having a fully fledged results framework at the CPF design stage when there is uncertainty surrounding the pipeline. Instead, they felt that the PLR stage was a more opportune time to elaborate the country program's results framework.

Country management significantly influences how much PLRs are used as a platform for collective reflection and adaptive management. For the most part, PLRs are used to report past decisions to the Board, not as an inflection point toward future change. Country teams, especially those from IFC, view the PLR primarily as a reporting requirement, not an adaptive management platform. PLRs contain little information on the adaptive management actions that IFC and MIGA take to maintain or improve the health of their portfolios, with fewer than half containing such information. Unless country management convenes GP and IFC industry staff for PLR collective learning, those groups are rarely involved, other than to provide updates to CPF results frameworks, and GPs generally found little value from the exercise. That said, IEG found examples of country leaders who proactively used PLRs to collectively reflect within the CMU team and with GPs and clients to uncover meaningful lessons. In those cases, country teams appreciated this moment for “taking a step back” as several put it. For instance, in the Maldives, the PLR was used to reorient the portfolio's focus from developing the capital city to developing secondary centers. However, this change in the country team's narrative and tactics was serendipitous rather than planned because the PLR process coincided with elections and a new government that had different priorities from the previous one.

Client Engagement

Clients engage country teams primarily when designing CPFs and carrying out portfolio health checks rather than when measuring and managing results. Clients that are most engaged with country teams at the CPF stage and when projects are being designed tend to have less interest in other aspects of the country engagement cycle, such as the PLR and the CLR. Country teams said that frequent government turnover incentivized officials to focus on short-term issues and “quick wins,” which made it difficult to engage them on medium-term outcomes and longer-term results measurement. Out of 25 CPFs, 18 rated political and governance risks as high in the Systematic

Operations Risk-Rating Tool framework. Peru and Sri Lanka are examples of high-capacity, high-turnover governments, where a CPF strategy can quickly become obsolete. As a response, Peru's country team has relied on signing Memorandums of Understanding with each incoming government because they serve as more realistic road maps for action than the CPF. By contrast, country teams more easily engage clients on results management issues in certain scenarios, including when the government has a strong "planning" track record (for example, in Colombia, Mexico, or Vietnam) or when the Bank Group contributes to a high-profile government priority, for example, Honduras's Dry Corridor Alliance.

The Bank Group plays a key role in strengthening clients' statistical systems but is little involved in developing their monitoring, evaluation, and results management systems. The Bank Group is among the world's largest providers of development cooperation to build the capacity of governments' National Statistical Offices and country statistical systems (World Bank 2017a, 21). This type of support is mainstreamed across country engagements. All sampled CPFs, except one, clearly articulated the Bank Group's support, through projects and ASA, for statistical capacity. However, only in rare cases, such as Serbia, does this support extend to strengthening national evaluation or results management systems (World Bank 2018c, 2019d). In interviews, Bank Group staff said the Bank Group does not rely on country-owned results frameworks to supply outcome evidence but instead relies on World Bank-funded project-level results frameworks, which it considers higher capacity. IEG could not identify any CPF that clearly uses country systems for monitoring and evaluations. Interviewed clients highlight that most development partners similarly rely on their own systems, which leads to a highly fragmented monitoring and evaluation landscape. This is a well-known and well-documented issue in the Sustainable Development Goal era when country-owned results frameworks and underlying monitoring and evaluation systems are supposed to be a core mechanism for measuring Sustainable Development Goals, yet are underperforming and overlooked by the donor community (OECD 2019, OECD and UNDP 2019; Vähämäki, Schmidt, and Molander 2011).

There is little coordination of results management activities among development partners. All 29 reviewed CPFs mention, often in general terms, some form of collaboration with donors to achieve CPF objectives. However, only

one CPF mentions monitoring and evaluation coordination. Country teams and clients report that development partners' results frameworks and indicators are driven by their own internal policies, practices, and approval cycles. Staff noted that the Bank Group and its development partners rarely aligned their operational systems even when there was an obvious upside, such as in procurement, so they saw a low likelihood they would align monitoring and evaluation frameworks. This lack of alignment across development partners makes it more difficult for governments to aggregate and assess donor contributions. Country teams see advantages when coordination takes place, often at times of crisis or in relation to budget support, but in their view, alignment's transaction costs rarely justify the payoff. In Haiti, for example, joint learning exercises and portfolio reviews were undertaken with the Inter-American Development Bank when both donors increased their programs after a catastrophic earthquake. In the Pacific Islands, major donors—including the World Bank, EU, Asian Development Bank, Australian Department of Foreign Affairs and Trade, and New Zealand Ministry of Foreign Affairs—have jointly provided budget support and policy reforms with a shared policy matrix and indicators. This had advantages in focusing efforts and achieving far-reaching policy changes.

5 | Toward Stronger Outcome Orientation

Highlights

- ||| The current country-level results system relies on specific principles of accountability that do not fit well with the way the World Bank Group pursues outcomes at the country level.
- ||| A renewed country-level results system could conceive accountability differently, based on evidence of achievement and failures and descriptions of learning and adaptation, rather than reaching predefined indicators.
- ||| The current country-level results system is too reliant on results frameworks and could benefit from introducing new approaches such as a monitoring, evaluation, and learning plans that are better suited to country programs. Under these plans, approaches would change in the following ways:
 - » Monitoring approaches would focus on selectively tracking critical country outcomes to assess progress on objectives and to support adaptive management.
 - » Evaluation approaches would focus on the World Bank Group's contribution to country outcomes, scale back ratings, and emphasize selectivity over coverage.
 - » Learning approaches would focus on creating safe spaces for teams to engage and discuss evaluative evidence.
- ||| Changes to the country-level results system would need to be supplemented by changes in signals and incentives to center on results achievement and management rather than overemphasizing approvals, commitments, and output delivery. Incentives could better match staff's intrinsic motivations.

This chapter rethinks the country-level results system’s accountability principles, tools, and incentives. To improve the accuracy, utility, and outcome orientation of the country-level results system, tools, principles, and incentives should better capture the Bank Group’s contribution to outcomes. The evaluation lays out the contours of a renewed country-level results system based on the findings and good practices presented above; a detailed review of literature on good practices in other development and government agencies; and an exercise to brainstorm solutions with a cross-section of Bank Group staff.

The challenges presented in the report are well known, but past corrections have ossified rather than improved these shortcomings, so a new vision is needed. Over time, the Bank Group, including IEG, has correctly identified some of the shortcomings of the current results system, including the over-emphasis on lending, short-termism, and output-focus, and the mismatch between objectives and indicators in the CPF results frameworks (World Bank 2016f, 2017a, 2017d). However, past recommendations and corrections to the system have been to double down on attribution and add more performance measures, creating a cascade of indicators that have become ever less useful without solving the underlying problems of the results system. This evaluation proposes to profoundly rethink the notion of accountability that underlies the current system; the tools for monitoring, evaluating, and learning from country engagements; and the incentives for staff to learn from experience and to prioritize development results. In changing the system there are inevitably trade-offs to consider. Some of the main trade-offs are laid out in box 5.1 at the end of the chapter.

Rethinking Accountability

The current country-level results system relies on specific principles of accountability that do not fit well the way the Bank Group pursues outcomes at the country level. The current country-level results system is premised on a narrow view of accountability that equates it with counting results that can be translated into metrics and ratings. The accountability system is hierarchical, based on a notion that accountability comes from senior management and the Board observing whether targets have been met and on exerting pos-

itive or negative pressure on subordinates based on results. Yet this model is not functioning effectively at the country level, because staff do not find that the country-level results system substantially drives their incentives. The country-level results system requires rating effectiveness against forecasted indicators. It measures results that can be quantified and delivered over a short time frame and that the Bank Group has direct control over achieving. These features are not suitable at the country level. At the country level, the Bank Group pursues results through indirect pathways (such as institution building, demonstration effects, market creation, and others) and multiple instruments (such partner convening, capacity building, policy dialogue, and analytical work) where its influence could be evaluated but rarely quantified or translated into indicators with targets and baselines. Moreover, these indirect pathways to outcomes are not fully within the Bank Group's control and so are not captured by the results system, which only considers results that are directly attributable to Bank Group interventions. Consequently, the results system does not answer key questions of how well the Bank Group's interventions, including ASA and IFC upstream efforts, build institutions, create markets, or shape policies and under what conditions these institutions contribute to sustained positive changes to societies' well-being. This evaluation has shown that this stringent notion of accountability constrains country teams' outcome measurement and management practices, thereby limiting the Bank Group's knowledge of its influence through indirect pathways over country outcomes.

A renewed country-level results system could conceive accountability differently. To enable a more strategic and collective focus on outcomes at the country level, a different model of accountability could be envisioned—one that is based on the idea of mutual accountability, collective learning, informed risk taking, and maintaining trust through both rewards and effective challenge mechanisms. This would start by (i) acknowledging that the Bank Group can influence country outcomes but is not in total control over them; (ii) recognizing that country teams can make informed decisions to adapt country engagements to evolving risks and to complex environments during implementation rather than deciding all targets and objectives during the design; and (iii) realizing that capturing contributions to country outcomes and assessing the cumulative effects from multiple interventions requires

dedicated evaluation inquiries, not just measurement of indicators. As such, country teams should be held accountable for providing well-evidenced descriptions of achievements and failures and for learning from them and adapting accordingly, not for reaching predefined indicators. Senior management or Board oversight (including independent evaluation) could make their own informed judgments on whether contributions have been substantive, given the context and available quantitative and qualitative evidence, and they could ensure that these judgments meaningfully incentivize behavior. The country-level results system should thus relax its current focus on metrics, attribution, and time-boundedness and adopt a more differentiated approach to capturing results—based on plausible contributions, an adequate evidence base, time-appropriateness, selectivity, and contestability mechanisms.

Rethinking the Tool Kit

The current country-level results system is too reliant on CPF results frameworks. The Bank Group uses results frameworks to monitor country programs in much the same way as it monitors individual projects, without much adaptation. However, results frameworks are best suited for measuring outcome areas where indicators are a reasonable proxy for success and collected routinely as part of sector work, where theories of change are well established and change is rather linear, and where there is relative certainty about interventions' efficacy within a specified context. Conversely, results frameworks are less effective when the measurability of outcomes is contested; there is little data; or there are multiple and diverse interventions contributing to a broader range of objectives, which is typical of country programs (Andrews 2013; Brinkerhoff and Brinkerhoff 2015; Mansbridge 2014; Muller 2018; Radin 2006). Also, country teams make little use of the results framework to engage clients or guide their adaptive management decisions.

A renewed country-level results system could introduce new monitoring, evaluation, and learning (MEL) approaches. Instead of focusing on results frameworks and their rating tools, the country-level results system could develop MEL plans that would enable country teams to use a wider set of distinct MEL approaches that best fit the type of outcome pursued and the type of results information needed to support learning and decision-making.

Several other development partners, including the United States Agency for International Development, have adopted the practice of MEL plans at the country level. This could also improve the evidence on the Bank Group's contribution to country outcomes and encourage evidence-informed adaptive management practices. Distinguishing between MEL approaches is necessary because one tool cannot meet all purposes. Table 5.1 shows how a MEL plan could address some of the limitations of the current country-level results systems. A more elaborate table is presented in appendix D, which provides more details on what a prototype MEL plan could entail.

Monitoring approaches would focus on selectively tracking critical country outcomes to assess progress on CPF objectives and to support adaptive management. A monitoring system could track key country outcomes to inform strategic perspectives on whether progress on CPF objectives is moving in the right direction but without expectations that these outcomes could be attributed to the Bank Group or that specific targets would be achieved. It would provide information that could be tracked and updated frequently and visualized and analyzed easily to provide useful and timely feedback for adaptation. The adaptive management literature highlights the importance of management information systems that work for the users. Such systems are simple to use, access a variety of relevant information sources, visualize progress toward objectives, and can be operated within existing project data platforms. During brainstorming sessions, many country teams highlighted two unmet needs that could inform their decision-making: (i) a portal that links project-level information to the CPF results framework for easier, more dynamic, and more accurate updates; and (ii) long-term tracking of key country outcomes that the Bank Group hopes to influence.

Evaluation approaches would focus on the Bank Group's contribution to country outcomes, scale back ratings, and emphasize selectivity over coverage. Each of these revisions to the current evaluation approach is described in table 5.1.

Table 5.1. How a Monitoring, Evaluation, and Learning Plan Could Address Current Limitations

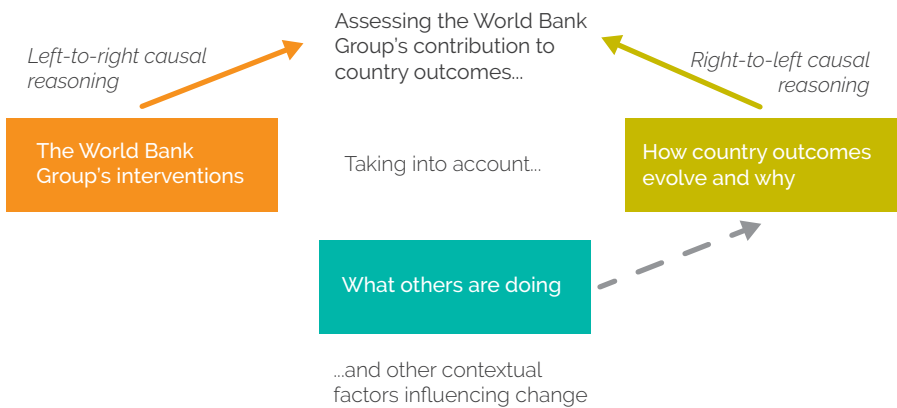
Intended Use	Current System	MEL Plan
Capturing contribution to country outcomes	Results frameworks' indicators do not capture country outcomes well because of expectations of quantification and attribution	The MEL plan and its terminal evaluation would report on the contribution of the World Bank Group to key CPF objectives and their influence on country outcomes
Capturing the contribution of multiple instruments and indirect development pathways	Results frameworks do not properly capture the contribution of multiple instruments, IFC and Multilateral Investment Guarantee Agency, complementarities, or indirect pathways	A MEL plan that embraced selectivity could allow deeper inquiry into critical areas of interest including these issues
Capturing contributions over time	Results frameworks primarily capture past operations, and do not capture medium- to long-term effects	Monitoring would take place across Country Partnership Framework cycles, and evaluation exercises would be timed to be able to capture long-term effects
Reporting on corporate priorities	Reporting on corporate priorities at the country level is carried out separately from the country results system.	The monitoring plan report on corporate priorities and International Development Association results frameworks
Adaptive management	The country-level results system does not effectively support adaptive management	Country teams would shape the MEL plan in part so that they fit their own adaptive management and learning needs
Knowledge of World Bank Group contribution to high-level outcomes	The country-level results system prioritizes accountability rather than knowledge generation	Increased knowledge would come from the combination of monitoring outcomes over the long term and from specific evaluations on critical areas
Organizational learning	Accountability and reporting requirements crowd out learning and reflection	Relaxing reporting requirements would reduce compliance burden, and MEL plans would place greater emphasis on evaluation that meets the needs of country teams and supports their learning
Incentives	Teams engage with the system out of compliance, not from a sense of utility or interest	Decentralization and emphasis on learning would improve ownership of and interest in the system by country teams
Governance and accountability	Staff report the system provides little effective accountability or impact on incentives	Conceive accountability differently, based on evidence of achievement and failures and descriptions of learning and adaptation

Source: Independent Evaluation Group.

Note: CPF = Country Partnership Framework; IFC = International Finance Corporation; MEL = monitoring, evaluation, and learning.

At the country level, the evaluation approach could focus on understanding the Bank Group’s contribution to country outcomes. There is an inherent conflict between expecting results to be attributable to the Bank Group and to achieving high-level outcomes that are not completely within the Bank Group’s control. It is, however, possible to evaluate the Bank Group’s contributions to country outcomes, including the magnitude, uniqueness, and catalytic nature of the Bank Group’s influence. Figure 5.1 shows that assessing the Bank Group’s contribution to country outcomes brings in right-to-left causal reasoning, rather than just the traditional left-to-right. This requires understanding all the factors that contribute to changes to country outcomes. Much of the Bank Group’s research and analytics already study changes in sectors and explain causes, though this available information is rarely used to discern the Bank Group’s contributions to these changes. Also, there are other evaluation methods and techniques that can assess the Bank Group’s contribution to outcomes, including techniques to measure objectives that are hard to quantify (Vaessen, Lemire, and Befani 2020).

Figure 5.1. Alternative Causal Reasoning at the Country Level



Source: Independent Evaluation Group.

Note: "Left-to-right causal reasoning" asks the following question: "Given what is known about the World Bank Group's portfolio of intervention, how likely is it that the World Bank Group contributed to changes in country outcomes?" Answering this question requires understanding and verifying theories of change, evidence of what works—for whom and under what circumstances—and information on what the Bank Group actually delivered. Conversely, "right-to-left" causal reasoning asks, "Given what is known about how country outcomes have changed over time and the factors that contributed to these changes, how likely is it that the Bank Group contributed to these changes?" Answering this question requires data on changes in outcome and research or analytics that explain what effected these changes.

The evaluation approach could scale back its use of ratings. Currently, the CLR and CLRR focus on arriving at and justifying a single development outcome rating for the entire country engagement that is based on deviations from the anticipated targets and indicators. There are arguments in favor of ratings: they provide discipline to the assessment, are simple to communicate, are attractive to shareholders, allow aggregation and a certain level of comparability, and provide a forum for contestability. However, there are also clear shortcomings to the current ratings system. The current ratings approach has methodological limitations because it combines several performance metrics that are not all logically linked, it faces uneven data quality and availability, and it requires aggregating into a single rating a very diverse set of objectives and achievements. Country teams are largely rated on whether they achieved indicators or not, rather than on whether they substantially influenced country outcomes. Ratings can divert attention from the substance, create “gaming” behaviors, and detract from learning (World Bank 2016f).

The evaluation approach could be more selective to achieve greater depth and to match specific learning and accountability needs. A country-level evaluation approach should provide sufficient evidence for teams to strategically reflect on how the Bank Group works with governments and development partners to contribute to medium- and long-term changes in client countries. A more selective approach would yield more valid and useful evidence of the Bank Group’s outcome contributions and help the organization bridge key knowledge gaps. As part of this, country teams could propose the evaluations that should be carried out during or after the engagement cycle, depending on the team’s learning and accountability needs. For example, evaluations might be warranted to inform key decision-making moments (like whether to continue or adjust an approach in a key sector or modality), or to assess the cumulative effectiveness of a substantial long-term engagement. Evaluations could be programmed to assess the effectiveness of critical indirect development pathways, which are difficult to capture with simple quantified indicators. Review and approval of the evaluation plan and selection of evaluations by Regions or the Board could ensure that strategically important interventions are covered by evaluation.

Learning approaches would focus on creating safe spaces for teams to engage and discuss the evaluative evidence. The PLR, CLR, and CLRR all include “learning” as a core objective. For the CLR, teams are asked to draw out lessons from the previous CPF’s implementation, which IEG also does when validating CLRs. Lessons captured in CLRs and CLRRs span a wide range of topics, but country teams often find them too generic and not substantiated enough to be useful. More important, teams question the effectiveness of using reports to share lessons when these reports are primarily geared toward external reporting to IEG and the Board and are rarely used by country teams themselves. Instead, country teams find that what facilitates learning is the process of coming together as country teams, sometimes with clients, to discuss the evaluative evidence and to reflect on the evidence’s implications for current and future engagements. The external literature on adaptive management shows that routine feedback meetings, used as safe spaces for teams to genuinely engage results data and evidence, are a cornerstone of organizational learning. They are also an effective mechanism to collectively reflect on the risks and rewards of alternative courses of action in changing environments.

Rethinking Incentives

Changing the tool kit will not lead to changes in practice and behavior if the underlying incentives are not aligned with the pursuit of outcomes. Country teams’ practices and behaviors are shaped by their intrinsic motivation to achieve impact in client countries and by incentives that can either support or hinder this intrinsic motivation. Incentives emerge from what is measured, tracked, and rated in managerial dashboards, regional vice presidents’ and country directors’ performance agreements, and Corporate Scorecards. Incentives also come from informal signals that staff receive from management, the Board, and other members of the authorizing environment, including IEG. Signals are less tangible than incentives and include what managers informally value or praise, what managers spend their time on, and what leadership prioritizes in meetings—including ROCs, Board discussions, and unit meetings. These signals and incentives are not always aligned with outcome orientation. IEG’s evaluations have shown that results systems will not uphold accountability if they are not embedded in an organizational set-

ting where teams are encouraged, through incentives and signals, to use the system to learn, course correct, and achieve results. A results system without such incentives and signals is performative and fuels a façade of accountability (World Bank 2013b, 2015h, 2016g, 2017a).

Current signals and incentives center on approvals, commitment, and output delivery, not results achievement and management. Through interviews and country cases, IEG gathered evidence on the types of incentives and signals that drive staff behavior. Interviewees felt that incentives of various types—carrots, sticks, and sermons—largely centered on the importance of receiving approvals and increasing the volume, disbursements, output delivery, and financial performance in the case of IFC. As one interviewee put it, “What really matters is getting new deals approved by the Board and making sure that those disbursement ratios in CD [country director] and Vice President’s dashboards don’t go into the red.” Regional agreements contain indicators on the number of CPFs submitted to the Board but nothing on the quality of PLRs or CLRs. The Board reviews new projects and CPFs at approval but pays less attention at closure. Staff recognition and promotion come from getting a project approved by the Board, not from the results of that project. IFC’s incentives are for the most part based on deal volumes that drive individual and sectoral behavior.

Although approvals, commitment, and output delivery are important to the Bank Group’s success, an overemphasis on these aspects can overshadow the drive for results. Staff interviews show concerns that pursuing these incentives can divert attention from quality and results. For example, for certain objectives, output delivery—such as operating newly built hydropower plants or delivering vaccines—is a good proxy for results. However, in other projects that involve analytical work or capacity or system building, output delivery is neither necessary nor sufficient to achieve results. In one theoretical example, a team that seeks to change a public agency’s management practices could lead a very successful reform process without ever delivering a specified output—or could deliver that output but make little traction on reform. Teams’ performance is also tracked by loan disbursements. Most times, this is a good proxy for a successful implementation and encourages action to unlock operational bottlenecks, but not always. For example, in policy reform loans,

coordinating and building consensus among divided stakeholder groups or administrative layers can indicate progress toward objectives, even if the lending disbursement is slow. Disbursement pressures encourage teams to unlock delivery bottlenecks but may discourage teams from applying strategic changes or drastically rethinking approaches to achieve greater impact.

A renewed country-level results system could experiment with new incentive mechanisms for better outcome management. Although reshaping incentives in the Bank Group goes well beyond what can be achieved through a renewed results system, some incentives could be corrected through changes in practice. For example, Regions could experiment with changing country leadership's incentives, including adjusting agreements between the managing directors and regional vice presidents and between regional vice presidents and their country directors as well as other formal incentive mechanisms to de-emphasize meeting disbursement targets while emphasizing outcome management. The right incentive need not take the form of a metric indicator or ratings but could stem from an assessment from senior management or the Board on whether sufficient progress was achieved on priority outcomes, based on compelling evaluative evidence.

Incentives could better match staff's intrinsic motivations to achieve outcomes, their learning methods, and the signals that matter most to them. The current results system relies on ratings and their independent validation as a "stick" and "reward" instrument but also as the primary "challenge mechanism" to ensure the contestability and credibility of the reporting. However, this report has shed light on a range of problems with the rating practices. A renewed system could envision alternative "challenge mechanisms." For example, the practice of the ROC for the CPF or IFC high-profile meetings with the chief operations officer and chief executive officer for country strategies could be extended to the evaluation stage. The Bank Group could experiment with dedicated high-profile meetings to discuss evaluative evidence on results achieved at multiple levels of the organization, including the regional vice presidential unit and the Board. The Regional Updates to the Board could more prominently feature evidence of results and their discussion. A challenge mechanism could be built into these meetings to ensure contestability of the evidence and decision-making. This

might require the presence of designated “provocateurs” and external voices with independent judgment who are assigned to ask questions that challenge tenuous assumptions and evidence.

Mechanisms to mitigate the short-termism created by rapid staff turnover and short country cycles could be envisioned. Country teams could improve staff succession plans and handover strategies by ensuring the continuity of outcome management. For example, exiting country leaders and teams could prepare knowledge packages that contain qualitative information on drivers of success and failure in the specific country political economy, to be passed on to the incoming country teams. Ensuring that results systems serve as a platform to capture evaluative evidence, including qualitative information, detailed lessons, and feedback on what works, what does not, and why in this context could mitigate the knowledge losses caused by staff turnover. Creating “results teams” related to key country outcomes—which exist in some country offices—gives GP staff a leadership role on key country outcomes. These teams could also better recognize and leverage the unique role played by local staff, who provide critical continuity to the Bank Group’s engagement across CPF cycles and have an acute sense of the unique development challenges, institutional capacities, and political constraints of the clients and support them for years on end to achieve outcomes.

Box 5.1. Trade-Offs in Adopting a New Country-Level Results System

The evaluation recognizes that there are trade-offs in changing some of the features of the current country-level results system. Some of the most important trade-offs include the following:

- » **External reporting versus internal use.** The current results system privileges information that is geared toward external reporting to the authorizing environment, with ratings that can be aggregated, are relatively easy to communicate and digest to shareholders, and contain reporting mechanisms that favor transparency. There is, however, little evidence of actual use by the authorizing environment of the information generated by the system. A renewed system would rebalance toward internal use of results information, privileging internal accountability and

(continued)

Box 5.1. Trade-Offs in Adopting a New Country-Level Results System (continued)

learning without losing the capacity to provide critical results information to the Board, clients, and the general public.

- » **Standardization versus customization.** The current results system uses a common methodology and approach for all country engagements, independent of the type of objectives pursued and nature of the portfolio and interventions, leading to some critical mismatches between what outcomes are pursued and what results are captured, especially through indirect development pathways. A renewed system would offer more flexibility to country teams to select the monitoring, evaluation, and learning approaches as well as types of data that are most suitable to the country engagements and objectives pursued.
- » **Centralized versus decentralized.** The current results system is centrally designed and managed, and the ownership of the system primarily lies with Operations Policy and Country Services and the Independent Evaluation Group. A renewed system would empower Regions and Country Management Units to play a more active role in designing, governing, and using the system. This is in line with the Bank Group's ongoing reform agenda to realign the matrix organization around Regional vice presidency units, enhance the global footprint, and move decision-making closer to clients.
- » **Comprehensive versus selective.** The current results system seeks to cover all Country Partnership Framework objectives and provide a broad picture of the overall portfolio of activities; at the same time, it inevitably provides a superficial picture of outcomes and does not adequately cover nonlending and indirect development pathways. A renewed system would have a more selective approach that would privilege outcome focus, utility for decision-making, and adequacy of the evidence base. Governance mechanisms can be put in place to avoid the risk of cherry picking.

Source: Independent Evaluation Group.

Recommendation

Based on the vision described above, IEG makes the following concrete recommendation:

Reform the country-level results system to ensure that it accurately captures the Bank Group's contribution to country outcomes and usefully informs decision-making on country engagements. The Bank Group should keep a country engagement model that articulates clear outcome-level objectives and lays out the pathways that will be pursued to achieve them, conducts periodic reviews to take stock of progress, and includes an end-of-cycle review of evidence and learning. These reviews need not be a Board deliverable but should be carried out in time to meaningfully inform course correction and the next strategy. It should discontinue the reliance on results frameworks in its country strategies and midterm and terminal reviews. And it should adopt MEL plans for its country engagements (box 5.2 and appendix D).

Box 5.2. How Could Monitoring, Evaluation, and Learning Plans Work?

Appendix D presents additional ideas on what a prototype monitoring, evaluation, and learning (MEL) plan could entail.

- » A monitoring plan would describe how the country team will monitor the country engagement's performance, changes to the country outcomes that the World Bank Group seeks to influence, and contextual risks that are relevant to specific Country Partnership Framework (CPF) objectives.
- » An evaluation plan would define the team's criteria for selecting evaluation activities, the evaluation's purpose and expected use, possible evaluation questions, and the evaluation modality (internal or commissioned out). The evaluation plan would also propose a terminal evaluation that would synthesize available information on the Bank Group's performance, including the results from evaluative activities under the MEL plan, portfolio information from project results systems, and stakeholder feedback surveys or narrative accounts.
- » A learning plan would identify knowledge gaps that the country team intends to fill during the CPF period. It would also plan activities for country teams to

(continued)

Box 5.2. How Could Monitoring, Evaluation, and Learning Plans Work? (continued)

periodically reflect on certain activities, processes, and findings related to evaluations, information monitoring, and stakeholder engagements. The MEL plan should include one or more midterm reviews, where the country team engages in an exercise to collectively consider whether progress toward objectives is on the right track and to identify lessons or opportunities for changing direction.

- » Country teams would propose a MEL plan at the CPF design stage and allocate resources for its implementation. The plan could be proposed by the Country Director and approved by the Region's vice president. However, the plan should not be comprehensive or fixed on approval, and country teams should be expected to update the plan over the life of the country engagement as part of the portfolio review process.

Source: Independent Evaluation Group.

Bibliography

- Aghion, Philippe, and Jean Tirole. 1997. "Formal and Real Authority in Organizations." *Journal of Political Economy* 105 (1): 1–29.
- Andrews, Matt. 2013. *The Limits of Institutional Reforms in Development: Changing Rules for Realistic Solutions*. Cambridge: Cambridge University Press.
- Andrews, Matt. 2018. "Public Policy Failure: 'How Often?' and 'What is Failure, Anyway? A Study of World Bank Project Performance.'" CID Faculty Working Paper 344, Center for International Development, Harvard University. https://bsc.cid.harvard.edu/files/bsc/files/public_policy_failure_cidwp344.pdf.
- Andrews, Matt, Lant Pritchett, and Michael Woolcock. 2013. "Escaping Capability Traps through Problem Driven Iterative Adaptation (PDIA)." Working Paper 299, Center for Global Development, Washington, DC. https://www.cgdev.org/sites/default/files/1426292_file_Andrews_Pritchett_Woolcock_traps_FINAL_0.pdf.
- Aronsson, Inga-Lill, and Hur Hassnain. 2019. "Value-Based Evaluations for Transformative Change." In *Evaluation for Transformational Change: Opportunities and Challenges for the Sustainable Development Goals*, edited by R. D. van den Berg, C. Magro, and S. S. Mulder, 89–104. Exeter: International Development Evaluation Association (IDEAS).
- ADB (Asian Development Bank) 2012. *Managing for Development Results: Relevance, Responsiveness, and Results Orientation*. Mandaluyong City: Asian Development Bank.
- Australian National Audit Office. 2009. *Innovation in the Public Sector: Enabling Better Performance, Driving New Directions*. Canberra: Australian National Audit Office.
- Barber, Michael, Paul Kihn, and Andy Moffit. 2011. "Deliverology: From Idea to Implementation." *McKinsey on Government* 6: 32–9. https://www.mckinsey.com/~/media/McKinsey/Industries/Public%20and%20Social%20Sector/Our%20Insights/Deliverology%20From%20idea%20to%20implementation/Deliverology_From_idea_to_implementation.pdf.
- Behn, Robert D. 2003. "Why Measure Performance? Different Purposes Require Different Measures." *Public Administration Review* 63 (5): 586–606. <http://www>.

csus.edu/indiv/s/shulockn/executive%20fellows%2004-05/behn%20why%20measure%20performance.pdf.

Bemelmans-Videc, Marie-Louise, Ray C. Rist, and Evert Vedung. 1998. *Carrots, Sticks and Sermons: Policy Instruments and Their Evaluation*. New Brunswick: Transaction Publishers.

Bester, Angela. 2012. “Results-Based Management in the United Nations Development System: Progress and Challenges.” Report prepared for the United Nations Department of Economic and Social Affairs, for the Quadrennial Comprehensive Policy Review. https://www.unjiu.org/sites/www.unjiu.org/files/rbm_full_report_english.pdf.

Birdsall, Nancy, ed. 2006. *Rescuing the World Bank: A CGD Working Group Report and Selected Essays*. Washington, DC: Center for Global Development. <https://www.cgdev.org/sites/default/files/9781933286112-Birdsall-rescuing-world-bank.pdf>.

Birdsall, Nancy, and D. Kapur. 2005. *The Hardest Job in the World: Five Crucial Tasks for the New President of the World Bank*. Washington, DC: Center for Global Development Working Group.

Brinkerhoff, Derick W., and Jennifer M. Brinkerhoff. 2015. “Public Sector Management Reform in Developing Countries: Perspectives beyond NPM Orthodoxy.” *Public Administration and Development* 35: 222–37.

Britt, Heather, and Richard Hummelbrunner. 2014. “Synchronizing Monitoring with the Pace of Change in Complexity.” Discussion Note. Monitoring and Evaluation Series. (September). Washington, DC: USAID. https://usaidlearninglab.org/sites/default/files/resource/files/pace_of_change_in_complexity_2014-09-25_final.pdf.

Britt, Heather, and Melissa Patsalides. 2018. “Complexity-Aware Monitoring.” Discussion Note. Monitoring and Evaluation Series. (December). Washington, DC: USAID. https://usaidlearninglab.org/sites/default/files/resource/files/cleared_dn_complexity-aware_monitoring.pdf.

Buffardi, Anne L., Paige Mason, Claire Hutchings, and Samuel Sharp. 2019. “Monitoring and Learning for Country-Level Portfolio Decision-Making and Adaptation.” Briefing Note (May), Overseas Development Institute, London.

- Buntaine, Mark T., Bradley C. Parks, and Benjamin P. Buch. 2017. "Aiming at the Wrong Targets: The Domestic Consequences of International Efforts to Build Institutions." *International Studies Quarterly* 61 (2): 471–88. <https://academic.oup.com/isq/article-abstract/61/2/471/3866882?redirectedFrom=fulltext>.
- Busan Partnership for Effective Development Co-Operation. 2011. "Fourth High Level Forum on Aid Effectiveness." Busan.
- Casebourne, Jo. 2014. "Why Motivation Matters in Public Sector Innovation." London: Nesta. https://media.nesta.org.uk/documents/why_motivation_matters_in_public_sector_innovation.pdf.
- Cashin, Cheryl. 2011. *Major Developments in Results-Based Financing (RBF) in OECD Countries: Country Summaries and Mapping of RBF Programs*. Washington, DC: World Bank.
- Clay, Edward J., Matthew Geddes, Luisa Natali, and Dirk Willem te Velde. 2008. "Thematic Study, The Developmental Effectiveness of Untied Aid: Evaluation of the Implementation of the Paris Declaration and of the 2001 DAC Recommendation on Untying ODA to the LDCs, Phase I Report." London: Overseas Development Institute.
- Dahler-Larsen, Peter. 2011. *The Evaluation Society*. Redwood City: Stanford University Press.
- Evans, Alison. 2019. "Evaluation Measuring and Achieving Development Outcomes for the Executive Directors' Seminar." World Bank, Washington, DC.
- Gaarder, Marie Moland, and Ulrich Bartsch. 2014. "Who Cares About Development Outcomes? Market Failures and the Role of the Evaluation Function." *Journal of Development Effectiveness* 6 (4): 361–377. <https://doi.org/10.1080/19439342.2014.966742>.
- Gauri, Varun, Sheheryar Banuri, and Stefan Dercon. 2015. *The Biases of Development Professionals*. World Bank, Washington, DC.
- Global Partnership for Effective Development Cooperation. 2016. *Nairobi Outcome Document*. Nairobi, Kenya.
- Gunasekaran, Angappa, Charbel José Chiappetta Jabbour, and Ana Beatriz Lopes de Sousa Jabbour. 2014. "Managing Organizations for Sustainable Development

- in Emerging Countries: An Introduction.” *International Journal of Sustainable Development and World Ecology* 21 (3): 195–7.
- Højlund, Steven. 2014. “Evaluation Use in Evaluation Systems—The Case of the European Commission.” *Evaluation* 20 (4): 428–46.
- Honig, Dan. 2018. *Navigation by Judgment: Why and When Top-Down Management of Foreign Aid Doesn’t Work*. New York: Oxford University Press.
- Honig, Dan. 2019. “When Reporting Undermines Performance: The Costs of Politically Constrained Organizational Autonomy in Foreign Aid Implementation.” *International Organization* 73 (1): 171–201. DOI: <https://doi.org/10.1017/S002081831800036X>.
- Honig, Dan. 2020. “Actually Navigating by Judgment: Towards a New Paradigm of Donor Accountability Where the Current System Doesn’t Work.” Policy Paper 169, Center for Global Development, Washington, DC.
- Kenny, Charles. 2019. “Marginal, Not Transformational: Development Finance Institutions and the Sustainable Development Goals.” Policy Paper 156, Center for Global Development, Washington, DC.
- Leeuw, Frans L., and Jan-Eric Furubo. 2008. “Evaluation Systems: What Are They and Why Study Them?” *Evaluation* 14 (2): 157–69.
- Levine, Ruth, and William D. Savedoff. 2006. “The Evaluation Agenda.” In *Rescuing the World Bank: A Working Group Report and Selected Essays*, 183–93. Washington, DC: Center for Global Development.
- López, Roberto García, and Mauricio García Moreno. 2011. *Managing for Development Results: Progress and Challenges in Latin America and the Caribbean*. Washington, DC: Inter-American Development Bank.
- Mansbridge, Jane. 2014. “A Contingency Theory of Accountability.” In *The Oxford Handbook of Public Accountability*, edited by Mark Bovens, Robert E. Goodin, and Thomas Schillemans, 55–68. Oxford: Oxford University Press.
- Mayne, John. 2015. “Useful Theory of Change Models.” *Canadian Journal of Program Evaluation* 30 (2): 119–42.

- Michie, Susan, Maartje M. Van Stralen, and Robert West. 2011. "The Behaviour Change Wheel: A New Method for Characterising and Designing Behaviour Change Interventions." *Implementation Science*, 6 (1): 42.
- Muller, Jerry. 2018. *The Tyranny of Metrics*. Princeton: Princeton University Press.
- Natsios, Andrew. 2011. *The Clash of the Counter-Bureaucracy and Development*. Washington, DC: Center for Global Development.
- OECD (Organisation for Economic Co-operation and Development). n.d. *Emerging Good Practice in Managing for Development Results: Emerging Good Practice in Managing for Development Results*. First Issue. Paris: OECD. www.mfdr.org/Sourcebook.html/.
- OECD. 2016. *Evaluation Systems in Development Co-Operation: 2016 Review*. Paris: OECD.
- OECD. 2017. *Strengthening the Results Chain: Synthesis of Case Studies of Results-Based Management by Providers*. OECD Development Policy Papers 7, OECD, Paris. <https://doi.org/10.1787/544032a1-en>.
- OECD. 2019. "Managing for Sustainable Development Results: Concept and Preliminary Draft Guiding Principles." Discussion Paper for the OECD/DAC Results Community Workshop, Paris, 4–5 April.
- OECD and UNDP. 2019. *Making Development Co-operation More Effective: 2019 Progress Report*. Paris: OECD Publishing. <https://doi.org/10.1787/26f2638f-en>.
- OECD and World Bank. 2014. *Measuring and Managing Results in Development Co-operation: A Review of Challenges and Practices among DAC Members and Observers*. Paris: OECD.
- Örtengren, Kari. 2016. *A Guide to Results-Based Management (RBM), Efficient Project Planning with the Aid of the Logical Framework Approach (LFA). The LFA Method*. Stockholm: Swedish International Development Cooperation Agency. <https://www.sida.se/contentassets/9d257b83f4124113a324c61715150722/21920.pdf>.
- Ouye, Michael Michio, and Steven Toye Crocker. 2018. Dynamic Evaluation of Access Rights. US Patent 10,033,700, issued July 24, 2018.

- Patton, Michael Quinn. 2008. *Utilization-Focused Evaluation*. Thousand Oaks: Sage Publications. https://www.betterevaluation.org/en/plan/approach/utilization_focused_evaluation.
- Patton, Michael Quinn. 2010. *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. New York: Guilford Press. https://www.betterevaluation.org/en/plan/approach/developmental_evaluation.
- Patton, Michael Quinn. 2017. *Principles-Focused Evaluation: The GUIDE*. New York: Guilford Press.
- Perakis, Rita, and William D. Savedoff. 2014. "An Introduction to Cash on Delivery Aid for Funders." Note. Washington, DC: Center for Global Development. <https://www.oecd.org/dac/peer-reviews/Introduction-cash-on-delivery-aid-funders.pdf>.
- Porter, Stephen, and Penny Hawkins. 2019. "Achieving Sustainability through Sustainable Organizational Evaluation Systems." *New Directions for Evaluation* 162: 87–101.
- Radin, Beryl A. 2006. *Challenging the Performance Movement: Accountability, Complexity, and Democratic Values*. Washington, DC: Georgetown University Press.
- Ramalingam, Ben, Leni Wild, and Anne Buffarid. 2019. "Making Adaptive Rigour Work: Principles and Practices for Strengthening Monitoring, Evaluation and Learning for Adaptive Management." Briefing Note, Overseas Development Institute, London. <https://www.odi.org/sites/odi.org.uk/files/resource-documents/12653.pdf>.
- Rist, Ray C., and Nicoletta Stame, eds. 2006. *Managing Evaluative Systems*. Vol. 12 of *From Studies to Streams*. New York: Transaction Publishers.
- Strode, Mary, Ian MacAuslan, Christine Spanneut, Mathew Powell, C. Willoughby, Philippe Ngango Gafishi, and T. Thomsen. 2009. *Evaluation of the Implementation of the Paris Declaration: Thematic Study-Support to Statistical Capacity Building*. Synthesis Report. London: DFID.
- Sweetman, Roger, and Kieran Conboy. 2018. "Portfolios of Agile Projects: A Complex Adaptive Systems' Agent Perspective." *Project Management Journal* 49 (6): 18–38.

- Thomas, Vinod. 2012. *Managing for Development Results: Relevance, Responsiveness, and Results Orientation*. Manila: Asian Development Bank.
- UNSDG (United Nations Sustainable Development Group). 2011. *Results-Based Management Handbook*. New York: United Nations Development Group.
- UNDP (United Nations Development Programme). 2000. "Results Based Management: Concepts and Methodology." Technical Note 2000, UNDP, New York. <http://web.undp.org/evaluation/documents/RBMConceptsMethodgyjuly2002.pdf>.
- UNFPA (United Nations Population Fund). 2019. *Developmental Evaluation of Results-Based Management at UNFPA*. New York: UNFPA. <https://www.unfpa.org/admin-resource/developmental-evaluation-results-based-management>.
- UNICEF (United Nations Children's Fund). 2017. *Theory of Change Paper, UNICEF Strategic Plan, 2018–2021. Realizing the Rights of Every Child, Especially the Most Disadvantaged*. New York: UNICEF. https://www.unicef.org/about/execboard/files/2017-EB11-Theory_of_Change-2017.07.19-EN.pdf.
- UNICEF. 2018. *UNICEF Strategic Plan 2018–2021: Executive Summary*. New York: UNICEF. https://www.unicef.org/media/48126/file/UNICEF_Strategic_Plan_2018-2021-ENG.pdf.
- USAID (United States Agency for International Development). 2018. *Discussion Note: Adaptive Management*. Washington, DC: USAID.
- Vaessen, Jos, Sebastian Lemire, and Barbara Befani. 2020. *Evaluation of International Development Interventions: An Overview of Approaches and Methods*. Independent Evaluation Group. Washington, DC: World Bank.
- Vähämäki, Janet, Martin Schmidt, and Joakim Molander. 2011. *Results Based Management in Development Cooperation*. Stockholm: Riksbankens Jubileumsfond. <https://www.oecd.org/dac/peer-reviews/Vahamaki%202011%20RBM%20in%20dev%20cooperation.pdf>.
- Vähämäki, Janet, and Chantal Verger. 2019. "Learning from Results-Based Management Evaluations and Reviews." OECD Development Co-operation Working Paper 53, OECD, Paris. https://www.oecd-ilibrary.org/development/learning-from-results-based-management-evaluations-and-reviews_3fda0081-en.

- World Bank. 2012a. *Impact Evaluations: Relevance and Effectiveness*. Independent Evaluation Group. Washington, DC: World Bank.
- World Bank. 2012b. *Nicaragua—Country Partnership Strategy for FY13–17*. Washington, DC: World Bank.
- World Bank. 2013a. *Assessing the Monitoring and Evaluation Systems of IFC and MIGA: Biennial Report on Operations Evaluation*. Independent Evaluation Group. Washington, DC: World Bank.
- World Bank. 2013b. *Results and Performance of the World Bank Group 2012*. Independent Evaluation Group. Washington, DC: World Bank.
- World Bank. 2014. *World Bank Group Directive: Access to Information*. OPCS 5.01-DIR.01. Issued and effective July 1, 2014. Operations Policy and Country Services. Washington, DC: World Bank.
- World Bank. 2015a. “Albania—Country Partnership Strategy for FY11–14.” Completion and Learning Review, World Bank, Washington, DC.
- World Bank. 2015b. “Arab Republic of Egypt—Country Partnership Strategy for FY06–14.” Completion and Learning Review, World Bank, Washington, DC.
- World Bank. 2015c. *Botswana—Country Partnership Framework for FY16–20*. Washington, DC: World Bank.
- World Bank. 2015d. “Honduras—Country Partnership Framework for FY11–15.” Completion and Learning Review, World Bank, Washington, DC.
- World Bank. 2015e. *Honduras—Country Partnership Framework for FY16–20*. Washington, DC: World Bank.
- World Bank. 2015f. *Indonesia—Country Partnership Framework for FY16–20*. Washington, DC: World Bank.
- World Bank. 2015g. *Learning and Results in World Bank Operations: Toward a New Learning Strategy: Evaluation 2*. Independent Evaluation Group. Washington, DC: World Bank.
- World Bank. 2015h. *Results Frameworks in Country Strategies: Lessons from Evaluations*. Independent Evaluation Group. Washington, DC: World Bank.

- World Bank. 2015i. *Serbia—Country Partnership Framework for FY16–20*. Washington, DC: World Bank.
- World Bank. 2015j. *Serbia—Systematic Country Diagnostic*. Washington, DC: World Bank.
- World Bank. 2015k. *World Development Report 2015: Mind, Society, and Behavior*. Washington, DC: World Bank.
- World Bank. 2016a. *Behind the Mirror. A Report on the Self-Evaluation Systems of the World Bank Group*. Independent Evaluation Group. Washington, DC: World Bank.
- World Bank. 2016b. *Bulgaria—Country Partnership Framework for FY17–22*. Washington, DC: World Bank.
- World Bank. 2016c. *Program-for-Results: An Early-Stage Assessment of the Process and Effects of a New Lending Instrument*. Independent Evaluation Group. Washington, DC: World Bank.
- World Bank. 2016d. *Results of the World Bank RAS Program in Romania, 2012–2015: Changing Institutions to Achieve Development Goals*. Washington, DC: World Bank.
- World Bank. 2016e. *Sri Lanka—Country Partnership Framework for FY17–20*. Washington, DC: World Bank.
- World Bank. 2016f. “Sri Lanka—Country Partnership Strategy for FY13–16.” Completion and Learning Review, World Bank, Washington, DC.
- World Bank. 2016g. *World Bank Operations Evaluation Department: The First 30 Years*. Operations Evaluation Department. Washington, DC: World Bank.
- World Bank. 2017a. *Data for Development: An Evaluation of World Bank Support for Data and Statistical Capacity*. Independent Evaluation Group. Washington, DC: World Bank.
- World Bank. 2017b. *The International Finance Corporation’s Approach to Engaging Clients for Increased Development Impact*. Independent Evaluation Group. Washington, DC: World Bank.
- World Bank. 2017c. *Nicaragua—Paving the Way to Faster Growth and Inclusion*. Systematic Country Diagnostic. Washington, DC: World Bank.

- World Bank. 2017d. *Results and Performance of the World Bank Group 2016*. Independent Evaluation Group. Washington, DC: World Bank.
- World Bank. 2017e. *World Bank Group Joint Projects: A Review of Two Decades of Experience. Lessons and Implications from Evaluation*. Independent Evaluation Group. Washington, DC: World Bank.
- World Bank. 2018a. *Benin—Country Partnership Framework for FY19–FY23*. Washington, DC: World Bank.
- World Bank. 2018b. “Benin—Country Partnership Strategy for FY13–18.” Completion and Learning Review, World Bank, Washington, DC.
- World Bank. 2018c. *Enhancing Self-Evaluation in the World Bank Group toward Achieving Better Performance through Monitoring and Evaluation and Results Measurement: An IEG Retrospective Review*. Independent Evaluation Group, Washington, DC: World Bank.
- World Bank. 2018d. *Nicaragua—Country Partnership Framework for FY18–22*. World Bank: Washington, DC.
- World Bank. 2018e. “Nicaragua—Country Partnership Strategy FY13–17.” Completion and Learning Review, World Bank, Washington, DC.
- World Bank. 2018f. *Results and Performance of the World Bank Group 2018*. Independent Evaluation Group. Washington, DC: World Bank.
- World Bank. 2018g. “Sustainable Financing for Sustainable Development: World Bank Group Capital Package Proposal.” Report to Governors at 2018 Spring Meetings. April 2018. Washington, DC: World Bank.
- World Bank. 2018h. *World Bank Group Guidance: Country Engagement Bank Access to Information*. OPS5.01-GUID.106. Operations Policy and Country Services. Washington, DC: World Bank.
- World Bank. 2018i. “World Bank/IMF Spring Meetings 2018: Development Committee Communiqué.” Press Release No. 2018/150/ECR. Washington, DC: World Bank. <https://www.worldbank.org/en/news/press-release/2018/04/21/world-bankimf-spring-meetings-2018-development-committee-communique>.

- World Bank. 2019a. “Albania—Country Partnership Framework for FY15–19.” Completion and Learning Review, World Bank, Washington, DC.
- World Bank. 2019b. “Arab Republic of Egypt—Country Partnership Framework for FY15–19.” Completion and Learning Review, World Bank, Washington, DC.
- World Bank. 2019c. “Assurance Review of the Use of the Bank’s Corporate Scorecard.” GIA Report IBRD FY19–08. July 16, 2019. Washington, DC: World Bank.
- World Bank. 2019d. *Bank Directive: Program-for-Results Financing*. OPS5.04-DIR.106. Issued June 20. Operations Policy and Country Services. Washington, DC: World Bank.
- World Bank. 2019e. “Benin—Country Partnership Framework for FY13–18.” Completion and Learning Review, World Bank, Washington, DC.
- World Bank. 2019f. *CODE Retreat Summary: Prioritizing CODE’s FY19–21 Work Program*. Independent Evaluation Group. Washington, DC: World Bank.
- World Bank. 2019g. “Ethiopia—Co-Chairs’ Summary.” IDA19 Second Replenishment Meeting, Addis Ababa, Ethiopia. Kristalina Georgieva, Chief Executive Officer of the World Bank and Antoinette Sayeh, External IDA19 Replenishment Co-chair. June 18–20, 2019. Washington, DC: World Bank.
- World Bank. 2019h. *Including a Lens on Fragile, Conflict and Violent Situations in CPFs*. Washington, DC: World Bank.
- World Bank. 2019i. *The IDA 19 Results Measurement System*. Washington, DC: World Bank.
- World Bank. 2019j. *India—Assessing Results and How to Use Lessons Learned*. Washington, DC: World Bank.
- World Bank. 2019k. *The Key Role of Results Frameworks in CPFs*. Washington, DC: World Bank.
- World Bank. 2019l. “A Look into the New World Bank Group and World Bank Corporate Scorecards.” Corporate Secretariat. Technical Briefing, Wednesday, July 31, 2019. Washington, DC: World Bank.
- World Bank. 2019m. *Morocco—Country Partnership Framework for FY19–24*. Washington, DC: World Bank.
- World Bank. 2019n. *Preparing a CPF: Practical Advice to Make This a Successful Experience*. Washington, DC: World Bank.

- World Bank. 2019o. *Program-for-Results: Proposal to Remove the Cap on Commitment Authority*. Washington, DC: World Bank.
- World Bank. 2019p. *Results and Performance of the World Bank Group 2018*. Independent Evaluation Group. Washington, DC: World Bank.
- World Bank. 2019q. “Serbia—Country Partnership Framework for FY16–20.” Performance and Learning Review. World Bank, Washington, DC.
- World Bank. 2019r. *Theory behind Country Engagement Cycle: Country Engagement—Overview*. Washington, DC: World Bank.
- World Bank. 2019s. *World Bank Group Joint Projects: A Review of Two Decades of Experience Lessons and Implications from Evaluation*. Independent Evaluation Group. Washington, DC: World Bank.
- World Bank. 2019t. *World Bank Group Procedure: Country Engagement*. Washington, DC: World Bank.
- World Bank. 2019u. “World Bank Group Procedure: Country Engagement Bank Access to Information.” IBRD/IDA/IFC/MIGA Procedure. OPS5.01-PROC.105. Issued May 6, 2019. Washington, DC: World Bank.
- World Bank. 2020a. “Honduras—Country Partnership Framework for FY16–20.” Completion and Learning Review, World Bank, Washington, DC.
- World Bank. 2020b. *Results and Performance of the World Bank Group 2020*. Independent Evaluation Group. Washington, DC: World Bank.
- World Bank. 2020c. “Serbia—Country Partnership Framework for FY16–20.” Completion and Learning Review, World Bank, Washington, DC.
- World Bank. 2020d. “Sri Lanka—Country Partnership Framework for FY17–20.” Completion and Learning Review, World Bank, Washington, DC.
- World Bank. 2020e. *The World’s Bank: An Evaluation of the World Bank Group’s Global Convening*. Independent Evaluation Group. Washington, DC: World Bank.
- World Bank Group. n.d. *Survey of Well-Being via Instant and Frequent Tracking and Iterative Beneficiary Monitoring*. Washington, DC: World Bank Group. <https://www.ifc.org/wps/wcm/connect/64f11adb-ab01-4207-93cd-dd2cc51af16c/SWIFT-booklet-05.pdf?MOD=AJPERES&CVID=m9Or9Ia>.

APPENDIXES

Independent Evaluation Group

*The World Bank Group Outcome
Orientation at the Country Level*

Appendix A. Methods





The evaluation assessed whether the country-level results system is fit for the purpose of bridging knowledge gaps on the World Bank Group's contribution to country outcomes, informing strategic decisions, and helping country teams adapt country engagements to achieve country outcomes.

The main question motivating this evaluation is, “What influences outcome management and measurement practices across Bank Group country engagements?” Two main lines of inquiry underlie this:

- What drivers and constraints influence outcome management and measurement practices across Bank Group country engagements? What can be learned from these influences? And, how did these influences change under different contexts?
- What are the opportunities for strengthening the outcome orientation of Bank Group country engagements? And, what behaviors and capacities would be necessary?

To answer these questions, the evaluation team developed a conceptual framework that unpacks various dimensions of outcome orientation (figure A.1). This framework was used to design the data collection and analysis instruments used in the evaluation, notably the protocol for the structured desk review and content analysis, the interview template, and the coding plan for sentiment analysis conducted in NVivo. It also helped organize the structured literature review around relevant modules.

Figure A.1. Four Dimensions of Outcome Orientation

	Design of country engagement and results frameworks		Monitoring, evaluation, and learning practices
	<ul style="list-style-type: none"> • Results chains and frameworks explain and measure the contribution to country outcomes • Use results frameworks and Country Partnership Framework as a road map • Explain and mitigate risks to outcomes 		<ul style="list-style-type: none"> • Monitoring, evaluation, and learning practices at portfolio level • Track progress towards the achievement of country outcomes • Feedback on what works and why, • Learn from failure
	Adaptive portfolio management practices		Client engagement practices
	<ul style="list-style-type: none"> • Use indicators and evaluations to adapt portfolio • Engage in adaptation at the portfolio level: <ul style="list-style-type: none"> - scale-up or down - balance risk-outcome - comparative advantage - track long term contributions 		<ul style="list-style-type: none"> • Strengthens clients' capacity to measure and manage development outcomes • Use country M&E systems • Engage clients in design, management and assessment of country portfolios.

Source: Independent Evaluation Group.

Note: CPF = Country Partnership Framework.

Based on a preliminary review of the literature, the evaluation team developed an archetypal definition of an outcome-oriented organization as one that promotes a shared view among staff, management, shareholders, and clients that success means

- Overcoming complex and dynamic development challenges. It involves catalyzing behavioral, institutional, and social changes that take place over the medium to long term, usually as the result of investments in organizational capacity for key development stakeholders (such as national governments, civil society, the private sector) or directly for the target population. Responding to complex development challenges demands contextual learning.
- Adopting monitoring, evaluation, and learning practices and a set of approaches, tools, staff capacities, and attributes as well as relationships

and incentives that are mobilized to systematically gather data and evidence on the organization's contribution to development outcomes and generate feedback on what is working, what is not working, why, and how can it be fixed.

- Using feedback on outcomes to adapt programs and portfolios. This requires a set of incentives, processes, and structures that supports and encourages staff and clients to apply evaluative thinking and collective decision-making to manage and adapt interventions, portfolios, and strategies. It has implications for how to engage with clients and partners.
- Ensuring outcome-oriented client engagements. This means supporting client countries' capacity to monitor, evaluate, and manage change in their key development outcome areas. Prioritizing development and use of clients' results frameworks, data, and evaluation systems, working with partners to avoid creating parallel systems that undermine this objective, and providing spaces to co-create and use outcome information.

Evaluation Approach

The evidence synthesized in this report comes from a sample of 39 country engagements carefully selected to represent the diversity of Bank Group's country clients' and portfolio. To generate findings that cut across country engagements while grasping the differences across types of portfolios and country contexts, the evaluation team relied on a typology of country engagements and selected 39 country engagements to study, enough cases to enable comparison within categories and identify challenges or opportunities that cut across country engagements. Figure A.2 shows the selected country engagements and the data collection and analysis associated with each. Among the 39 countries, 10 were studied through interviews only, 14 through desk review only, and 15 through both. IEG had

planned to carry out a field visit for six countries to be able to interview clients and partners. Because of the outbreak of the coronavirus (COVID-19) pandemic two of the missions were canceled.

The evaluation team conducted three types of data collection and analysis: (i) structured document review of country engagement cycles, (ii) key informant interviews and facilitated group discussions, and (iii) structured literature review. Figure A.2 provides an overview of data collection and analysis.

Figure A.2. Overview of Data Collection and Analysis

1. Case studies of country team practices

- 25 country engagements covered in interviews
- 185 interviews
- 5 Group brainstorming on solutions

2. Classification and structured document reviews

- 65 country products coded for 29 countries (CPF, PLR, CLR)
- 144 CPF objectives and 361 indicators classified for 5 sectors
- Content analysis of adaptive management based on GLAM typology

3. Literature Reviews

- 16 IEG and World Bank evaluation report on results management
- 25 external reports from other development partners on RBM
- 65 external reports on adaptive management (government, private sector, academia)

Source: Independent Evaluation Group.

Note: CLR = Completion and Learning Review; CPF = Country Partnership Framework; GLAM = Global Learning for Adaptive Management; PLR = Performance and Learning Review.

The evaluation design is based on principles from case-based evaluation approaches. The cases focused on country teams' outcome measurement and management practices. The evaluation team assembled a database of 52 countries that went through at least three steps of the engagement cycle between FY14 and Q2 of FY20. After a preliminary review of country and portfolio data, an initial typology of country engagements was created (table A.1). Based on this, countries were selected for in-depth inquiry across

several factors thought to influence outcome orientation. Selection factors included the following:

- Country characteristics; notably, country capacity (Country Policy and Institutional Assessment score, Worldwide Governance Indicators, Open Data Barometer, Statistical Capacity Indicator, and so on); fragility status; aid dependency and significance of the Bank Group in the country’s budget; and income group
- Bank Group country engagement and portfolio characteristics: notably, portfolio composition (portfolios dominated by advisory services and analytics and reimbursable advisory services or IFC advisory services versus portfolios with a majority of development policy financing, versus portfolios with a mix of instruments, as well as size of IFC portfolio) and country engagement experience (for example, number of cycles completed and Bank performance ratings)
- Bank Group country teams’ characteristics, including size and composition of country teams

Table A.1. Universe of Countries Organized by World Bank Group Portfolio Composition

Country Type	World Bank Lending Projects	IFC Investment			
		IFC investment dominated	IFC investment and advisory	IFC advisory only	No IFC projects
	IPF only	Uruguay	Bolivia,		
		Argentina	Bosnia and Herzegovina		
		North Macedonia	El Salvador		
			Belarus		

Country Type	World Bank Lending Projects	IFC Investment			
		IFC investment dominated	IFC investment and advisory	IFC advisory only	No IFC projects
IDA non-FCS	IPF and DPF	Armenia	Bosnia and Herzegovina	Albania	
			Indonesia, Serbia		
			Sri Lanka, Tunisia		
			Croatia, Georgia, India		
			Peru, Vietnam		
	DPF significant	Guatemala	Colombia		
		Poland	Egypt, Arab Rep.		
		Romania	Panama		
			Morocco, Paraguay		
	No projects		Bulgaria		
			Nicaragua, Tajikistan	Lesotho	Maldives
	IPF only		Zambia		
			Cameroon, Honduras	Niger	
			Uzbekistan		
			Benin, Burkina Faso		

Country Type	World Bank Lending Projects	IFC Investment			
		IFC investment dominated	IFC investment and advisory	IFC advisory only	No IFC projects
IDA FCS	IPF only		Guinea, Ghana ^a , Kyrgyz Republic, Moldova, Nepal, Senegal ^a		
			Papua New Guinea		
	IPF and DPF		Afghanistan, Côte d'Ivoire	Chad	
			Haiti		
	DPF significant		Kosovo, Myanmar, Solomon Islands		
			Burundi		

Note: IBRD-IDA lending group is based on the list of economies listed by the World Bank Group in June 2019.

Country FCS status is based on the World Bank Group's fiscal year 2019 list of fragile situations. The colors have the following meanings: **blue** = SCD-CPF-PLR; **green** = PLR-CLR-SCD-CPF; **orange** = PLR-CLR-SCD-CPF-PLR. IPF and DPF: Among IPF and DPF projects, IPF projects count greater than 70 percent and less than 100 percent. DPF significant: Among IPF and DPF projects, DPF projects count 30 percent or more. IFC investment dominated: Among IFC projects, investment projects count 90 percent or more. IFC investment and advisory: Among IFC projects, investment projects count greater than 0 percent and less than 90 percent. IFC advisory only: No IFC investment projects. IDA lending group is based on the World Bank Group's list of economies in June 2019. CLR = Completion and Learning Review; CPF = Country Partnership Framework; DPF = development policy financing; FCS = fragile and conflict-affected situation; IBRD = International Bank for Reconstruction and Development; IDA = International Development Association; IFC = International Finance Corporation; IPF = investment project financing; PLR = Performance and Learning Review; SCD = Systematic Country Diagnostic.

a. Ghana and Senegal had their CPF discussed by the Board of Executive Directors after the cutoff date for the evaluation sample but were included to correct for sample bias in the Africa Region.

Figure A.3. Typology of Sampled Country Engagements

FCS	IDA (non-FCS)	Large IBRD	Small IBRD
<ul style="list-style-type: none"> • Afghanistan • Cameroon • Chad • Haiti • Kosovo^a • Myanmar • Papua New Guinea^a • Solomon Islands 	<ul style="list-style-type: none"> • Benin • Burkina Faso^a • Botswana • Côte d'Ivoire • Ghana^a • Honduras • Lesotho • Maldives • Kyrgyz Republic^a • Nicaragua^a • Tajikistan^a • Zambia^a 	<ul style="list-style-type: none"> • Colombia • Egypt • Indonesia • Morocco^a • Peru • Sri Lanka • Tunisia • Uzbekistan • Vietnam 	<ul style="list-style-type: none"> • Albania • Bolivia • Bosnia and Herzegovina • Bulgaria • El Salvador • North Macedonia • Panama • Romania • Serbia • Uruguay

Source: Independent Evaluation Group.

Note: Countries in **bold** were studied through a combination of desk review and interviews with country teams and clients in a few cases. FCS = Fragile and Conflict-Affected Situations; IDA = International Development Association; IBRD = International Bank for Reconstruction and Development.

a. These countries were only studied through interviews because they did not have a CPF-PLR sequence; the others were only studied through desk review.

Interviews

Given that much of the outcome-oriented behaviors and decisions of country teams are not captured in formal documents, the evaluation methods focused on eliciting the tacit knowledge of teams. Through interviews and facilitated group discussions the evaluation identified, explained, and amplified informal practices. Several questions guided the case inquiry (box A.1).

Box A.1. Questions Guiding Case Inquiry

Should we give the case study questions?

- i. What are the main factors that influence the design of the Country Partnership Framework results chain, frameworks, and choices of indicators?
- ii. What kind of monitoring, evaluation, and learning activities are taking place throughout the country engagement cycle? For what purpose? How do they capture the World Bank Group's contribution to country-level outcomes?
- iii. What kind of adaptive management is carried out in the country engagement? To what extent does information from the results framework or other evaluative information or sources of data inform course correction or changes in the country engagement?
- iv. What role do clients play in outcome measurement or management? What kind of monitoring and evaluation functions exist in governments that the World Bank can leverage? Does the World Bank contribute to strengthening its clients' capacity to measure and manage outcomes?
- v. Are there any practices or behaviors adopted by the World Bank Group in this country that have helped or hindered outcome measurement or management?
- vi. What else could be done to strengthen the country teams' outcome measurement and management practices?

Cross-cutting: (i) What factors of the country context and World Bank Group engagement context are especially important in how they affect outcome measurement or adaptive management, and why?

Cross-cutting: (ii) Are there tensions/trade-offs in the current country engagement frameworks, Bank's policies, or processes that constrain the outcome measurement and management practices at the country level?

Source: Independent Evaluation Group.

Interviews with Bank Group country teams were conducted for 19 of the 30 countries in the sample. In four countries, the evaluation team also interviewed clients and partners. Table A.2 shows the targeted interviews by type at the country level. Country-level interviews were complemented with interviews with members of the authorizing environment, including senior management in the three institutions and consultations with Committee on Development Effectiveness members. Table A.3 shows the overall breakdown.

Table A.2. Targeted Key Informants by Type

World Bank Country Teams	IFC and MIGA	Clients and Partners
Country manager or operations manager	IFC country manager IFC country officer	Formal counterpart (for example, head of international financial institution division in ministry of finance, ministry of planning)
Country director	IFC investment officer for any sectors with large exposure	Other relevant staff for finance and planning ministries
Resident representative	IFC Regional director	Head of project management office (when a single project management office manages all World Bank projects)
Country Partnership Framework task team leader	IFC Regional industry directors for any sectors with large exposure	National statistical office
Performance and Learning Review or Completion and Learning Review task team leader	IFC manager covering Region in Country Economics Unit	Government evaluation functions as exist
Country program coordinator	Investment officers and operations officer within industries that are key contributors to IFC country strategies and Country Partnership Framework outcomes	Office of planning in line ministries for any sectors with large World Bank engagements
Country operation officer	MIGA management	Coordinating ministries that might run monitoring and evaluation for multiple ministries
Country economist		Delivery units within ministries
Program leaders		Development Partners head of mission / evaluation officers
Regional development effectiveness unit		
Regional director of strategy and operations		
Lead technical specialist(s) for any Global Practices with very large engagements		

Source: Independent Evaluation Group.

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

Table A.3. Breakdown of Interviews Conducted

Key Informants	Number
World Bank Group management or Committee on Development Effectiveness	23
World Bank country teams	91
IFC and MIGA country teams	43
Clients and partners in country	28
Total	185

Source: Independent Evaluation Group.

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

Structured Desk Review

The overall objective of the structured document review was to identify patterns of outcome orientation exhibited in Bank Group country engagement documents and, to the extent possible, identify explanatory factors driving differences in outcome measurement and management practices. More precisely, the structured document review was guided by the following specific questions:

1. What is the outcome level of Country Partnership Framework (CPF) objectives and indicators?
2. To what extent do country engagement documents capture contributions from International Finance Corporation (IFC) / Multilateral Investment Guarantee Agency (MIGA) and other development partners to CPF objectives and country development outcomes?
3. To what extent do country engagement documents articulate Bank Group contribution to Sustainable Development Goals?
4. What is the quality and reporting rate of CPF indicators?

5. To what extent do country engagement documents refer to evaluation activities for assessing CPF objectives (other than metrics from the results frameworks)?
6. How are CPF objectives and indicators adapted throughout country engagements?
7. What adaptive management practices can be identified throughout country engagements? What type of information drives these practices?
8. How does the Bank Group support clients' outcome orientation in country engagements?

To answer the structured document review questions, the team developed a coding template to extract information from the country engagement documents of a sample of countries selected for review. Appendix B provides more details on the structured review approach, including sampling/selection, coding protocols, and limitations. As part of the structured review exercise, the evaluation team classified CPF objectives and CPF results frameworks indicators using a classification plan developed for the RAP 2020 and described in box A.2.

Box A.2. Classifying Objectives and Indicators along a Change Pathway

The Independent Evaluation Group's *Results and Performance 2020* team developed a framework to classify the various objectives of World Bank Group's interventions along a change pathway. To ensure coherence within the Independent Evaluation Group, this evaluation applied this framework to the classification of Country Partnership Framework objectives and indicators.

The main value of this framework is that it mirrors and systematizes the widely shared perspective within the Bank Group about the degrees of ambition and results achieved through its interventions. It is thus intuitive and easily understood by internal stakeholders. It also provides a single framework to classify the gamut of objectives pursued by the Bank Group and enables a cross-sector and cross-country picture.

(Level 1) Outputs: Activities deliver outputs, such as knowledge products, goods, equipment, and services. From a monitoring and evaluation standpoint, outputs tend to be readily observable and countable and their delivery can easily be tagged to a specific intervention or "attributed" to a specific organization.

(Level 2) Immediate outcomes: Outputs can have immediate effects on stakeholders or their environment. Government, private sector, and nonstate actors can gain new skills or capabilities, and citizens can have enhanced access to better quality services or environmental benefits and see a material improvement in their living conditions. From a monitoring and evaluation standpoint, with the right approach to evidence gathering, these types of changes can be identified and often measured, and a plausible link between interventions and the observed changes can be established.

(Level 3) Intermediate outcomes: Such immediate effects can trigger further changes if they have ripple effects within a sector or if they build on the changes created by other interventions;

(Level 4) Long-term outcomes: If this is sustained, long-term outcomes eventually arise with sustained changes in delivery, governance, or citizens' well-being.

The framework, however, has several limitations: (i) it uses an intervention perspective, which makes it harder to apply to packages of intervention; (ii) it conflates multiple dimensions of change processes, including time to results (how long it takes until the results materialize), longevity of results (how long the results last), types of results (encompassing functional, behavioral, and institutional), as well as target group for results (for example, government agents or rural people); and (iii) it does not distinguish between the result and its observation.

Source: Independent Evaluation Group, based on Belcher and Palenberg 2018.

Structured Literature Review

Since the advent of the “Managing for Development Results” (MfDR) agenda in the early 2000s, a vast body of evidence and research has emerged on the strengths and weaknesses of results-based management and results systems in development partners (bilateral and multilateral). Several World Bank-specific studies (IEG and non-IEG) as well as comparative reviews of various systems have been conducted, along with more specific research on various aspects of the MfDR agenda and its institutionalization within the development industry.

Given the array of information already existing, the evaluation team designed the literature review with two main objectives: First, to provide an evaluative synthesis of the existing evidence on the strengths and weaknesses of the Bank Group’s current various results systems concerning measuring, managing, evaluating, and learning from the Bank Group’s contribution to country development outcomes. Second, to identify and substantiate evidence-based good practice principles for improving the outcome orientation of development agencies.

The literature review was organized as follows:

- Internal literature review (World Bank Group documents): A systematization of common good practices, limitations, and

recommendations of World Bank Group authored documents relevant to the nature of the evaluation.

- External literature review: The review of external literature had two components:
 - A systematization of common limitations, good practices, and recommendations of gray and academic literature on results systems and results-based management. The team also consulted with experts on outcome orientation within and beyond development organizations. The core streams of literature included policy innovation, administrative science, public service delivery, government labs, and deliverology. The evaluation team also explored bodies of literature that grapple with outcome-oriented behaviors, incentives, and motivations within and outside of public service delivery, such as medicine and business administration.
 - A review on emerging solutions to common challenges. The team performed a targeted in-depth literature review of both academic evidence and good practices addressing the identified shared limitations. The team derived from this literature review a number of adaptive management practices that have the potential to overcome some of the challenges identified in the evaluation.

Table A.4 provides a summary of the literature reviewed by types of documents that the evaluation team analyzed and their authors or institutions.

Table A.4. Overview of the Reviewed Literature

Document Type	Documents Revised (no.)	Category	Examples of Authors or Topics
Internal Review (World Bank Group Documents)	16	World Bank Group	Self-Evaluation Systems, Corporate Scorecards, Program-for-Results, Results Frameworks Impact Evaluations, Monitoring and Evaluation Systems
External Review	80	Multilateral or international organizations	Asian Development Bank, United Nations Population Fund, United Nations Children's Fund, United Nations Development Programme, Inter-American Development Bank, Organisation for Economic Co-operation and Development, United States Agency for International Development, and so on.
		Academia	Harvard's Policy Performance Lab, MIT Media Lab, Oxford Policy Management, University of Wisconsin Madison, Stanford Center for International Development, and so on.

Document Type	Documents Revised (no.)	Category	Examples of Authors or Topics
		Governments	Australian Audit Office, Finland Ministry of Foreign Affairs, Government Labs and Government Centers (Denmark, United States, Chile, Argentina, Brazil), and so on.
		Private sector and International nongovernmental organizations	McKinsey Center for Government, Overseas Development Institute, Delivery Associates, Center for Global Development, Oxfam, Nesta, and so on.
		Development cooperation agencies	United States Agency for International Development, Department for International Development (UK), UK Aid, and so on.

Source: Independent Evaluation Group.

Limitations

The limitations of this evaluation revolve around three main challenges: delimitations, perspectives, and generalizability of the findings:

- This evaluation’s scope focused on the country-level results system that deliberately leaves out the project-level results system, which is outside the boundaries of this evaluation. The project-level results system had been evaluated extensively in the evaluation *Behind the Mirror: A Report on the Self-Evaluation Systems of the World Bank Group* (World Bank 2016) and in the report *Assessing the Monitoring and Evaluation Systems of IFC and MIGA* (World Bank 2013), and this evaluation systematically reviewed and integrated in its analysis the findings from these and other IEG reports that tackled the project evaluation system, including the RAP 2016 on Managing

for Results, and the Learning Evaluation (2015). However, the country-level and project-level results system are intertwined because most indicators and data presented at the country level come directly from the project level. Thus, this evaluation focused on assessing the adequacy of the project-level data for a country-level results system; it did not cover the adequacy of the project-level system as such.

- The evaluation also did not manage to cover in depth the issue of client capacity in monitoring and evaluation and World Bank's efforts to build capacity through its projects and other interventions as well as the strengths and limitations of these interventions. This topic was partially covered in the IEG evaluation on Data for Development (World Bank 2017).
- This evaluation purposefully focused on country teams' practices, which is reflected in the sampling and selection of interviews. Although Global Practice staff were interviewed (primarily Practice Leads and Technical Leads), the Global Practice perspective is not covered in depth. Whereas the evaluation had planned to gather the clients' perspective through field missions, the team was able to do so only in four countries, instead of the planned six countries, because of the interruption of missions for COVID-19 and the inability to reach clients by phone.

The evaluation focused on 39 country engagements but built into its methodology several principles for reaching analytical generalizability. First, the countries were selected to cover well the diversity of Bank Group country engagements, following a transparent and elaborate typology. Second, the number of cases was selected to allow corroboration of findings within and across types. Third, the team applied principles of data saturation—analyzing and gathering more evidence until no new information or themes are observed in the data—to interviews and content analysis of structured document. There was an unexpected amount of convergence in the findings across country contexts, informants, and data sources, despite the heterogeneity in the sample of countries. Finally, the team found a very high degree of convergence between its findings and themes emerging from the

literature review, which gives further credence to the transferability of the findings outside of the sampled cases.

References

- World Bank. 2013. *Results and Performance of the World Bank Group 2012*. Independent Evaluation Group. Washington, DC: World Bank.
- World Bank. 2016. *World Bank Operations Evaluation Department: The First 30 Years*. Independent Evaluation Group. Washington, DC: World Bank.
- World Bank. 2017. *Data for Development: An Evaluation of World Bank Support for Data and Statistical Capacity*. Independent Evaluation Group. Washington, DC: World Bank.

Appendix B. Structured Document Review

Objective

The overall objective of the structured document review was to identify patterns of outcome orientation exhibited by World Bank Group country engagements and, to the extent possible, identify explanatory factors driving differences in outcome measurement and management practices. More precisely, the structured document review was guided by the following specific questions:

- What is the outcome level of Country Partnership Framework (CPF) objectives and indicators?
- To what extent do country engagement documents capture contributions from the International Finance Corporation (IFC) and the Multilateral Investment Guarantee Agency (MIGA) and other development partners to CPF objectives and country development outcomes?
- To what extent do country engagement documents articulate Bank Group contribution to Sustainable Development Goals (SDGs)?
- What is the quality and reporting rate of CPF indicators?
- To what extent do country engagement documents refer to evaluation activities for assessing CPF objectives (other than metrics from the results frameworks)?
- How are CPF objectives and indicators adapted throughout country engagements?
- What adaptive management practices can be identified throughout country engagements? What type of information drives these practices?

- How does the Bank Group support clients' outcome orientation in country engagements?
- What are costs associated with running the country-level results system?

To answer the above questions, the evaluation team developed a coding template and applied it to a sample of countries. Sections below detail the following approaches: (i) coding strategy; (ii) sampling and selection; (iii) data analysis; and (iv) limitations.

Methodology

Coding

To answer the structured document review questions, the team developed a coding template to extract information from the country engagement documents of a sample of countries selected for review. The coding template had five sections: (i) outcome orientation at the design stage; (ii) measuring outcomes; (iii) adapting country programs and portfolios; (iv) use of adaptive management practices; and (iv) supporting clients' outcome orientation. For each section, the team developed key questions (box B.1).

Box B.1. Coding Template Sections and Questions

1. Outcome Orientation at the Design Stage

What is the outcome level of Country Partnership Framework (CPF) objectives?

What is the outcome level of the indicators?

What is the share of objectives with International Finance Corporation and Multilateral Investment Guarantee Agency involvement?

To what extent do results frameworks include indicators to capture International Finance Corporation and Multilateral Investment Guarantee Agency contributions under CPF objectives?

To what extent CPFs discuss collaboration or coordination with other donors for achieving CPF objectives?

To what extent CPFs discuss World Bank Group contribution to the Sustainable Development Goals?

2. Measuring Outcomes

Type of indicators used (project versus holistic).

What is the share of CPF indicators kept as originally formulated at the Performance and Learning Review stage?

What is the share of reliable quality CPF indicators?

Use of evaluation exercises (other than metrics from the results framework).

3. Adapting Programs and Portfolios

Rate of CPF objectives and indicator change.

Rationales for CPF objectives and indicator changes.

4. Adaptive Management Practices

Portfolio health check.

Portfolio management actions.

Adapting to context.

Assessing scale-ups or scale-downs.

Assessing cumulative effects.

Balancing risks and outcomes.

Reflecting on comparative advance.

5. Supporting Clients' Outcome Orientation

Country engagement provides support for statistical capacity building.

Country engagement articulates use of country systems.

Source: Independent Evaluation Group.

Outcome orientation at the design stage. The objective of the first section was to extract information on the outcome level of CPF objectives and indicators. To achieve this objective, the team used a deductive-inductive approach for developing a typology of different outcome levels for objectives and indicators (box B.2). In addition, the section also included questions to capture (i) whether results frameworks measure IFC and MIGA contributions under CPF objectives; (ii) whether CPFs discuss collaboration or coordination with other donors for achieving CPF objectives; and (iii) whether CPFs discuss the Bank Group's contribution to the SDGs. To answer these questions, the team reviewed CPF documents.

Box B.2. Outcome-Level Typology

Level 1—Output. Product or service provided within the control of the client.

Level 2—Immediate outcome. Development of the capability of a group or organization, an initial benefit to people, infrastructure delivered.

Level 3—Intermediate outcome. Stakeholders apply a new capability to solve an issue; change in the lives of ultimate beneficiaries.

Level 4—Long-term outcome. A sustained change in delivery or governance or a sustained benefit to an ultimate beneficiary.

Source: Independent Evaluation Group.

Measuring outcomes. The second section sought to measure (i) the type of indicators used (project-level indicators versus holistic indicators attempting to measure the combined effect of several Bank Group interventions at the country or sector level); (ii) the share of CPF indicators kept at the Performance and Learning Review (PLR) stage as originally formulated; (iii) the share of reliable quality indicators for measuring achievement of CPF objectives; and (iv) the use of evaluation exercises (other than the metrics in the results framework) to measure the achievement of CPF objectives. Indicators were considered of “reliable quality” when they included baselines and targets consistent with the CPF time period and when they specified their data sources. To answer these questions, the team reviewed CPF documents and, when applicable, PLR documents (that is, for answering the question on whether CPF indicators were kept at the PLR stage as originally formulated).

Adapting country program and portfolios. The goal of the third section was to measure the extent to which CPF objectives and indicators are adapted throughout country engagements. To achieve this goal, the team reviewed PLRs to identify the extent to which objectives and indicators are modified at the PLR stage. In addition, the team used a deductive-inductive coding approach to classify the rationales for changing objectives and indicators. To

answer the questions in this section, the team reviews PLR documents whereby country teams record changes to CPF objectives and indicators.

Adaptive management practices. The fourth section sought to identify adaptive management practices in country engagements. Adaptive management is defined as “an intentional approach to making decisions and adjustments in response to new information and changes in context.” To identify adaptive management practices in country engagements, the team followed a deductive approach that relied on analytical categories developed by the Global Learning for Adaptive Management (GLAM)¹ initiative (box B.3). To answer the questions in this section, the team reviewed CPFs and, when applicable, PLRs and CLRs.

Box B.3. Type of Adaptive Management Practices with Guiding Questions for Identification

1. Portfolio health check

Is there a discussion of the projects that require more attention? Is there an overall health check of the portfolio performance? Is there a discussion of what is happening and how are things going?

2. Portfolio management actions

What are the key actions taken to maintain or improve World Bank Group portfolio performance?

3. Adapting to context

Is there a discussion of what all projects need to consider in response to shifts in context?

4. Assessing scale-ups or scale-downs

Is there a discussion of demonstration projects and the lessons learned from them? Is there a discussion of projects or approaches that need to be scaled up or down?

5. Assessing cumulative effects

Is there an assessment of the cumulative effect of multiple projects/interventions? What do results narratives look like? Are they a mechanical repetition of results framework indicators or do they provide a plausible evidence-based narrative assessing the cumulative effect of multiple projects/interventions?

6. Balancing risks and outcomes

Is there a discussion of how to balance risks against payoffs and of how the country engagement will adjust if certain risks were to materialize?

7. Reflecting on comparative advance

Is there a discussion of how the Bank Group can maximize the value of its investments and unique contributions relative to others? Is there a discussion of the areas where the Bank Group should move out or expand into ("future positioning")?

Source: Independent Evaluation Group, based on Global Learning for Adaptive Management analytical categories.

Supporting clients' outcome orientation. The fifth section aimed to measure the extent to which the Bank Group supports country statistical capacity in country engagements. The section was guided by two specific questions: (i) To what extent is the Bank Group providing support for statistical capacity building (supply side)?; and (ii) Is there an articulation in country engagement documents of the extent to which the Bank Group will leverage the use of country statistical systems for measuring CPF objectives (demand side)? To answer these questions, the team reviewed CPF documents.

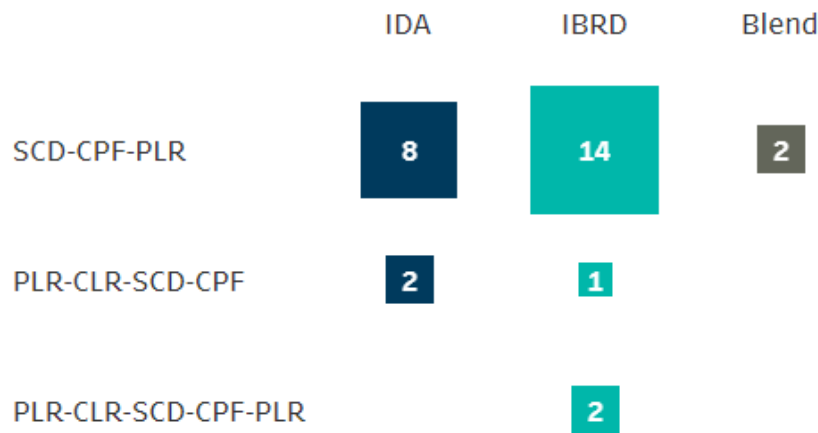
Sampling

The sampling strategy involved two approaches, namely (i) one approach for sampling countries and (ii) one approach for sampling objectives and indicators within the sampled countries. This sampling strategy was used for answering the following sections of the coding template: section 1: "Outcome Orientation at Design"; section 2: "Measuring Outcomes," and section 5: "Supporting Client's Outcome Orientation." For "Adapting

Programs and Portfolios” (Coding Template section 3) and “Adaptive Management Practices” (Coding Template section 4), the team carried out different sampling strategies that are described when the findings for these two topics are discussed in the Summary of Findings section of this appendix. See annex B.1 for a detailed mapping of coding template questions by sampling strategy.

Sampling countries for review. The team chose a purposive sample of 29 countries for review out of a sampling frame of 52 countries that had at least three steps of the new country engagement model completed between FY14 and FY20.² Within the sampling frame, countries were classified into three categories: (i) countries with a SCD-CPF-PLR cycle; (ii) countries with a PLR-CLR-SCD-CPF cycle; and (iii) countries with a PLR-CLR-SCD-CPF-PLR cycle (figure B.1). These categories reflect different levels of progress under the new country engagement model. The team used these categories as a criterion for its purposive sampling. In addition, the team also sought to ensure representation by lending group (International Development Association, International Bank for Reconstruction and Development, and Blend). The characteristics of the purposive sample chosen are displayed in figure B.1 and included (i) 24 countries with an SCD-CPF-PLR (out of 24); (ii) three countries with a PLR-CLR-SCD-CPF (out of 26); and (iii) two countries with a PLR-CLR-SCD-CPF-PLR cycle (out of two). The number of country engagement documents associated with these sampled countries was 55 (29 CPFs and 26 PLRs; table B.1). See table B.2 for detailed list of countries sampled.

Figure B.1. Sampled Countries by Lending Group and Country Engagement Cycle



Source: World Bank, Independent Evaluation Group based on Operations Policy and Country Services Data.

Note: *n* = 29 countries. CLR = Completion and Learning Review; CPF = Country Partnership Framework; IBRD = International Bank for Reconstruction and Development; IDA = International Development Association; PLR = Performance and Learning Review; SCD = Systematic Country Diagnostic.

Table B.1. Country Engagement Documents Reviewed

Stage	CPF	PLR
SCD – CPF – PLR	24	24
PLR – CLR – SCD – CPF	3	—
PLR – CLR – SCD – CPF – PLR	2	2
Total	29	26

Source: Independent Evaluation Group.

Note: *n* = 29 countries. CLR = Completion and Learning Review; CPF = Country Partnership Framework; PLR = Performance and Learning Review; SCD = Systematic Country Diagnostic.

Table B.2. Detailed List of Countries Sampled

Country	Lending Group	Stage
Afghanistan	IDA	SCD-CPF-PLR
Albania	IBRD	SCD-CPF-PLR
Benin	IDA	PLR-CLR-SCD-CPF
Bolivia	IBRD	SCD-CPF-PLR
Bosnia and Herzegovina	IBRD	SCD-CPF-PLR

Country	Lending Group	Stage
Botswana	IBRD	SCD-CPF-PLR
Bulgaria	IBRD	SCD-CPF-PLR
Cameroon	Blend	SCD-CPF-PLR
Chad	IDA	SCD-CPF-PLR
Colombia	IBRD	SCD-CPF-PLR
Côte d'Ivoire	IDA	SCD-CPF-PLR
Egypt, Arab Rep.	IBRD	SCD-CPF-PLR
El Salvador	IBRD	SCD-CPF-PLR
Haiti	IDA	SCD-CPF-PLR
Honduras	IDA	SCD-CPF-PLR
Indonesia	IBRD	SCD-CPF-PLR
Lesotho	IDA	SCD-CPF-PLR
Maldives	IDA	SCD-CPF-PLR
Myanmar	IDA	SCD-CPF-PLR
Panama	IBRD	SCD-CPF-PLR
Peru	IBRD	PLR-CLR-SCD-CPF-PLR
Romania	IBRD	PLR-CLR-SCD-CPF
Serbia	IBRD	SCD-CPF-PLR
Solomon Islands	IDA	PLR-CLR-SCD-CPF
Sri Lanka	IBRD	SCD-CPF-PLR
Tunisia	IBRD	SCD-CPF-PLR
Uruguay	IBRD	SCD-CPF-PLR
Uzbekistan	Blend	SCD-CPF-PLR
Vietnam	IBRD	PLR-CLR-SCD-CPF-PLR

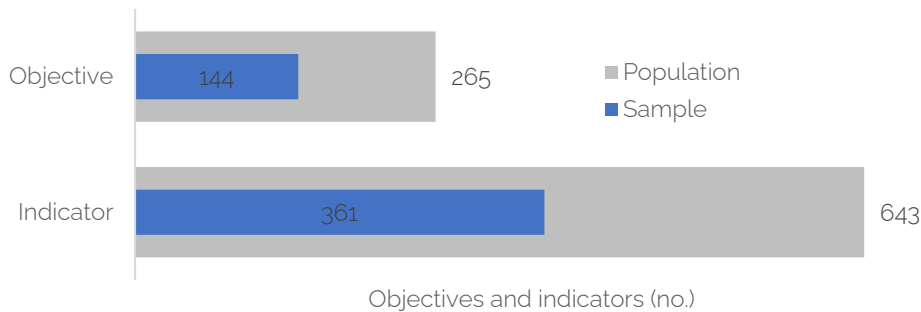
Source: Independent Evaluation Group.

Note: *n* = 29 countries. CLR = Completion and Learning Review; CPF = Country Partnership Framework; IBRD = International Bank for Reconstruction and Development; IDA = International Development Association; PLR = Performance and Learning Review; SCD = Systematic Country Diagnostic.

Sampling objectives and indicators for review. The 29 countries sampled for review included 265 objectives and 643 indicators. The team mapped these objectives against World Bank Global Practices and chose a purposive sample of 144 objectives and 361 indicators for review.³ This purposive sample accounted for more than 50 percent of the population of objectives and indicators in the 29 countries chosen for review (figure B.2). The

purposive sampling approach followed two criteria. First, the team sought to include the most representative Global Practices from each Practice Group (that is, the Global Practices with the largest numbers of objectives). Second, the team sought to include Global Practices that were assumed to differ in the outcome level of their objectives and their ability to measure achievement of objectives. For instance, the team chose the Education Global Practice because their objectives were assumed to have a higher outcome level than, for instance, objectives in the Transport and Water Global Practices. In addition, when sampling Global Practices, the team sought to include variation in the ability to measure results by including practices that have strong measurement capacity, such as the Social Protection and Education Global Practices, and others where results measurement is more challenging and less developed (Governance, Transport, and Water Global Practices).

Figure B.2. Country Partnership Framework Objectives and Indicators by Population and Sample Size



Source: Independent Evaluation Group.

Note: n = 29 countries.

Analysis

The team used NVivo software to build a database of objectives and indicators for the 29 countries purposively chosen for review and to administer the coding template and extract data from country engagement

documents. In a subsequent step, the team used Tableau software for carrying out the data analysis and elaborating graphs to illustrate findings for each of the specific questions guiding the structured document review. The Summary of Findings section of this appendix provides a summary of the findings organized as follows: (i) Outcome Orientation at the Design Stage; (ii) Measuring Outcomes; (iii) Adapting Program and Portfolios; and (iv) Supporting Clients' Outcome Orientation.

Limitations

The findings from this structured document review have limitations that, broadly speaking, can be classified into two categories: (i) sampling and selection and (ii) data availability. First, with respect to sampling and selection, the team chose purposive sampling strategies that sought to capture a wide range of perspectives relating to outcome orientation, measurement, and adaptive management practices. Because of these purposive sampling strategies, it is not possible to generalize sample findings to the population with a known confidence and margin of error. Second, with respect to data availability, the team used country engagement documents as the data source for answering the questions included in the coding template (box B.1). For many of the “Adapting Management Practices” questions, the review found little evidence of these practices. The lack of data on these practices should not be interpreted as evidence that the Bank Group management is completely devoid of these practices but only as evidence that they are not reflected in country engagement documents.

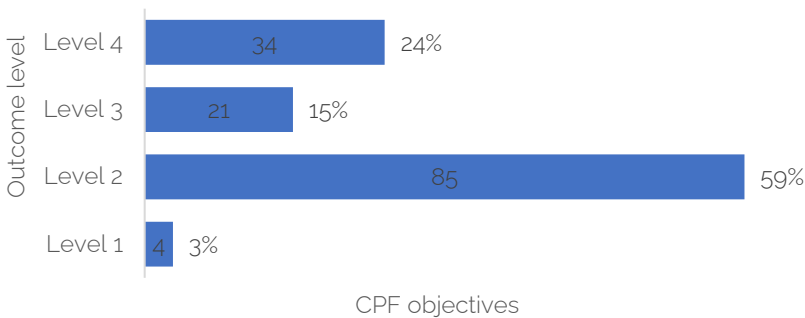
Summary of Findings

Outcome Orientation at the Design Stage

CPF Objectives

Most CPF objectives are pitched at the level of immediate outcomes. Out of the 144 CPF objectives reviewed, the team found that, in 59 percent of the cases, the formulation of the objective represents a change that happens as a direct consequence, effect, or result of a Bank Group intervention (level 2). Among this pool of objectives, 58 percent capture improved access to or quality of services as a result of a Bank Group intervention, whereas 46 percent capture improved capacity in a sector. Finally, objectives capturing intermediate and long-term outcomes represent 15 percent and 24 percent, whereas the objectives capturing level 1 represent only 3 percent (figure B.3). See box B.4 for examples of CPF objectives by outcome level.

Figure B.3. CPF Objectives by Outcome Level



Source: Independent Evaluation Group.

Note: n = 144 CPF objectives. CPF = Country Partnership Framework.

Box B.4. Examples of CPF Objectives by Outcome Level

Level 1—Output

Creating national performance management framework (Botswana CPF FY16–20).

Improved regulatory and institutional framework for key sectors (Cameroon CPF FY17–21).

Level 2—Immediate outcome

Improve access to quality education and health-related services in targeted rural areas (Indonesia CPF FY16–20).

Strengthen capacity to build transport infrastructure (Romania CPF FY19–23).

Level 3—Intermediate outcome

Building resilience at the community level (Benin CPF FY19–23).

Enhanced school outcomes for better employability (Bulgaria CPF FY17–22).

Level 4—Long-term outcome

Improving living standards in the lagging area (Sri Lanka, CPF FY17–20)

Increase accountability and transparency in public expenditures (Côte d'Ivoire, CPF FY16–19)

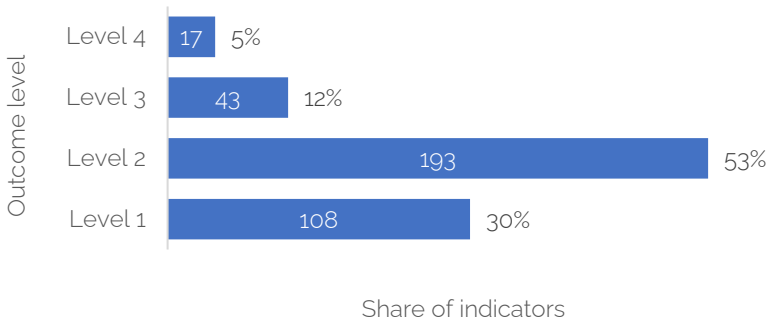
Sources: World Bank 2015c, 2015e, 2015j, 2016b, 2016g, 2017b, 2018a, 2018d.

Note: CPF = Country Partnership Framework.

CPF Indicators

CPF indicators have a lower outcome level than CPF objectives. Thirty percent of the indicators reviewed are formulated at the output level (level 1), which compares unfavorably with a 3 percent share for CPF objectives (figure B.4). Likewise, only 5 percent and 12 percent of CPF indicators are formulated at the long-term outcome (level 4) and intermediate outcome (level 3) versus 24 percent (level 4) and 15 percent (level 3) percent for CPF objectives. See box B.5 for examples of CPF indicators by outcome level.

Figure B.4. CPF Indicators by Outcome Level



Source: Independent Evaluation Group.

Note: n = 361 CPF indicators. CPF = Country Partnership Framework.

Box B.5. Examples of CPF Indicators by Outcome Level

Level 1 – Output

Percentage of line ministries collecting, analyzing, and reporting on gender data as part of their annual plans (Afghanistan CPF FY17–20)

Development of system of School Performance Indicators to measure multiple aspects of school performance in Bulgaria to inform future government interventions (Bulgaria CPF FY17–22)

Level 2 – Immediate outcome

Transaction dwell time in Douala port in days (Cameroon CPF FY17–21)

Number of primary students attending effective nonpublic schools (Haiti CPF FY16–19)

Level 3 – Intermediate outcome

Compliance rate for individual and corporate taxpayers (Indonesia CPF FY16–20).

Number of municipalities with enhanced institutional performance (Tunisia CPF FY16–20)

Level 4 – Long-term outcome

Domestic revenue-to-gross domestic product ratio (Afghanistan CPF FY17–20)

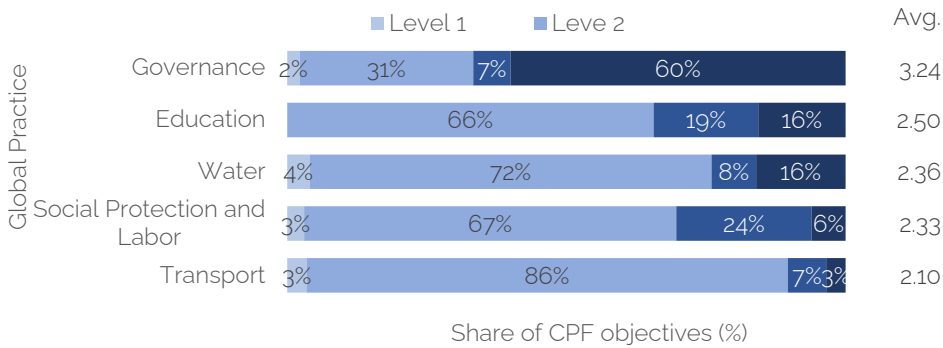
Level of fiscal deficit in percent of gross domestic product baseline (Solomon Islands CPF FY18–23)

Sources: World Bank 2015g, 2015i, 2016a, 2016b, 2016h, 2017b, 2018e.

Note: CPF = Country Partnership Framework.

Several factors are associated with the outcome level of CPF objectives and indicators. The Governance Global Practice exhibits the highest share of objectives formulated at the level of long-term outcomes (60 percent) whereas the Transport Global Practice exhibits the lowest share (3 percent; figure B.5). Likewise, this pattern is repeated for indicators, with the Governance Global Practice exhibiting the largest share of indicators (12 percent) formulated at the level of long-term outcomes (level 4) with the Transport (0 percent) and Water (1 percent) Global Practices displaying the lowest shares (figure B.6). Country income level is not associated with the outcome level of CPF objectives and indicators, but Country Policy and Institutional Assessment ratings exhibit some association. The team found that higher Country Policy and Institutional Assessment ratings in the economic management criteria⁴ are associated with a lower outcome level in CPF objectives. Finally, the team found that the outcome level in CPF indicators is higher among IBRD countries with more aid dependency (as measured by net official development assistance received as a percent of gross national income and net official development assistance received per capita).

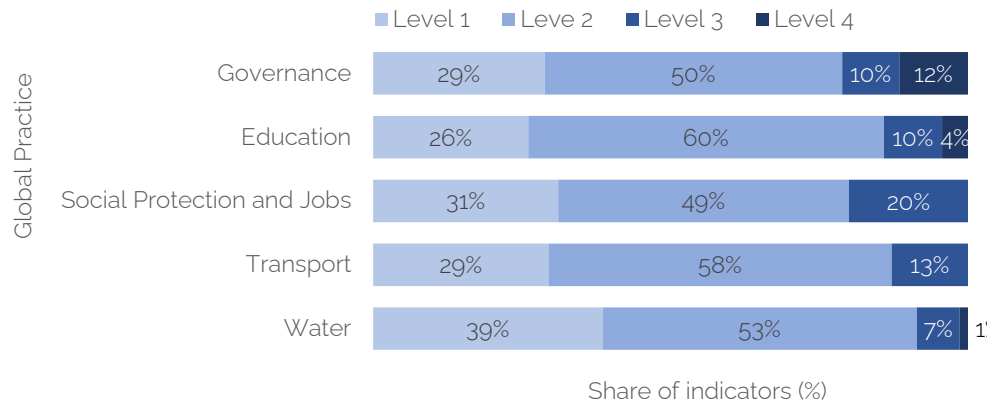
Figure B.5. CPF Objectives Outcome Level by Global Practice



Source: Independent Evaluation Group.

Note: n = 144 CPF objectives. CPF = Country Partnership Framework.

Figure B.6. CPF Indicators Outcome Level by Global Practice



Source: Independent Evaluation Group.

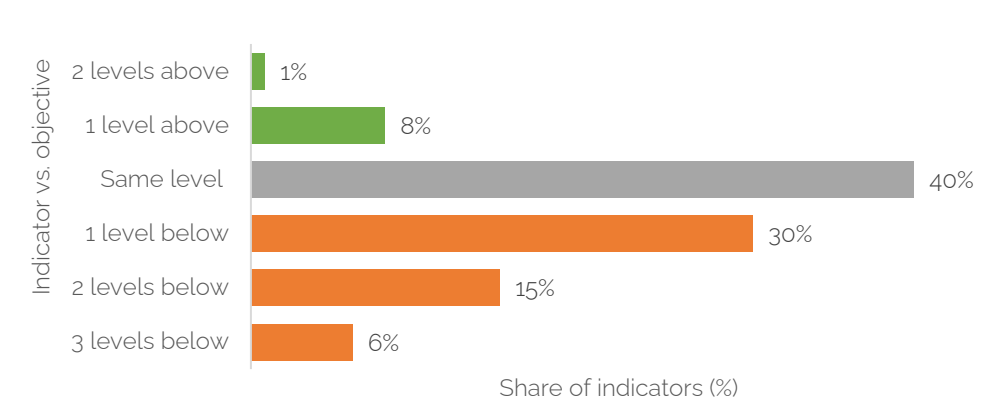
Note: n = 361 CPF indicators. CPF = Country Partnership Framework.

Alignment between Objectives and Indicators

A comparison between objectives and indicators shows a sizable mismatch between the outcome level of objectives and their associated indicators. The

team found that half of the indicators reviewed (51 percent) are below the outcome level of their measured objective, with 30 percent being one level below, 15 percent being two levels below, and 6 percent being three levels below (figure B.7). The team also calculated the average difference between the outcome level of objectives and their attached indicators and found that, on average, indicators are formulated at an outcome level that is below the outcome level of their measured objectives. This measurement gap is higher for the Governance Global Practice, where the formulated indicators are, on average, 1.3 levels below the outcome level of their measured objectives (figure B.8). A detailed mapping of the objectives’ outcome level against the outcome level of their respective indicators shows that objectives at the level of intermediate (level 3) or long-term outcomes (level 4) tend to be measured predominantly via indicators at the immediate outcome level (level 2). Figure B.9 illustrates this mismatch: almost half of the indicators (47 percent) associated with objectives at the long-term outcome level (level 4) are formulated at the immediate outcome level (level 2). See box B.6 for examples of mismatches between the outcome level of objectives and indicators.

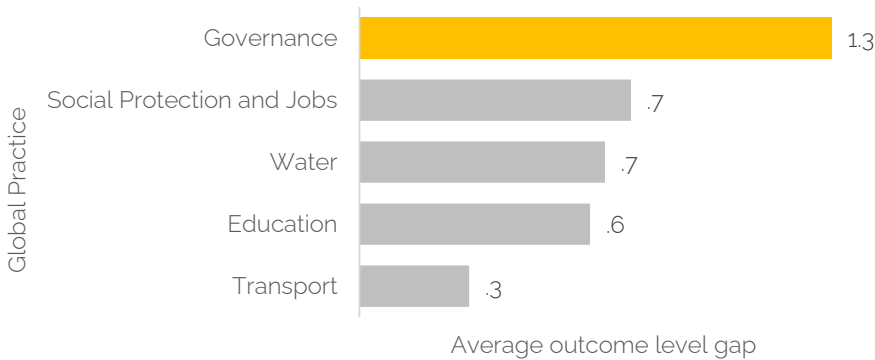
Figure B.7. Comparison of Indicator Outcome Level Against Objective Outcome Level



Source: Independent Evaluation Group.

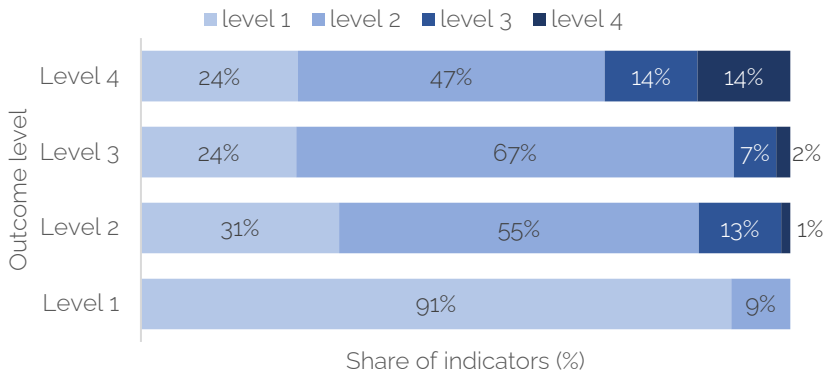
Note: n = 361 CPF indicators. CPF = Country Partnership Framework.

Figure B.8. Average Outcome-Level Gap between Indicators and Objectives, by Global Practice



Source: Independent Evaluation Group.
Note: n = 361 CPF indicators nested within 144 CPF objectives. CPF = Country Partnership Framework.

Figure B.9. Objective Outcome Level by Indicator Outcome Level



Source: Independent Evaluation Group.
Note: n = 361 CPF indicators. CPF = Country Partnership Framework.

Box B.6. Examples of High-Level Objectives Measured by Low-Level Indicators

Vietnam CPF FY18–22

Objective: Strengthen economic governance and market institutions (level 4)

Indicator: Level 1: Modern debt management law enacted by 2018 (level 1)

Afghanistan CPF FY17–20

Objective: Improved public financial management and fiscal self-reliance (level 4)

Indicator: Number of provinces where the Pension Management Information System is fully used, including the associated biometric identification (level 1)

El Salvador CPF FY16–19

Objective: Improve Secondary-school attainment (level 3)

Indicator: Number of Integrated School Systems that fully adopt the full-time school model (level 1)

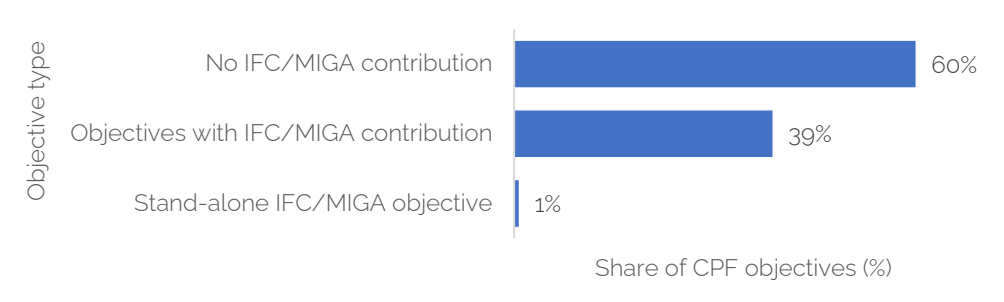
Sources: World Bank 2015f, 2016a, 2017d.

Note: CPF = Country Partnership Framework.

Capturing IFC and MIGA Contributions

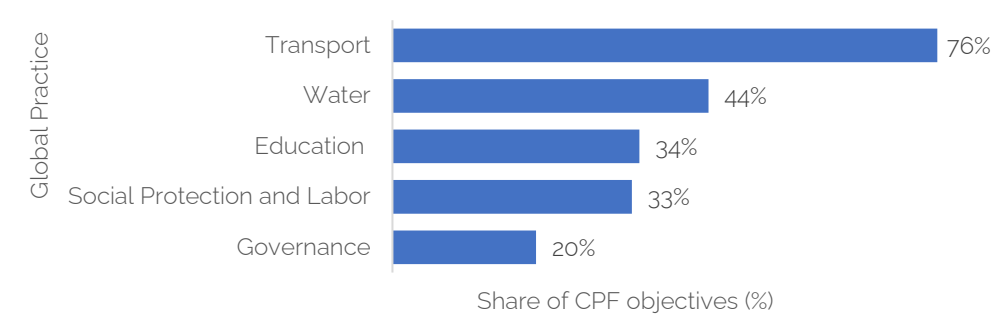
There is limited and selective involvement of IFC and MIGA under CPF objectives, and results frameworks do not fully capture the contributions from their interventions. The team found that IFC and MIGA are involved in only 40 percent of the 144 CPF objectives reviewed (figure B.10). Moreover, the team found that CPF objectives under the Transport Global Practice exhibit the largest share of IFC and MIGA involvement (76 percent), whereas objectives under the Governance Global Practice exhibit the lowest (20 percent; figure B.11). Finally, the team reviewed 31 objectives,⁵ including IFC and MIGA contributions, and found that only 29 percent of them had indicators capturing their contributions.

Figure B.10. CPF Objectives with IFC and MIGA Contributions



Source: Independent Evaluation Group.
Note: *n* = 144 CPF objectives. CPF = Country Partnership Framework; IFC = International Finance Corporation; MIGA = Multilateral Investment Guarantee Agency.

Figure B.11. Share of Objectives with International Finance Corporation Involvement in Each Global Practice



Source: Independent Evaluation Group.
Note: *n* = 29 CPF Transport Objectives; *n* = 25 CPF Water Objectives; *n* = 32 Education Objectives; *n* = 33 CPF Social Protection and Labor Objectives; *n* = 45 CPF Governance Objectives. Because objectives can be tagged to multiple sectors, they do not add up to the total number of sampled objectives (*n* = 144). CPF = Country Partnership Framework.

Capturing Other Development Partners' Contributions

In general terms, CPFs highlight collaboration or coordination with other development partners for achieving CPF objectives. Of the 29 CPFs reviewed, the team found that all mentioned collaboration and coordination with other

development partners (see box B.7 with frequency distributions and examples). By and large, these discussions are generic, with no identification of specific partners and including calls for avoiding overlap of activities, coordinating to implement complementary projects, and leveraging each other's resources. Almost 60 percent (17 out of 29) of CPF documents include more detailed discussions in which country teams identify other donors' presence in a specific area under a CPF objective. A similar share (16 out of 29) provides a more detailed discussion by identifying and, in some cases, describing other donors' contribution to a CPF objective through cofinancing, grants, and technical assistance. Finally, only one out of the 29 CPFs reviewed discussed donor coordination or collaboration for monitoring and evaluating outcomes.

Box B.7. Development Partners' Contributions to Achieving and Measuring Achievement of CPF Objectives

Mentioning collaboration or coordination with other donors (29 of 29 CPFs)

The Sri Lanka CPF FY17–20 notes that the World Bank Group will collaborate closely with other development partners to ensure coherence in approaches and maximize impact (World Bank 2016g, 37).

Identifying other donors' presence in a specific area (17 of 29 CPFs)

When discussing the infrastructure sector, the Colombia CPF FY16–21 identifies other relevant donors such as the European Union, Asian Development Bank, African Development Bank, Arab development partners, Germany, and China (World Bank 2016c, 58).

The Honduras CPF FY16–20 mentions that Bank Group interventions will not be enough for supporting the modernization of the road network and energy sector in Honduras and that, for this reason, the Bank Group support will build on efforts of partners such as the Central American Bank for Economic Integration, the German Development Agency, the Inter-American Development Bank, and the United States Agency for International Development (World Bank 2015h, 32).

Identifying other donors' contributions through parallel financing and cofinancing (16 of 29)

The Peru CPF FY17–21 describes the support provided by the World Bank Digital Inclusion Project for increasing connectivity and notes that the project will be cofinanced by the Inter-American Development Bank, and CAF (Development Bank of Latin America; World Bank 2017c, 17).

The Arab Republic of Egypt CPF FY15–19 describes parallel financing for World Bank projects supporting additional gas-fired conventional power supply. The partners expected to provide the parallel financing are the Islamic Development Bank, the Arab Fund for Economic and Social Development, the Kuwait Fund, the OPEC Fund for International Development, the African Development Bank, and the European Investment Bank (World Bank 2015b, 31).

Discussing donor coordination or collaboration for monitoring and evaluating outcomes (1 of 29)

The Lesotho CPF FY16–20 notes that the World Bank will partner with the European Union and the United Nations Development Programme to leverage its work on monitoring the Sustainable Development Goals and World Bank–financed work in statistics and evidence-based policy making (World Bank 2016d, 16).

Note: $n = 29$ CPF documents. CPF = Country Partnership Framework.

Capturing Contributions to SDGs

The Bank Group's contribution to the SDGs is not systematically addressed in CPFs. Of the 29 CPFs reviewed, the team found that 14 were dated January 1, 2016, or later; that is, after the SDGs officially came into force. The team reviewed these 14 CPFs and found that 6 included discussions addressing the Bank Group's contribution to SDGs.⁶ Out of these six CPFs, the team found (i) four linking CPF objectives to relevant SDGs with no further elaboration on how these objectives would contribute to the SDGs; (ii) two indicating how certain activities under CPF objectives align with SDGs; and (iii) two stating alignment between the CPF and SDGs in general terms (table B.3).

Table B.3. CPFs Discussing Linkages to Sustainable Development Goals

CPF	Links CPF Objectives to Relevant SDGs	Notes How Certain Activities Align with SDGs	States Alignment between CPF Objectives and SDGs
Benin FY19–23	Yes	No	No
Cameroon FY17–21	Yes	No	No
Lesotho FY16–20	No	Yes	No
Sri Lanka FY17–20	Yes	Yes	Yes
Tunisia FY16–20	No	No	Yes
Vietnam FY18–22	Yes	No	No

Sources: World Bank 2016d, 2016g, 2016h, 2017b, 2017d, 2018a.

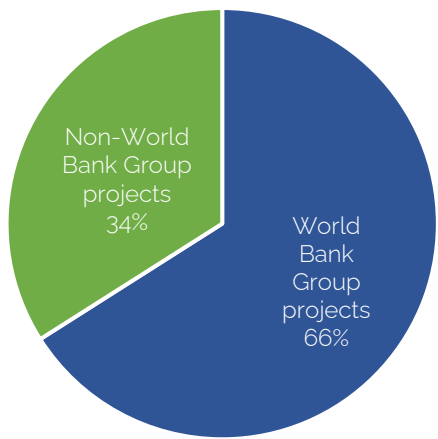
Note: *n* = 14 CPF documents. CPF = Country Partnership Framework; SDG = Sustainable Development Goal.

Measuring Outcomes

Type of Indicator

The Bank Group contribution to country-level outcomes is mainly captured through project-based indicators. Out of the 361 indicators sampled for review, the team found that 66 percent can be traced back to indicators coming from Bank Group projects (figure B.12), whereas the other 34 percent could not be traced back to Bank Group projects.⁷ The team found that these non-project-based indicators are concentrated under CPF objectives mapped to the Governance Global Practice (figure B.13) and that they typically either measure an action taken by government with Bank Group support (usually development policy financing) or measure government performance in some specific criteria (box B.8).

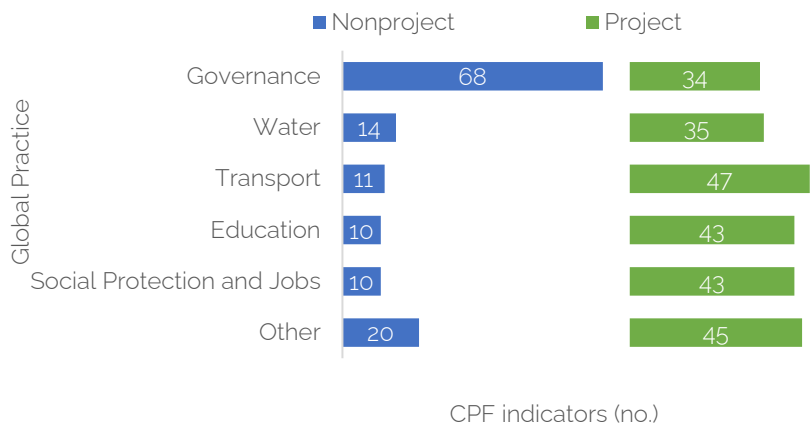
Figure B.12. CPF Indicators by Source



Source: Independent Evaluation Group.

Note: *n* = 361 CPF indicators. CPF = Country Partnership Framework.

Figure B.13. CPF Indicators Source, by Global Practice



Source: Independent Evaluation Group.

Note: *n* = 361 CPF indicators. CPF = Country Partnership Framework.

Box B.8. Examples of Non-Project-Based Indicators under the Governance Global Practice

Government action

Indicator 1.3.5. Official poverty line endorsed by government (Maldives CPF FY16–19, 40)

Indicator for objective 1. Modern debt management law enacted by 2018 (Vietnam CPF FY18–22, 49)

Government performance

Indicator 1.1.2. Domestic revenue-to-gross domestic product ratio (Afghanistan CPF 17–20, 45)

Indicator for objective 10. Corruption Perception Index (Côte d'Ivoire CPF FY16–19, 48)

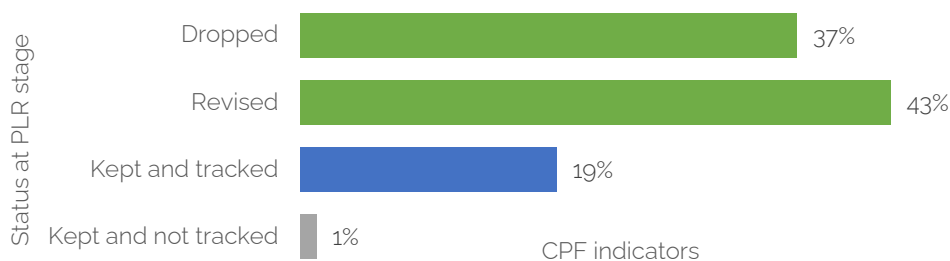
Indicator 1 under objective 1a. No accumulation of new central government arrears and arrears backlog cleared (Albania CPF FY15–19, 42)

Sources: World Bank 2015a, 2015e, 2016a, 2016e, 2017d.

Adjusting Indicators

Only a small share of CPF indicators is kept and reported as originally formulated at the PLR stage, with most of them being revised or dropped. The team reviewed 334 indicators from all sampled CPFs that had a PLR (26 out of 29). Three CPFs lacking a PLR were excluded from the review. The team found that at the PLR stage, only 19 percent of the CPF indicators were kept and tracked as originally formulated. Most of the indicators were either dropped (37 percent) or revised (43 percent; figure B.14).

Figure B.14. Status of Sampled CPF Indicators in Performance and Learning Review



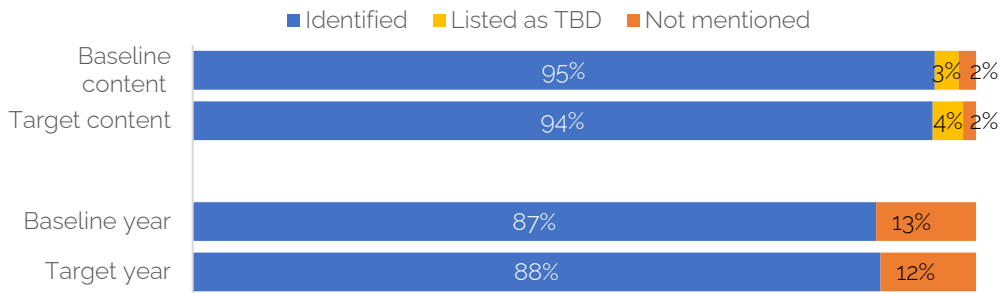
Source: Independent Evaluation Group.

Note: $n = 334$ CPF indicators corresponding to 26 sampled CPFs. CPF = Country Partnership Framework.

Indicator Quality

Almost all indicators include baselines and targets but lack a measurement time frame aligned with the CPF period and adequate documentation of data sources. The team found that most indicators identified baselines (95 percent) and targets (94 percent), and slightly smaller shares listed baseline (87 percent) and target years (88 percent; figure B.15). The team also compared the baseline and target years with the CPF period and found that, among indicators with baseline and target years mentioned, 16 percent of baselines (49 out of 315) and 24 percent of targets are not aligned with the CPF period (76 out of 317; figure B.16). In this sense, the team found that 53 percent of the indicators provided both baseline and target years consistent with the CPF period (191 out of 361).⁸ Finally, the team found that 63 percent of the indicators did not list their data sources (227 out of 361).

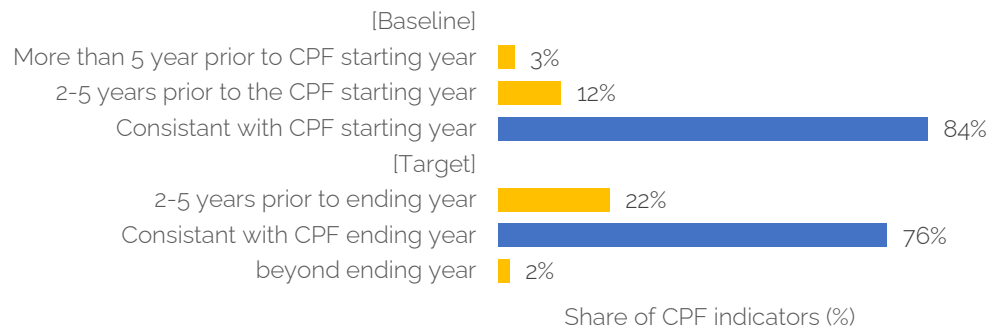
Figure B.15. Baselines and Targets in CPF Indicators



Source: Independent Evaluation Group.

Note: *n* = 361 CPF indicators. CPF = Country Partnership Framework; TBD = to be determined.

Figure B.16. Comparison between CPF Period and Years of Baseline and Target



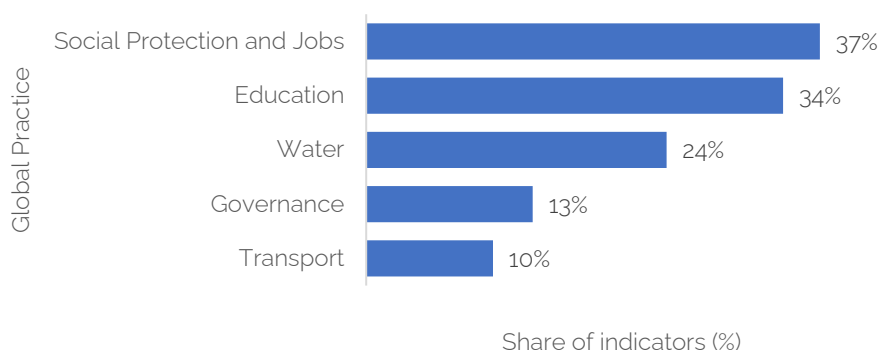
Source: Independent Evaluation Group.

Note: *n* = 343 CPF indicators with baseline; 339 CPF indicators with target. CPF = Country Partnership Framework.

Overall, the share of reliable quality indicators is low and associated with some specific sectors and the overall number of indicators in a CPF. The team coded an indicator as being “reliable” if it met two criteria: (i) an adequately specified baseline and target consistent with the CPF period, and (ii) data source specification. Based on this coding, the team found that only

20 percent (73 out of 361) of indicators can be considered “reliable.” Moreover, the team found that CPF objectives mapped to the Social Protection and Jobs and Education Global Practices exhibit a higher share of “reliable” indicators than objectives mapped to the Governance and Transport Global Practices (figure B.17). Finally, the team found a statistically significant negative correlation between the number objectives and indicators in a CPF and the likelihood that an indicator will be of “reliable quality.”

Figure B.17. Share of Reliable Indicators, by Global Practice



Source: Independent Evaluation Group.

Note: $n = 361$ CPF indicators; Reliable = (1) adequately specified baseline and target consistent with the CPF period and (2) data source specification. CPF = Country Partnership Framework.

Use of Evaluation

Planning for other evaluation activities (other than collecting data on indicators included in results frameworks) is not uncommon in CPFs and PLRs. The team found that 41 percent of the CPFs (out of 29 reviewed) and 50 percent of the PLRs (out of 26 reviewed) referred to plans for carrying out impact evaluations either in the main text or in their annexes. Other planned monitoring and evaluation activities mentioned in CPFs and PLRs include client/beneficiary satisfaction surveys, impact assessments, third-party monitoring and evaluations, using data from census and household

surveys, and gender-sensitive analyses. See box B.9 with examples of evaluation activities mentioned in CPFs and PLRs.

Box B.9. Examples of Evaluation Activities Mentioned in CPFs and PLRs

Impact Evaluation

The Benin CPF FY19–23 notes that the World Bank Group's Gender Innovation Lab will conduct an impact evaluation to assess the gender differential and overall impact of information campaigns, sensitization activities, and short-term technical retraining in productive trades (31).

The Indonesia PLR FY19–23 mentions that the Bank Group will carry out an impact evaluation to assess the effectiveness of a pilot approach for improving teacher performance and accountability through citizen engagement and performance-based pay (15).

Surveys

The Haiti CPF FY16–19 indicates that the country engagement will include the use of client surveys to assess the relevance, satisfaction, and impact of the Bank Group's work (43).

Sources: World Bank 2015g, 2018a, 2018c.

Note: CPF = Country Partnership Framework; FY = fiscal year; PLR = Performance and Learning Review.

Adapting Programs and Portfolios

To answer the questions included in this section, the team chose a purposive sample of six country engagements and reviewed the adjustments to objectives and indicators that the country teams performed at the PLR sage. For these six country engagements, the team identified 60 changes to objectives and 215 changes to indicators (table B.4). The criteria for inclusion in the purposive sample was (i) the country had been selected for either desk or field-based case study, and (ii) the country had at least both a CPF and PLR prepared. In addition, the team sought to ensure representation by lending group (International Development Association,

International Bank for Reconstruction and Development, and Blend), Regions, and fragility status.

Table B.4. Changes to Country Objectives and Indicators

Country Engagement	Changes to Objectives	Changes to Indicators
Chad FY16–20	7	28
Honduras FY16–20	7	15
Indonesia FY16–10	16	50
Peru FY17–21	8	39
Serbia FY16–20	11	25
Vietnam FY18–22	11	59
Total	60	215

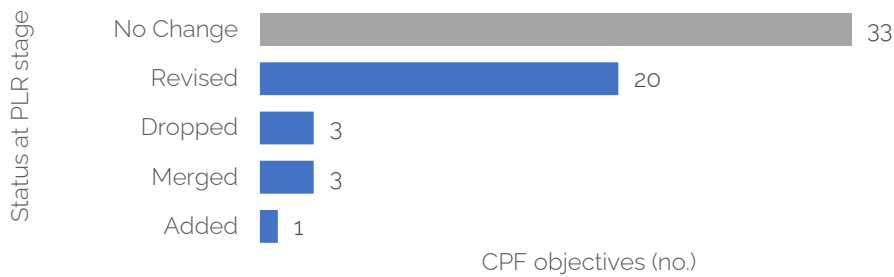
Sources: World Bank 2015d, 2015h, 2015i, 2015j, 2017c, 2017d.

Note: FY = fiscal year.

Adapting Objectives

Country teams use the PLR stage to adjust objectives and reflect changes that have taken place since the CPF’s inception. Country teams can carry out any of the following actions at the PLR stage with respect to a CPF objective: (i) leave objective unchanged (“no change”); (ii) revise objective; (iii) drop objective; (iv) merge with other objectives; and (v) add new objectives. For the six country engagements reviewed, the team identified 60 actions carried out on CPF objectives by country teams at the PLR stage. Almost half of these actions (27 out of 60) include changes to the CPF objective (20 revised, three dropped, three merged, and one added; figure B.18). The rationales for these changes are varied and include (i) redefining the scope or outcome level of the objective; (ii) aligning with the focus or progress of Bank Group’s projects; and (iii) reflecting the situation in the field or political economy reasons. See box B.10 for illustrative examples of these changes.

Figure B.18. Actions on CPF Objectives at the PLR Stage



Source: Independent Evaluation Group.

Note: *n* = 60 actions on CPF objectives at the PLR stage. CPF = Country Partnership Framework; PLR = Performance and Learning Review.

Box B.10. Examples of Changes in Objectives from the CPF to the PLR

Changing scope of objective

CPF: "Improve rural access to and quality of education"

PLR: "Improve access to and quality of education"

Explanation: "This objective must consider actions carried out throughout the territory"
(Chad, PLR FY16–20, 25)

Changing objective outcome level from outcome level 2 (immediate outcome) to outcome level 3 (intermediate outcome)

CPF: "Develop capacity for more effective management of hydrocarbon resources"

PLR: "More effective and transparent Hydrocarbon resources management."

(Chad, PLR FY16–20)

Changing or dropping the objective to reflect operational delays

CPF: "Improve fisheries management practices"

PLR: Dropped

Explanation: "Dropped as the project that contributes to this indicator is undergoing major restructuring with the PDO [project development objective] changed from improving fisheries management to coastal ecosystems monitoring and research" (Indonesia, PLR FY16–20, 51).

Reflecting country situation or government needs

CPF: "Expand coverage of social programs"

PLR: "Improve targeting of social programs"

Explanation: "Given the government's intention to maintain the same coverage of the conditional cash transfer program and the fiscal constraints, the PLR proposed to refocus objective 1 from expanding coverage to improving targeting and efficiency of social programs" (Honduras, PLR FY16–20, 7)

Source: World Bank 2019a, World Bank 2018d, World Bank. 2019b.

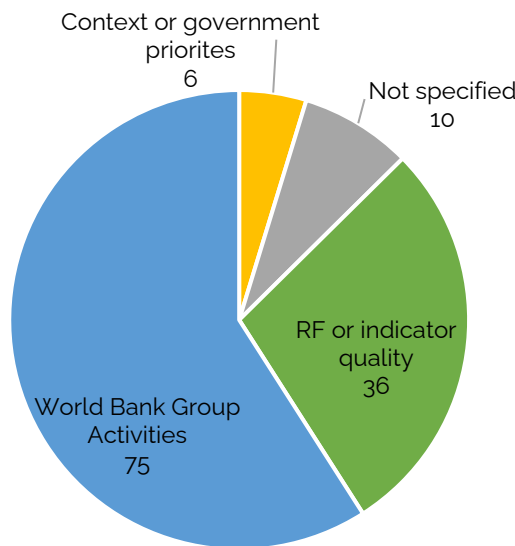
Note: CPF = Country Partnership Framework; PLR = Performance and Learning Review.

Adapting Indicators

Country teams use the PLR stage to adjust indicators in the results framework and reflect changes in Bank Group activities, country context, and government priorities that affect these indicators. For the six country engagements reviewed, the team identified 215 actions carried out on CPF indicators by country teams at the PLR stage. Of the 215 actions, the team found that 60 percent reflect changes in indicators at the PLR stage (127 indicators). These changes can be categorized into three main groups: (i) adjustments related to changes in Bank Group activities (for example, introduction of new activities, redefined scope or delays in implementation

of already existing activities, and changes in results framework indicators to align them with Bank Group’s project indicators); (ii) adjustments to changes in country context (for example, presence of external shocks such as natural disasters or economic shocks) or changes in government priorities; or (iii) adjustments to strengthen the CPF results frameworks (figure B.19). See table B.5 for a more detailed breakdown of these categories into subcategories and box B.11 for illustrative examples.

Figure B.19. Rationales for Changing Indicators at the PLR Stage



Source: Independent Evaluation Group.

Note: *n* = 127 changes to CPF Indicators at the PLR stage. CPF = Country Partnership Framework; PLR = Performance and Learning Review; RF = results framework.

Table B.5. Rationales for Adjusting Indicators

Rationales	Frequency
Related to Changes in World Bank Group Activities	
Added/dropped to align with the presence of Bank Group activities	25
Added/dropped/revised to align with supporting project's indicator	23

Rationales	Frequency
Dropped since supporting projects were canceled, dropped, delayed, restructured, scaled down, failed, unable to complete or meet the target by the end of CPF	20
Dropped/ revised to align with the focus, scope, coverage of supporting activities	7
Related to Changes in Context or Government Priorities	
Dropped because of shift in country context, political issue, lack of government support	4
Added to reflect government support	2
Related to Improving the Results Framework or Indicator Quality	
Dropped to simplify the results framework and avoid overlapping with other CPF indicator or having multiple indicators from the same project	7
Added to capture CPF achievement	6
Added to reflect the focus of updated CPF objective	5
Dropped since indicator is not relevant for objective	1
Merged to simplify the results framework	1
Revised to refine and clarify wording	3
Revised to change unit of measurement	3
Dropped due to unfeasible data collection, unavailable data, or indicator not measured by supporting project	3
Added to measure gender dimension	2
Revised to remove gender dimension	2
Added as a more measurable indicator	1

Rationales	Frequency
Split to avoid one indicator with multiple dimensions	1
Revised to correct target error in CPF	1
Other	
Reason not specified	10

Source: Independent Evaluation Group.

Note: $n = 127$ changes to CPF Indicators at the PLR stage. CPF = Country Partnership Framework; PLR = Performance and Learning Review.

Box B.11. Examples of Indicator Changes at the Performance and Learning Review Stage

Indicator added to align with the presence of World Bank Group activities

CPF: No indicator

PLR: "National Stunting Reduction Coordination mechanism launched and operational"

Explanation: "Added to track results of a new nutrition project that is scheduled to be approved by the Board in FY18" (Indonesia, PLR FY16–20, 55)

Indicator revised to align it with an indicator from a World Bank project

CPF: "Number of households benefiting from PKH (Program Keluarga Harapan), disaggregated by gender"

PLR: "PKH beneficiary families"

Explanation: "Revised to match the exact wording used and baseline and target recorded in the Project Appraisal Document of the Social Assistance Program-for-Results Project" (Indonesia, PLR FY16–20, 61)

Indicator dropped to reflect World Bank implementation challenges

CPF: "Communities implementing evidence-informed interventions that address the main risk factors affecting them"

PLR: dropped

Explanation: "Given the implementation challenges, the [World] Bank-financed Safer Municipalities Project closed without achieving its development objectives" (Honduras, PLR FY16–20, 20).

Source: World Bank 2018d, World Bank 2019b.

Note: CPF = Country Partnership Framework; FY = fiscal year; PLR = Performance and Learning Review.

Adaptive Management Practices

The team conducted a content analysis of CPF, PLR, and CLR documents to identify the presence of adaptive management practices. The categories of analysis were derived from an analytical framework elaborated by the GLAM initiative and structured into a coding template with guiding questions (box B.12). For the purpose of this review, the team reviewed a purposive sample of 25 countries. This purposive sampling approach covered all the countries chosen for either desk- or field-based case studies by the evaluation team. The following sections describe the main findings for the questions in the coding template.

Box B.12. Coding Analysis Template for Adaptive Management Practices

Portfolio Health Check

1. Identification or discussion of problem projects / portfolio at risk and overall discussion of the health status of the World Bank Group and International Finance Corporation and Multilateral Investment Guarantee Agency portfolio.
2. Use of Country Portfolio Performance Review for assessing the health status of Bank portfolio.

Key Portfolio Management Actions

3. Key actions to address Bank Group and IFC/MIGA portfolio health issues.

Adapting to Context

4. Discussion of what country portfolios need to consider in response to changes in context.

Assessing Cumulative Effects

5. Discussion of the cumulative effects of multiple projects (that is, how they collectively “move the needle” toward objectives).

Assessing Scale-Ups and Scale-Downs

6. Discussion of projects or approaches that need to be scaled down or up based on results information.

Balancing Risks and Outcomes

7. Assessment of country portfolio risks against payoffs.
8. Discussion of how country portfolios can maintain a pipeline of outcomes over different time frames.

Reflecting on Comparative Advantage

9. Discussion of how the Bank Group can maximize the value of its portfolios relative to others.
10. Identification of areas where the Bank Group portfolio should move out or expand to.

Source: Independent Evaluation Group adapted from the Global Learning for Adaptive Management initiative framework.

Portfolio Health Check

Country teams conduct portfolio “health checks” at the PLR stage focusing on disbursements, implementation bottlenecks, and problem projects. This review found that 24 out of the 25 country engagements reviewed carry out “health checks” for the Bank Group portfolio in the Portfolio Overview or Portfolio Performance section of the PLR. In 11 out of the 25, country teams report conducting the health checks via Country Portfolio Performance Reviews. With some exceptions (box B.13), the substance of the health checks (with or without Country Portfolio Performance Review) focuses on delivery issues, with little discussion of higher-level results. With respect to IFC and MIGA, the team found similar types of health checks but carried out for a slightly lower share of PLRs (18 out of 26). By and large, the discussion of IFC and MIGA health checks is less detailed than for the Bank Group portfolio.

Box B.13. Results-Oriented “Health Check”: Using Country Portfolio Performance Reviews for Engaging the Client and Discussing Results

Vietnam

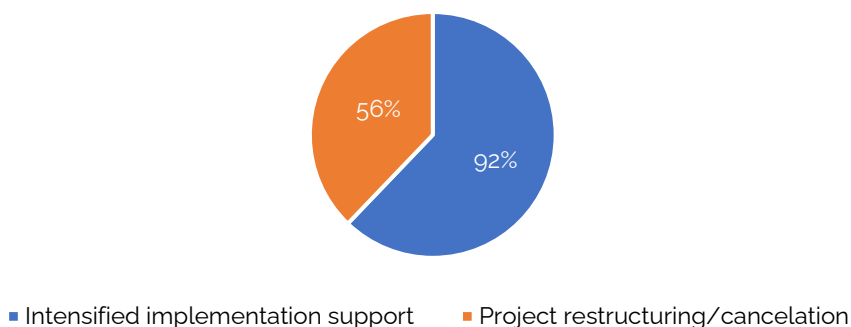
The country team put in place a three-year Programmatic Country Portfolio Performance Review (fiscal year 2013–15) to improve the operational efficiency of the Bank Group portfolio and achieve faster results. With a set of quantifiable indicators to measure the achievements, the team maintained a continuous assessment of the health of the portfolio with a view to achieving the intended results and acting on identified problems. Moreover, in consultation with the government, the country team addressed obstacles to implementation and strengthened the capacity of decentralized Project Management Units. The team noted that, thanks to this approach, the program generated strong development outcomes.

Source: World Bank 2015k.

Portfolio Management Actions

Country teams document key actions taken to address “portfolio health issues” at the PLR stage. To maintain or improve Bank Group portfolio performance, this review found that country teams document the following two actions in PLRs: (i) intensified implementation support; and (ii) project restructurings or cancellations (figure B.20). By and large, these actions are not driven by results monitoring information. Finally, the review found no detailed information on IFC and MIGA portfolio adaptations, and for this reason it was unable to generate meaningful categories.

Figure B.20. Key Actions to Maintain or Improve World Bank Portfolio Performance



Source: Independent Evaluation Group.

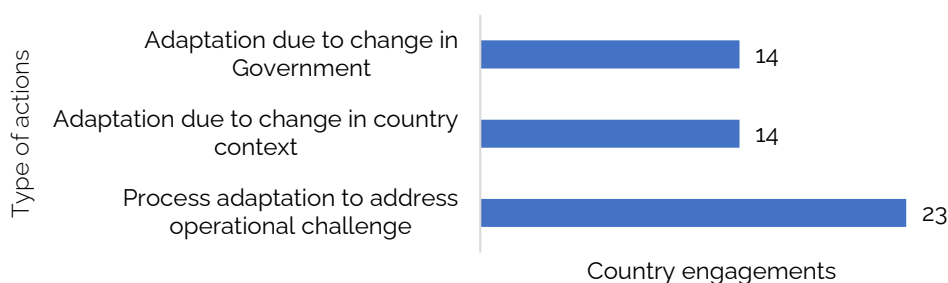
Note: $n = 25$ country engagements.

Adapting to Context

Country teams extensively report and document how they have adapted portfolios to changing contexts at the PLR and CLR stages. However, there is no evidence that PLR and CLR processes have been used to make these adaptive decisions. The review found evidence of country teams adapting portfolios, particularly at the PLR stage, through the following typology of actions: (i) adaptation due to changes in government; (ii) adaptation due to changes in country context; and (iii) process adaptation to address operational challenge (figure B.21). The first category corresponds to

adaptive actions taken to accommodate new government priorities and changes in key government actors, whereas the second category corresponds to adaptive actions driven by unexpected events such as natural disasters, pandemics, sudden increase of migration flows, and civil unrest. Finally, the third category corresponds to adaptive actions taken to address and improve Bank Group portfolio implementation (thus, this category is usually about actions addressing operational challenges). Overall, the review found no evidence indicating that PLR and CLR exercises instigated these adaptive actions. If anything, they seem to have captured actions that country teams had already put in place before conducting the PLR and CLR exercises.

Figure B.21. Adapting to Context, by Type of Action



Source: Independent Evaluation Group.

Note: $n = 25$ country engagements.

Assessing Scale-Ups or Scale-Downs

Few country engagements discuss whether to scale up or scale down new programs or approaches based on results information. Out of the 25 countries reviewed, the review found that (i) 13 lacked discussions on whether to scale up or down programs or approaches based on results evidence; (ii) 4 discussed whether to scale up or scale down Bank Group engagement in certain areas based on results information from the portfolio or from specific demonstration projects; (iii) 3 discussed scale-ups or scale-downs based on output data; and (iv) 5 justified the need for scale changes by simply asserting the relevance or nonrelevance of an area without

reference to results or outputs information (figure B.22). See box B.14 for examples of scaling-up decisions based on results information.

Box B.14. Examples of Scale Changes based on Results Information

North Macedonia

The North Macedonia Country Partnership Framework for fiscal year (FY)19–23 reports that rigorous results information of the performance of a World Bank–supported social protection program influenced the government's decision to replicate the program's design features in a newly created cash entitlement called the "education allowance."

Colombia

While discussing the lessons learned from the Bank Group support for the Fourth Generation Toll Roads Program through a Maximizing Finance for Development (MFD) approach, the Colombia Performance and Learning review for FY16–21 discusses Bank Group plans for expanding the MFD approach to support Colombia's energy sector and Bogotá's first metro line. In addition, the Performance and Learning Review also reflects on the demonstration effect that the Fourth Generation Toll Roads Program had for the MFD approach in Colombia.

Sources: World Bank 2019c, 2019d.

Assessing Cumulative Effects

PLRs and CLRs lack a discussion of how Bank Group projects collectively contributed toward achieving CPF objectives and country development outcomes. At the CPF stage, country teams articulate *ex ante* how Bank Group interventions will contribute to objectives (20 out of the 25 countries). This articulation is thorough and typically included in both the narrative of the CPF document and the results framework annex. However, at the PLR and CLR stages, there is only a mechanical discussion ("bean counting") of project results achieved under each CPF objective. For the most part, these discussions do not include a thorough assessment of cumulative effects (namely, a discussion of how the different Bank Group

interventions “moved the needle” toward CPF objectives and country development outcomes).

Balancing Risks and Outcomes

Country engagement documents include few discussions on how to balance risks against payoffs and on how the Bank Group engagement will adjust if certain risks materialize. Except for two cases (box B.15), the review found that country engagements lack a discussion of how the Bank Group will “hedge” against different types of risk (for example, if engagement in sector X were to fail, how would the Bank Group adapt to still deliver a stream of results over the engagement cycle?). The absence of these discussions does not necessarily mean that this type of thinking is not present in Bank Group management decisions, but this review was unable to find examples of it in country engagement documents. Finally, the review did not find country engagements including reflections on how to maintain the outcome trajectory across different CPF time frames.

Box B.15. Examples of Balancing Risks against Payoffs

Romania Performance and Learning Review, fiscal year (FY)14–17

In the Romania Performance and Learning Review for FY14–17, the country team notes that the “Bank [Group] will not avoid pursuing high-risk, high-reward operations where it has a comparative advantage.” (13) In this sense, the country team notes that the Bucharest Urban Development Project (FY18) is a high-risk project that could have a transformative impact on urban development and that, for this reason, it is worth pursuing. The country team decision is supported by the Bank Group’s accumulated knowledge and its ability to bring global best practices to support project design and implementation.

Bosnia and Herzegovina Performance and Learning Review, FY16–20

In the Bosnia and Herzegovina Performance and Learning Review for FY16–20, the country team discusses the Bank Group's readiness for taking a high-risk, high-reward approach and concludes that taking such an approach is warranted when there is strong government commitment for reform and program implementation.

Source: World Bank 2016h, 2018e.

Reflecting on Comparative Advantage

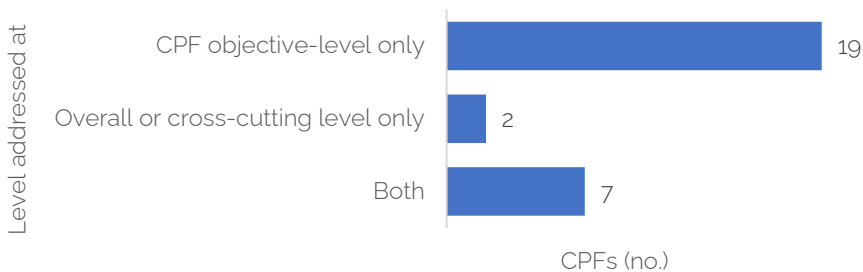
Country engagements only include weak reflections on the Bank Group's comparative advantage, and discussions about the Bank Group's future positioning are driven by a combination of SCD evidence and country demand. The review found that most country engagement documents make generic references to donors' coordination and some refer to the division of labor with other donors, but few, if any, discuss the Bank Group's unique contribution relative to other donors and the areas where the Bank Group should scale up or down based on its comparative advantage. Finally, the review found that the ex ante analytical exercises (primarily, the SCD) and country priorities and demand play a key role in the reflections about the future positioning of the Bank Group in the country. This sort of reflection, driven by the SCD and country priorities and demand, is present in all country engagements reviewed, which is hardly surprising; it is what Operations Policy and Country Services guidelines prescribe (and country teams diligently follow guidelines). The review only found three examples of reflections on future positioning influenced by results information.⁹

Supporting Clients' Outcome Orientation

CPFs provide direct support for countries' statistical capacity but provide no clear articulation of how the Bank Group will leverage the use of country systems. There are two routes through which the Bank Group can support country statistical capacity building: (i) by using country systems (demand side) and (ii) by providing direct support for statistical capacity (supply side).

The team found that all CPFs except one (El Salvador FY16–19) clearly articulate Bank Group support for statistical capacity via projects and analytical and advisory work. In contrast, the team did not find any CPFs clearly articulating how the Bank Group will leverage the use of country systems. With respect to direct support for statistical capacity building, the team found that it is provided under specific objectives in most CPFs and that, in a minority of cases, as cross-cutting support (figure B.22). Some examples of this support include developing a data system or platform for a sector, enhancing indicator measurement, improving data collection such as household surveys, and conducting gender-related data collection and analysis (box B.16).

Figure B.22. CPFs Addressing the World Bank Group's Support for Statistical Capacity Building



Source: Independent Evaluation Group.

Note: n = 29 CPFs. CPF = Country Partnership Framework.

Box B.16. Examples of the Bank Group's Support for Statistical Capacity Development in Country Partnership Frameworks

Overall Level or Cross-Cutting Area

The Afghanistan Country Partnership Framework (CPF) for fiscal year (FY)17–20 (40) notes that the Bank Group is providing technical assistance to the Central Statistics Organization on national accounts, trade and price statistics, data analysis, and the Integrated Business Establishment Survey, and to the Ministry of Economy on the Afghanistan Living Conditions Survey.

The Romania CPF FY19–23 (19) indicates that the Bank Group will support mainstreaming evidence-based policy making through two reimbursable advisory services (RASs). The National Statistical System RAS will support the development of an integrated, IT-based statistical system, whereas the Modeling the Romanian Economy RAS will help the government evaluate the impact of policy changes and shocks to the economy (Modeling the Romanian Economy RAS).

Under a CPF Objective

Under objective 4 (“Enhanced school outcomes for better employability”), the Bulgaria CPF FY17–22 (24) notes that the Bank Group will provide analytical and advisory services to the Ministry of Education and Science for developing a system of School Performance Indicators, which will measure multiple aspects of school performance in Bulgaria.

Under objective 5 (“Broaden economic participation of ethnic minorities, women, and vulnerable groups”), the Vietnam CPF FY18–22 (30) notes that Bank Group will seek to strengthen the policy-making capacity of leading Vietnamese institutions related to gender issues through analytical work and technical assistance that will improve gender statistics and data

Sources: World Bank 2016a, 2016b, 2017b, 2018d.

Costs Associated with Country-Level Results System

Budgeting and tracking expenses involved in operating the country-level results system is inherently challenging. The time recording system is likely to capture the time of dedicated staff working on CPFs/PLRs/CLRs, as well as development effectiveness support. But this is only a fraction of the time spent by many others who are involved in providing information and data. This includes many GP team members, IFC colleagues, and reviewers during the quality assurance processes associated with Board deliverables, who may not charge their time to the budget code of the deliverable. Moreover, there is no distinct code for the CLR, which is part of the CPF budget. Finally, the CPF budget encompasses many activities, consultations, and so on that are well beyond the preparation of the results frameworks. In box B.17 are a

number of “costing elements” and examples of costs associated with the country-level results system.

Box B.17. Costs Associated with the Country-Level Results System

- IEG’s 2016 evaluation of Bank Group self-evaluation systems (IEG 2016) estimated the average cost of a CLR to be about \$45,000. There were 13 CLRs produced per year on average for FY18–FY20. Together these suggest an average annual cost of \$585,000
- Based on IEG budget data, the average cost of a CLRR is \$45,000 with an average annual budget of \$615,000 for FY18–FY20
- The average costs of PLRs vary by portfolio size as shown in the table below. As expected, the larger the portfolio, the more costly the PLR. There were 10 PLRs produced per year between FY18–FY20 at an average cost of \$80,000 and a rough average budget of \$800,000.

Table BB17.1. Portfolio

Portfolio size (approved since FY10) (\$, millions)	CPF (\$, thousands)	PLR (\$, thousands)
>5,000	349	134
1,000–5,000	309	76
<1,000	195	41
Average cost	284	80

Source: Independent Evaluation Group based on data from standard report and Independent Evaluation Group resource management.

Note: CLR = Completion and Learning Review; CLRR = Completion and Learning Report Review; CPF = Country Partnership Framework; PLR = Performance and Learning Review.

References

World Bank. 2015a. *Albania—Country Partnership Framework for FY15–19*. Washington, DC: World Bank.

- World Bank. 2015b. *Arab Republic of Egypt—Country Partnership Framework for FY15–19*. Washington, DC: World Bank.
- World Bank. 2015c. *Botswana—Country Partnership Framework for FY16–20*. Washington, DC: World Bank.
- World Bank. 2015d. *Chad—Country Partnership Framework for FY16–20*. Washington, DC: World Bank.
- World Bank. 2015e. *Côte D’Ivoire—Country Partnership Framework for the Republic of for the period FY16–19*. Washington, DC: World Bank.
- World Bank. 2015f. *El Salvador—Country Partnership Framework for FY16–19*. Washington, DC: World Bank.
- World Bank. 2015g. *Haiti—Country Partnership Framework for FY16–19*. Washington, DC: World Bank.
- World Bank. 2015h. *Honduras—Country Partnership Framework for FY16–20*. Washington, DC: World Bank.
- World Bank. 2015i. *Indonesia—Country Partnership Framework for FY16–20*. Washington, DC: World Bank.
- World Bank. 2015j. *Serbia Country Partnership Framework for FY16–20*. Washington, DC: World Bank.
- World Bank. 2015k. “Vietnam—Country Partnership Framework for FY12–16.” Performance and Learning Review, World Bank, Washington, DC.
- World Bank. 2016a. *Afghanistan—Country Partnership Framework for FY17–20*. Washington, DC: World Bank.
- World Bank. 2016b. *Bulgaria—Country Partnership Framework for FY17–22*. Washington, DC: World Bank.
- World Bank. 2016c. *Colombia—Country Partnership Framework for FY16–21*. Washington, DC: World Bank.
- World Bank. 2016d. *Lesotho—Country Partnership Framework for FY16–20*. Washington, DC: World Bank.

- World Bank. 2016e. *Maldives—Country Partnership Framework for FY16–19*. Washington, DC: World Bank.
- World Bank. 2016f. “Romania Country Partnership Framework for FY14–17.” Performance and Learning Review, World Bank, Washington, DC.
- World Bank. 2016g. *Sri Lanka—Country Partnership Framework for FY17–20*. Washington, DC: World Bank.
- World Bank. 2016h. *Tunisia—Country Partnership Framework for FY16–20*. Washington, DC: World Bank.
- World Bank. 2017a. “Benin—Country Partnership Framework for FY13–17.” Performance and Learning Review, World Bank, Washington, DC.
- World Bank. 2017b. *Cameroon—Country Partnership Framework for FY17–21*. Washington, DC: World Bank.
- World Bank. 2017c. *Peru—Country Partnership Framework for FY17–21*. Washington, DC: World Bank.
- World Bank. 2017d. *Vietnam—Country Partnership Framework for FY18–22*. Washington, DC: World Bank.
- World Bank. 2018a. *Benin—Country Partnership Framework for FY19–23*. Washington, DC: World Bank.
- World Bank. 2018b. “Bosnia—Country Partnership Framework and Herzegovina for FY16–20.” Performance and Learning Review, World Bank, Washington, DC.
- World Bank. 2018c. “Indonesia—Country Partnership Framework for FY16–20.” Performance and Learning Review, World Bank, Washington, DC.
- World Bank. 2018d. *Romania—Country Partnership Framework for FY19–23*. Washington, DC: World Bank.
- World Bank. 2018e. *The Solomon Islands—Country Partnership Framework for FY18–23*. Washington, DC: World Bank.
- World Bank. 2019a. “Chad—Country Partnership Framework for FY16–20.” Performance and Learning Review, World Bank, Washington, DC.

World Bank. 2019b. “Colombia—Country Partnership Framework for FY16–21.” Performance and Learning Review, World Bank, Washington, DC.

World Bank. 2019c. “Honduras—Country Partnership Framework for FY16–20.” Performance and Learning Review, World Bank, Washington, DC.

World Bank. 2019d. *North Macedonia—Country Partnership Framework for January 2019 – June 2023*. Washington, DC: World Bank.

¹ The Global Learning for Adaptive Management initiative, funded by the Department for International Development and the United States Agency for International Development, is a globally networked learning alliance that aims to actively identify, operationalize, and promote rigorous evidence-based approaches to adaptive management.

² In 2013, the Bank Group established a new country engagement model consisting of two instruments: (i) the Systematic Country Diagnostic (SCD) and (ii) the Country Partnership Framework (CPF). The former uses data and analytic methods to help country clients and Bank Group teams identify the most critical constraints to, and opportunities for, reducing poverty and building shared prosperity. The latter determines focus areas for Bank Group support that align with the country’s development agenda, addresses the key constraints and opportunities identified in the SCD, and reflects the comparative advantages of the Bank Group.

³ The 361 indicators were nested within the 144 objectives chosen for review.

⁴ Economic Management Criteria: (i) Monetary and Exchange Rate Policies; (ii) Fiscal Policy; (iii) Debt Policy and Management.

⁵ To assess whether indicators for capturing IFC contributions were present, the team reviewed objectives that had IFC contributions mentioned in the results framework annex of the CPF document or in the text narrative under each objective.

⁶ Operations Policy and Country Services guidance requires that country teams elaborate on the alignment between CPF objectives and Country Development Goals. The guidance does not include any reference to Sustainable Development Goals (SDGs).

⁷ To identify whether CPF indicators were linked to project-based indicators, the team triangulated information from CPFs and follow-up Performance and Learning Reviews,

project documents, and several World Bank databases (Operations Portal, World Bank operations database).

⁸ At the country level, there is significant variation—ranging from 100 percent for Benin and Myanmar to zero percent for Bosnia and Herzegovina, Bulgaria, Chad, and Romania.

⁹ For instance, the Benin Performance and Learning Review for fiscal years 2013–17 reported that the World Bank performed labor market diagnostics to provide evidence-based recommendations on how to implement appropriate job creation programs (World Bank 2017a). According to the Performance and Learning Review, these recommendations will play a key role in guiding the World Bank support to job creation in Benin.

Annex B.1. Coding Template Questions by Sampling Strategy

Question	Sample
Section 1: Outcome Orientation at the Design Stage	
What is the outcome level of Country Partnership Framework (CPF) objectives?	Purposive sample of 29 countries out of a sampling frame of 52 countries.
What is the outcome level of CPF indicators?	Purposive sample of 144 objectives and 361 indicators out of a sampling frame of 265 objectives and 643 indicators.
What is the share of objectives with International Finance Corporation (IFC) and Multilateral Investment Guarantee Agency (MIGA) involvement?	
To what extent do results frameworks include indicators to capture IFC and MIGA contributions under relevant CPF objectives?	Purposive sample of 144 objectives out of 256 objectives
To what extent CPFs discuss collaboration or coordination with other donors for achieving CPF objectives?	Purposive sample of 29 countries out of a sampling frame of 52 countries.
To what extent CPFs discuss Bank's Group contribution to the SDGs?	Purposive sample of 29 countries out of a sampling frame of 52 countries.
Section 2: Measuring Outcomes	
Type of Indicator	Purposive sample of 361 indicators out of a sampling frame of 643 indicators.
What is the share of indicators kept as originally formulated at the PLR stage?	
What is the share of reliable quality CPF indicators?	
Use of evaluation exercises (other than metrics from results framework).	Purposive sample of 29 countries out of a sampling frame of 52 countries.

Question	Sample
Section 3: Adapting Programs and Portfolios	
Rate of CPF objectives and indicator change	Purposive sample of six countries (Chad, Honduras, Indonesia, Peru, Serbia, Vietnam) out of sampling frame of 29 countries. The total number of objectives and indicators reviewed under these country engagements was 60 and 215.
Rationales for CPF and Indicator change	
Section 4: Adaptive Management Practices	
Portfolio health check	Purposive sample of 25 countries out of sampling frame of 29 countries.
Portfolio management actions	
Adapting to context	
Assessing scale-ups or scale-downs	
Assessing cumulative effects	
Balancing risks and outcomes	
Reflecting on comparative advantage	
Section 5: Supporting Client's Outcome Orientation	
Country engagement provides support for statistical capacity building	Purposive sample of 29 countries out of a sampling frame of 52 countries.
Country engagement articulates use of country systems	

Appendix C. Summary of Findings by Country Engagement Type

Country Engagements in Small International Bank for Reconstruction and Development Countries

Table C.1 presents summary findings from the small International Bank for Reconstruction and Development (IBRD) countries of Bosnia and Herzegovina, North Macedonia, Romania, and Serbia.

Table C.1. Evidence from Bosnia and Herzegovina, North Macedonia, Romania, and Serbia

Evidence Type	Evidence
Aiming for outcomes	<ul style="list-style-type: none">• Prioritization: Small portfolio size, which reinforces Global Practice competition for a few lending operations, and the need to be opportunistic to meet specific client demands make it difficult to prioritize interventions so they adhere to a theory of change leading to higher-level outcomes. Finding a niche (World Bank Group value added), and carving out opportunities for a programmatic approach is key. Strong country leadership and the use of filters have proven critical both in Romania and in the Western Balkans.• One World Bank Group: IFC has a lot of competitors in small IBRD countries and is relatively more expensive than others. It thus needs to be even more opportunistic in this context. In EU accession countries, specifically, IFC’s advisory portfolio can make a substantive contribution to help the country adhere to EU regulation. However, the multiplicity of strategic documents from within the World Bank Group, from the government, and donors can create difficulties. The preparation of country strategies backed by data and analytics presents an opportunity to enhance the internal coherence of IFC interventions.

Evidence Type	Evidence
Capturing outcomes	<ul style="list-style-type: none"> • Institutional capacity building versus direct services: Institutional capacity building and fostering governments' ability to implement policy reforms are prevalent modes of influence in small IBRD countries, especially through instruments such as RASs, IFC advisory and ASA. Yet, the Bank Group still does not have robust theories of change that link these types of Bank Group contributions to country outcomes. <p>Availability of data on key sector outcomes is not a major issue; however, the Bank Group's results system does not tap into available country data because of the attribution and time-boundedness requirements of the model.</p> <p>The most meaningful contributions to country outcomes in small IBRD countries are institutional strengthening and building agencies' capacity to implement policies. They do not fit well in the Bank Group's results reporting system, especially when delivered through analytical and technical assistance instruments. Thus, they are not well tracked, evaluated, and learned from. The Romania CMU is investing in developing tools and approaches to assess the value of the RAS portfolio. Other CMUs are trying to ensure that ASAs are tied to CPF objectives and have an indicator in the CPF results framework. Country teams find it hard to do justice to the actual change they are creating through indirect pathways.</p>
Adaptive management	<p>Changes in client demand are the main factor influencing adaptation in small IBRD countries. Given the rapid government turnaround, country teams need to be flexible to adjust portfolios to maintain a Bank Group footprint and influence. Regular health checks with clients and a well-trained and coordinated World Bank portfolio team have proven essential, especially in multicountry CMUs such as the Western Balkans.</p>

Evidence Type	Evidence
Client engagement	Governments' monitoring and evaluation functions are nascent in small IBRDs. Some have invested in results-based management, including with World Bank support, such as in Serbia, but the culture of evaluation and results remains weak. Results-based financing such as IPF with disbursement-linked indicators combined with technical assistance have proven useful for clients. Clients take an active role in ensuring the health of the portfolio from a delivery and disbursement point of view.

Note: ASA = Advisory Services and Analytics; CMU = Country Management Unit; CPF = Country Partnership Framework; EU = European Union; IBRD = International Bank for Reconstruction and Development; IFC = International Finance Corporation; IPF = investment policy financing; RAS = reimbursable advisory services.

Country Engagements in Large IBRD Countries

Table C.2 presents summary findings from the small IBRD countries of Colombia, Indonesia, Morocco, Peru, Sri Lanka, and Vietnam.

Table C.2. Evidence from Colombia, Indonesia, Morocco, Peru, Sri Lanka, and Vietnam

Evidence Type	Evidence
Aiming for outcomes	<ul style="list-style-type: none">• Prioritization: Client demand and internal incentives to lend substantially drive Bank Group priorities. In large IBRD countries, governments usually have many financing alternatives to Bank Group lending and investments. The Bank Group is a relatively small lender, despite sizable portfolios, and faces competition from other development banks and from the market, which can offer cheaper financing. Thus, country teams also need to be opportunistic in selecting projects and need to offer clear comparative advantage through technical

Evidence Type	Evidence
Capturing outcomes	<p>skills or global knowledge. This can lead to broad CPFs, so that any project will be "relevant," and country teams with reduced leverage to push priorities. The difficulty of maintaining a medium-term trajectory of change depends on the nature of the client's institutional arrangements. In countries with strong planning culture (for example, Colombia, Vietnam, Morocco), it is easier to maintain a medium- to long-term view of outcomes and trajectories of change. In countries with strong political cycles (Philippines, Peru, Sri Lanka) and less of a planning structure, development needs may be separate from government priorities, and country teams have difficulty in maintaining a focus on long-term outcomes.</p>
	<ul style="list-style-type: none"> • Institutional capacity building versus direct services: Bank Group portfolios are more mixed in large IBRD countries, with a combination of direct and indirect pathways to outcome. However, World Bank Group financing is only a small fraction of government and private sector financing in direct service delivery. Country teams working in large IBRD countries find that the most important contributions they can make to country outcomes come from their catalytic or demonstration effect on improving or influencing the efficiency of government spending, not particularly the direct effects of World Bank projects. • Measurability capabilities and data availability are less of an issue than in IDA FCV countries. The issue is the lack of CMU

Evidence Type	Evidence
	<p>incentive or resources for drawing together that data.</p> <ul style="list-style-type: none">• Demonstration effects, institutional capacity building, analytics-supported policy dialogue, and advocacy and convening work are some of the main contributions of the Bank Group to country outcomes in large IBRD countries but are not captured well in Results Frameworks, which primarily draw from investment policy financing data and thus are biased toward direct services improvements.• An added complexity in large IBRD countries is that the Bank Group is often engaging subnationally. What might not work in northern Vietnam might be possible in southern Vietnam. Therefore, monitoring national or sector outcomes or others can also be misleading.
Adaptive management	<p>Most adaptive management decisions are not driven by results; they are driven by client demand, country context, and crises or shocks. Client demand is by far the strongest factor in large IBRD countries. Large IBRD countries are usually more resilient and less vulnerable to shocks than IDA FCV, hence, adaptive management is mostly about managing the forward portfolio. Aide-mémoire, Memorandums of Understanding with incoming finance ministers, and Joint Reviews and feedback prove useful in countries with high turnover in government (for example, Peru).</p>
Client engagement	<p>Governments have clear and specific demands and are heavily involved at the CPF design stage, less so later in the cycle. Government staff turnover quickly and are usually accountable to</p>

Evidence Type	Evidence
	their governments for output delivery. In some cases, they have detailed monitoring and reporting systems, but these remain output-based. Country staff suggest improvements in their data or monitoring and evaluation management would depend on country political culture but also mention the opportunities that remain in refocusing evaluative systems on clients, particularly in countries with very capable clients.

Note: ASA = Advisory Services and Analytics; CMU = Country Management Unit; CPF = Country Partnership Framework; FCV = fragility, conflict, and violence; IBRD = International Bank for Reconstruction and Development; IDA = International Development Association; IFC = International Finance Corporation.

Country Engagements in International Development Association Countries

Table C.3 presents summary findings from the International Development Association (IDA) countries of Benin, Burkina Faso, Ghana, Honduras, Kyrgyz Republic, Nicaragua, Tajikistan, and Zambia.

Table C.3. Evidence from Benin, Burkina Faso, Ghana, Honduras, Kyrgyz Republic, Nicaragua, Tajikistan, and Zambia

Evidence Type	Evidence
Aiming for outcomes	<ul style="list-style-type: none"> • Prioritization: In IDA countries, the World Bank Group's concessional finance means that it has much more leverage (relative to IBRD countries) in pushing clients to adopt Bank Group priorities. Yet, significant internal competition between Global Practices for program space makes it difficult to achieve a strategic focus. Geographic location filters, higher project overlap/interconnectedness, and higher Country Director ownership throughout the whole country engagement have proven useful for increasing prioritization.

Evidence Type	Evidence
Capturing outcomes	<ul style="list-style-type: none"> One World Bank Group: IFC's involvement in the design of the country engagement in IDA countries is marginal, although the Upstream work of IFC and its new country products are starting to shift this. IDA countries are considered priority countries for MIGA, whose involvement in the country engagement cycle has improved over time. Institutional capacity building versus direct services: In IDA countries portfolios are more balanced between direct investments in service delivery and indirect pathways to strengthen government's capacity to deliver. Measurability is still an issue in IDA countries, with frequent data gaps, especially at the local level, which contributes to an overreliance on project data in Bank Group results frameworks. In IDA countries there is more effort to coordinate with other development partners, including to set priorities and division of labor. However, most of the time this coordination does not extend to monitoring and evaluation. Donors retain their own system and government counterparts have to comply with multiple reporting requirements, which does not favor the establishment of a country system for monitoring and evaluation. Results frameworks fail to capture other important intangible aspects of IDA investments such as capacity building, policy dialogues, or scale-up or demonstration effects. CPF results frameworks also aim for alignment with the IDA results frameworks, which is a driving force of indicator selection. Teams suggest more qualitative monitoring and the use of already existing World Bank tools such as the Survey of Well-Being via Instant and Frequent Tracking, the Iterative Beneficiary Monitoring, and "Listening to Tajikistan," as examples of high-frequency surveys to monitor life satisfaction, fragility, and shocks during country engagements.

Evidence Type	Evidence
Adaptive management	Adaptations to country engagements are common in IDA, but few are driven by information coming from the results framework. Other factors usually carry more weight. These include (i) responding to changes in country demands and political economy; (ii) incorporating corporate priorities, particularly IDA commitments; and (iii) recalibrating donor coordination and establishing a division of labor.
Client engagement	The level of attention to outcome measurement varies significantly across and within IDA countries. The World Bank contributes to strengthening government statistical capacity in IDA countries via stand-alone projects and through specific activities built into projects. In some countries there has been significant improvement in data availability, but the use of data remains weak.

Note: IBRD = International Bank for Reconstruction and Development; IDA = International Development Association; IFC = International Finance Corporation; MIGA = Multilateral Investment Guarantee Agency.

Country Engagements in Fragility, Conflict, and Violence Countries

Table C.4 presents summary findings from the fragility, conflict, and violence–categorized countries of Afghanistan, Chad, Haiti, Kosovo, Papua New Guinea, and the Solomon Islands.

Table C.4. Evidence from Afghanistan, Chad, Haiti, Kosovo, Papua New Guinea, and the Solomon Islands

Evidence Type	Evidence
Aiming for outcomes	<ul style="list-style-type: none"> • Prioritization: FCV countries present a particular challenge because teams need to balance long-term development agendas with shorter-term shocks or crisis response, some of which are of humanitarian nature. Political economy challenges are severe, and flexibility is required to take advantage of opportunities as they arise based on political economy (for example, a supportive minister). Very high turnover in government officials can make long-term planning and sustained agendas difficult. Coordination with

Evidence Type	Evidence
	<p>other donors and partnerships are all the more important, because the Bank Group does not have a comparative advantage in addressing all of the complex challenges posed by FCV, including security issues. Joint donor planning mechanisms have been helpful in coordinating priorities for policy reforms notably (for example, Pacific Islands and Afghanistan). In some FCV countries, such as Afghanistan, Trust Funded activities are a substantial part of the World Bank's portfolio and have their own strategy processes, and the Bank Group's country strategy can be designed to follow and align with those processes.</p>
	<ul style="list-style-type: none">• Institutional capacity building versus direct service: Contributions to stability, capacity building and basic state-building and service delivery are prevalent channels. For example, in the Solomon Islands, World Bank projects on rural development and youth employment have played a contributing role to peace and stability through project activities that build dispute resolution mechanisms, increasing rural incomes, improving service delivery, and providing jobs. Yet, it is hard to have big sector impacts except over the long term. There is an institutional culture that prioritizes delivery over results, but staff argue that the sheer difficulty of operational delivery means that it is appropriate for substantial effort to be focused here.
Capturing outcomes	<ul style="list-style-type: none">• Data availability is a critical issue. The scarcity and low quality of data are major problems in FCV countries. As a result, country teams depend on project data, which can rely on third-party monitoring or remote and extenuating factors associated with fragility (that vary by sectors and timeframes) hinder measurability.• Development outcomes in FCV countries can be achieved through maintaining development gains, or by preventing them from slipping backward.
Adaptive management	<p>Because of rapidly changing circumstances in the field and political economy, high levels of conflict and insecurity, and the capacity constraints of counterparts, frequent adaptations are necessary in</p>

Evidence Type	Evidence
	FCV countries. Country teams must have a higher tolerance for risk and expect failures. Yet, this remains difficult in a results system that requires forecasting and justifying departure from what was planned.
Client engagement	Client involvement in outcome measurement and management is limited to the project level. Other than this, in FCV countries, clients have either no system or very limited monitoring and evaluation systems for tracking development progress. There are few signs of World Bank support for capacity building on monitoring and evaluation at the country level; those efforts that exist are usually about improving monitoring and evaluation capacity for specific projects (monitoring systems) not on evaluation or learning.

Source: Independent Evaluation Group.

Note: FCV = fragility, conflict, and violence.

Appendix D. Options for a Prototype Monitoring, Evaluation, and Learning Plan

This appendix seeks to provide more detail on what a monitoring, evaluation, and learning plan might look like and how it could improve some of the limitations of the World Bank Group's country-level results system. The vision laid out here is one possible means of reforming this system to ensure that it accurately captures the Bank Group's contribution to country outcomes and usefully informs decision-making on country engagements. The vision here is not intended to be prescriptive but rather to respond to Bank Group management requests for greater granularity in a proposed reform and provide ideas that could be built on further.

Objectives of a Monitoring, Evaluation, and Learning Plan

At the country-level A monitoring, evaluation, and learning (MEL) plan is a tool for country teams, Regions, management, and the Board of Executive Directors for planning, managing, and governing three processes:

- **Monitoring** strategic progress, portfolio performance, operational risks and contexts, and the development hypotheses (programmatic assumptions) that underly the CPF
- **Evaluating** the results and performance of the country program and the contribution of the Bank Group to country outcomes
- **Learning** and adapting portfolios, programs, and interventions based on evidence.

MEL plans are intended to be a **country-level complement to the existing project-level results system**, which is anchored on results frameworks.

IEG’s evaluation on outcome orientation at the country level does not suggest changes to the project-level results system.

Content and Intended Use

Table D.1 summarizes the types of MEL approaches, their intended use, and associated evidence needs.

The MEL plan would address some of the limitations of the current systems as laid out in table D.2. However, the MEL plan approach would not be sufficient to address the range of challenges laid out in the report. Other issues such as incentives or selectivity also deserve attention.

Table D.1. Monitoring, Evaluation, and Learning Approaches and Their Intended Use

	Approach	Intended Use	Type of Evidence Needed
Monitoring	Tracking country outcomes	Monitor progress toward key country outcomes that the Bank Group seeks to influence across CPF cycles. Help management and the Board identify areas where progress is going in the right direction and those where it is not and might warrant further attention.	A selection of country outcome-level indicators that would not change drastically with each CPF cycle. Some could be aligned with SDGs. No assumption of attribution to Bank Group interventions and no need for specific targets to be met.

	Approach	Intended Use	Type of Evidence Needed
Evaluating	Tracking corporate goals at the country level	Ensure that corporate priorities are being addressed by the program and provide the data necessary for corporate reporting mechanisms such as the corporate scorecard and for IDA.	Key performance indicators to track corporate commitments at the country level, such as gender equality, climate change, and so on, as well as key indicators from the IDA results framework.
	Monitoring country portfolio performance	Help country teams with operational management and delivery.	Information from the unchanged project-level monitoring and results system.
	Evaluating the contribution of the Bank Group to country outcomes	Inform teams and the Board on how Bank Group engagements influence country outcomes. This could include assessing the combined effects of Bank Group interventions over time in countries where multiple instruments are used to contribute to key outcomes. This could also include assessing the Bank Group's contribution to indirect pathways (market creation, institution building, and so on.) through analytical work, technical assistance, capacity building, policy dialogue, market creation, and convening work.	Mixed-methods studies that focus on contribution rather than attribution could help provide a narrative that informs teams and the Board on how Bank Group engagements connect to country outcomes. A range of evaluation methods can be used to collect and analyze this information.

	Approach	Intended Use	Type of Evidence Needed
Learning	Assessing the results of country portfolios	Provide an overall performance proxy that mimics most of the accountability elements of the existing system (which relies heavily on project data) but without the need for a separate and costly country-level results framework.	A synthesis of evidence from project ratings and assessments.
	Assessing the satisfaction of key stakeholders	Provide qualitative feedback on the effectiveness of the country engagement and elements that have been more or less effective.	These and other qualitative information sources are particularly important for understanding the relevance of Bank Group's engagement.
	Stocktaking and Mid-Term reviews	Occasional stocktaking exercises when circumstances warrant (for example, Mid-Term review; change in government; major shock) to consider shifting priorities, necessary adaptations. These stocktaking would go to the Board only if a major shift in the program is warranted.	A synthesis of existing data and information available from the monitoring and evaluation activities to inform specific decisions or adaptation.
	Facilitating reflection sessions	Encourage and facilitate the use of evaluative evidence for decisions and learning.	Create platforms for person-to-person learning and information sharing that can be more effective than learning from formal reports or codified lessons.
	Filling knowledge gaps	Bridging knowledge gaps around specific CPF objectives or country outcomes.	A set of questions or knowledge gaps and how they can be filled through research, analytics, evaluation, and so on.

Note: CPF = Country Partnership Framework.

Table D.2. How a MEL Plan Addresses Limitations of the Current Country Results System

Intended Use	Current System	MEL Plan
Capturing contribution to country outcomes	Though CPF objectives are framed in terms of outcomes, their results frameworks' indicators do not capture well these objectives, and about half the time are framed below the level of the objectives they seek to capture, particularly because of the expectation that indicators are quantified and clearly attributable to Bank Group interventions.	Rather than focusing on quantitative and attributable effects of Bank Group interventions, the MEL plan and its terminal evaluation would report on the contribution of the Bank Group to key CPF objectives and their influence on country outcomes. The design of evaluations would embrace principles from contribution analysis and incorporate qualitative data and evidence.
Capturing the contribution of various and multiple instruments, and of indirect development pathways	Most country strategy indicators come from lending projects, and do not properly capture the contribution of advisory services, analytics, policy dialogues, or stakeholder convening. IFC's and MIGA's contributions are also not well captured. CLRs rarely captured complementarities across instruments so are not able to establish whether the Bank Group's contribution is more than the sum of its parts.	By removing the expectation that effects of the full CPF be captured, a MEL plan that embraced selectivity could allow deeper inquiry into critical areas of interest. These could include specific evaluation exercises on the influence of critical nonlending activities and indirect development pathways. A well-selected set of study would assess the contribution of packages of interventions, including IFC and MIGA, to key country outcomes pursued over the long term by the Bank Group. Not restricting the evidence base to quantifiable indicators would provide opportunities for teams to leverage and tap into a wider range of evidence, including surveys, narrative accounts, and better leverage existing evidence from the project-level results systems in the three institutions.

Intended Use	Current System	MEL Plan
Capturing contributions over time	Country strategies lay out future objectives, but results frameworks primarily capture past operations and do not capture medium- to long-term effects.	Monitoring would take place across CPF cycles and evaluation exercises would be timed to be able to capture the contribution of long-term effects of programs.
Reporting on corporate priorities	The current results framework-based system does not facilitate reporting on corporate priority at the country level and these are carried out as separate exercises.	The monitoring plan would have a set of indicators dedicated to reporting on corporate priorities and IDA Results Frameworks. This would make it easier for Regions to report to Senior Management and the Board on corporate commitments.
Adaptive Management	<p>The country-level results system does not effectively support adaptive management. Decisions on balancing risks and rewards are also not well informed by the results system.</p> <p>Much of the time and effort put into PLRs is taken up by revising the results frameworks, which can crowd out collective reflection.</p>	<p>Country teams would shape the Monitoring and Evaluation plan in part so that they fit their own adaptive management and learning needs. Less time spent on reporting and adapting the results framework and more ownership in the system would encourage collective reflection.</p> <p>Accountability and incentive mechanisms that would encourage evaluative thinking and learning would support the use of data and evidence that are more fit for purpose.</p>
Knowledge of Bank Group contribution to high level outcomes	<p>The country-level results system prioritizes accountability rather than knowledge generation.</p> <p>Some country teams carry out other evaluative exercises to assess achievements of outcomes, but these are rare and overshadowed by the results system's emphasis on</p>	<p>Increased knowledge of the Bank Group's contribution to country outcomes would come from the combination of monitoring country outcomes that the Bank Group seeks to influence over the long term and specific evaluations to assess the contribution of the World Bank to such outcomes for accountability.</p> <p>The multiple evaluative studies that already take place at the project or country level</p>

Intended Use	Current System	MEL Plan
	targets and indicators, which also consumes the time of monitoring and evaluation staff.	would be better recognized and their content used for outcome management because they would be part and parcel of the MEL plan.
Organization learning	Requirements to update and fill out CPF results frameworks and prepare PLRs and CLRs as Board deliverables tend to crowd out learning and reflection. Lessons captured in the PLR and CLR are considered too generic to be useful and there is little evidence of these products being used beyond reporting purposes.	The MEL would place greater emphasis on evaluation that meets the needs of country teams and supports their learning. The Learning Plan would help Country Teams better manage their learning and have dedicated reflection sessions to grapple with the findings of the monitoring and evaluation activities. Relaxing requirements for Board deliverables for midterm reporting would reduce the compliance burden on teams.
Incentives	Teams engage with the system out of compliance, not out of a sense of utility or interest.	By placing greater emphasis on the needs and uses of country teams, the MEL would improve ownership of and interest in the system by these teams. Dedicated time and forums at all levels of the organization to discuss the findings from monitoring and evaluation activities would enhance interest and use of the results system.

Note: CLR = Completion and Learning Review; CPF = Country Partnership Framework; MEL = monitoring, evaluation, and learning; PLR = Performance and Learning Review.

Section 1: Monitoring Plan

A monitoring plan would describe how the country team will monitor the country engagement’s performance, changes to the country outcomes that the Bank Group seeks to influence, and contextual risks that are relevant to specific CPF Objectives. A monitoring plan could encompass the following:

- **(Across CPF cycles) Monitoring progress toward key country outcomes:** a selection of country outcome–level indicators that the World Bank Group is trying to influence across CPF cycles. The set of selected indicators would not change drastically with each CPF cycle unless there were fundamental shifts in CPF strategy (which is rare), and there would also be no assumption of attribution to Bank Group interventions and no need for specific targets to be met.
 - Indicators might include human capital index, education enrollment, electricity access, prevalence of stunting, share of population using safely managed water services, proportion of youth not in education or employment, natural capital depletion, and so on.).
 - These indicators could align with SDG indicators in some cases. Data for indicators could come from government data, new monitoring efforts, country-specific or global World Bank analytical work (such as the Human Capital Index or work on natural capital accounting) but would be unlikely to come from specific World Bank project indicators.¹
- **(Within CPF cycle) Monitoring progress toward key corporate objectives:** This would include key performance indicators to track corporate commitments at the country level, such as gender equality, climate change, and so on, as well as key indicators from the IDA results framework. This would help streamline and facilitate corporate reporting on key goals and priorities.
- **(Periodic) Monitoring of outcome risks:** a list of relevant indicators for monitoring country-level risks that could affect progress toward specific outcomes, this could also include existing World Bank and IFC/MIGA tools to monitor conflict or fragility. There could be synergies with the updated risk management framework currently being rolled out.

- **(Routine) Monitoring of portfolio performance:** a streamlined set of indicators that country teams use for Health Check of the portfolio, on a regular basis. This would be extracting information from the unchanged project-level results system.
 - For example, disbursement rate, problem projects, financial return for IFC, and so on.

Intended Use

Monitoring country outcomes would help track progress toward country outcomes that the Bank Group is seeking to influence and would support oversight and governance. It would be a complement to the existing monitoring of results frameworks at the project level. It could help senior management and Board members to identify areas where progress is going in the right direction and those where it is not and might warrant further attention. It could help meet the Board's demand for information on high-level outcomes. Harmonized monitoring with other development partners could help focus attention and incentives around priority outcomes. Tracking SDG indicators could help show to the Board how the Bank Group program connects to the SDGs.

- Monitoring corporate objectives would help ensure that corporate priorities were being addressed by the program and provide the data necessary for corporate reporting mechanisms such as the corporate scorecard and for IDA.
- Monitoring of outcome risks would support adaptive management by providing early warning of trends that might affect progress toward outcomes and serve as a trigger for country teams to identify options for proactive risk management through risk mitigation or exposure reduction.

- Monitoring of portfolio performance, as occurs already, helps country teams with operational management and delivery. This is an important function that should continue.

Section 2: Evaluation Plan

At the country level, an evaluation plan would describe evaluations planned to be carried out during the strategy period. The plan would be a complement to the existing project-level self-evaluation system and focus on activities that would bridge knowledge gaps that are not covered by project-level assessments. The plan would define the team's criteria for selecting evaluation activities, the evaluation's purpose and expected use, possible evaluation questions, and the evaluation modality (internal or commissioned out). The evaluation plan would also include a terminal synthetic report that would be intended to inform development of the next strategy. This report would synthesize available information on the Bank Group's performance, including the results from evaluative activities under the MEL plan, portfolio information from project results systems, and stakeholder feedback surveys or narrative accounts. The evaluation plan would lay out any planned particular areas of focus for this terminal assessment. The plan could include the following:

- A synthesis of evidence from project evaluations and ratings, which could include information stemming from World Bank Implementation Completion and Results Report Reviews, IFC's Expanded Project Supervision Report, Project Completion Reports, information from Anticipated Impact Measuring and Monitoring, and MIGA's IMPACT assessments.
- Planned stakeholder surveys or narrative accounts; for example,
 - Client or beneficiary survey
 - Other citizen engagement, civil society, and nonstate actor engagement tools at the country level

- **Planned evaluations with specific data collection:**
 - Of interventions or projects: (for example, [quasi] experimental impact evaluations or mixed-methods assessment)
 - Of contributions to specific outcomes over multiple CPF cycle of packages of instruments; for example,
 - Operations Policy and Country Services pilots
 - RAS evaluation in Romania
 - (Hypothetical) contribution of Bank Group to state-owned enterprise reforms in Serbia
 - (Hypothetical) contribution of Bank Group to SDG 4 (quality education) in Vietnam
 - (Hypothetical) contribution of Bank Group to improving social contract in Tunisia
 - Joint evaluation with other development partners of shared approaches; for example,
 - Joint evaluation of budget support to Ghana
 - Joint evaluation of hurricane response in Haiti
- For each planned evaluation, the plan could lay out the purpose and expected use of the evaluation, the decision points that require evidence from evaluation, possible evaluation questions, whether the evaluation will be conducted internally (for example, by country teams, by poverty GP, Development Economics, Regional Development Economics units, and so on.) or externally (outsourced to consultant or consulting firm); estimated budget, and start/completion date.

Intended Use

A terminal synthesis report can support accountability and learning functions. Timing it to precede rather than accompany the CPF would enable it to meaningfully inform CPF preparation. Enabling it to focus evaluative efforts on specific areas of interest would allow deeper treatment of priority issues, rather than wide but superficial coverage. Using a wider set of methods and a focus on contribution rather than attribution could help provide a narrative that informs teams and the Board on how Bank Group engagements connect to country outcomes.

A synthesis of evidence from project evaluations and ratings would be a low-cost means of providing an overall performance proxy that mimics most of the accountability elements of the existing system (which relies heavily on project data) but without the need for a separate and costly country-level results framework.

Planned stakeholder surveys or narrative accounts can provide valuable qualitative feedback on the effectiveness of the country engagement and elements that have been more or less effective. These and other qualitative information sources are particularly important for understanding Bank Group contributions to indirect development effects, including analytical work, technical assistance, capacity building, policy dialogue, market creation, and convening work.

Specific planned evaluations can create platforms for gathering evidence around particular issues of strategic importance, especially when there are known decision points that could be informed by such evidence. For example, a deeper evaluation on a particular modality would be useful for informing decisions on whether to continue, scale up, or shift away from a particular modality.

An evaluation plan is intended to be context-specific and allow the Bank Group to best capture its contribution to country outcomes. Evaluation plans are meant to be adapted to the nature and composition of the country

portfolio (for example, heavy reliance on advisory services and analytics, policy lending versus primarily investment lending); the development pathways that are privileged (for example, indirect versus direct) and the country context (for example, data rich versus data poor).

For example, in countries where the World Bank primarily pursues country outcomes through indirect pathways to strengthen institutions, improve the quality of governments' spending, or catalyze private sector development and uses primarily policy lending, technical assistance, or IFC Upstream approaches, the evaluation plan would rely less on collecting information on standard sector indicators and more on dedicated mixed-methods evaluations that would trace progress toward institutional change. Conversely, in countries where the Bank Group primarily works through direct investments in sectors where outcome data can be routinely collected and where sectoral Management Information Systems can effectively track progress, indicator frameworks can be prominently featured in the evaluation approach.

Section 3: Learning Plan

A learning plan would identify knowledge gaps that the country team intends to fill during the CPF period. It would also plan activities for country teams to periodically reflect on certain activities, processes, and findings related to evaluations, information monitoring, and stakeholder engagements. The MEL could include one or more midterm reviews, where the country team engages in an exercise to collectively consider whether progress toward objectives is on the right track or not, and identify lessons or opportunities for changing direction.

This plan would lay out the country teams' approach to learning and adapting, what types of processes or activities will help the team ensure that the program is coordinated (with Practice Groups, across Bank Group institutions, with partners, and with clients), that it is grounded in evidence,

and adjusted as necessary to remain effective. Such a plan could include the following types of activities:

- Identifying key knowledge gaps around specific CPF objectives or country outcomes and how they can be filled through research, analytics, evaluation, and so on.
- Planning for facilitated reflection sessions based on collected or synthesized evidence to inform decisions and adaptation. Some of these sessions would be internal, others would be joint with IFC and MIGA, some might be within the context of “country platforms,” or with the clients only, and so on.
- Occasional stocktaking exercises when circumstances warrant it (for example, Mid-term Review; change in government; major shock) to consider shifting priorities, necessary adaptations. These stocktakings would go to the Board only if a major shift in the program is warranted (light PLR).

Intended Use

A learning plan would help encourage country teams to learn from evidence, and to create platforms for person-to-person learning and information that can be more effective than learning from formal reports or codified lessons. They would also enable oversight from senior management and the Board, who would be able to send signals that learning is a priority and observe whether country management teams are establishing environments that support learning. These learning plans are not meant to be standardized across the Bank Group but should be adapted to the current practices and team culture of country teams that support continuous learning and adaptation.

¹ Exceptions might occur for example in very small countries, for policy lending operations, or for sector-wide approaches, where the project/program truly engages on key country sectoral outcomes.



IEG
INDEPENDENT
EVALUATION GROUP

WORLD BANK GROUP
World Bank • IFC • MIGA



The World Bank
1818 H Street NW
Washington, DC 20433