



Program-for-Results

An Early-Stage Assessment of the Process and Effects
of a New Lending Instrument



IEG
INDEPENDENT
EVALUATION GROUP

WORLD BANK GROUP
World Bank • IFC • MIGA





Program-for-Results

An Early-Stage Assessment of the Process and Effects
of a New Lending Instrument

An Independent Evaluation

© 2016 International Bank for Reconstruction and Development / The World Bank
1818 H Street NW
Washington DC 20433
Telephone: 202-473-1000
Internet: www.worldbank.org

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.

Suggested citation: World Bank. 2016. *Program-for-Results: An Early-Stage Assessment of the Process and Effects of a New Lending Instrument*. Washington, DC: World Bank.
<https://hubs.worldbank.org/docs/imagebank/pages/docprofile.aspx?nodeid=26943793>.

Contents

ABBREVIATIONS	V
ACKNOWLEDGMENTS	VII
OVERVIEW.....	IX
MANAGEMENT RESPONSE	XVIII
MANAGEMENT ACTION RECORD	XXIII
CHAIRPERSON'S SUMMARY: COMMITTEE ON DEVELOPMENT EFFECTIVENESS	XXIX
1. THE PROGRAM-FOR-RESULTS INSTRUMENT AND THE EVALUATION APPROACH	1
The PforR Instrument	1
The Evaluation.....	7
2. EXPERIENCE WITH THE PFORR PROGRAM.....	11
Growth and Diversification of the PforR Portfolio.....	11
Design Coherence.....	16
Emerging Messages.....	22
3. THE QUALITY OF PROGRAM ASSESSMENTS.....	25
Technical Assessments.....	25
Fiduciary Assessments.....	28
Political Economy Aspects.....	31
Environmental and Social Systems	31
Emerging Messages.....	34
4. IDENTIFICATION, ASSESSMENT, AND MITIGATION OF RISKS	36
Appropriateness of the Identification and Assessment of Program Risks	37
Appropriateness of Risk-Based Exclusions	39
Risk Management Implementation Experience	41
Emerging Findings.....	43
5. STRENGTHENING NATIONAL SYSTEMS.....	45

CONTENTS

The Identification and Assessment of Capacity-Building Measures	45
Links with IPFs.....	50
Links with Development Partners and Technical Assistance.....	51
Implementation Experience with Capacity-Building Measures	52
Emerging Findings.....	53
6. FINDINGS AND RECOMMENDATIONS	55
Main Findings	55
Recommendations.....	60
Boxes	
Box 2.1. Scope and Results Supported by the Ethiopia Health MDGs Program-for-Results	18
Box 2.2. Example of a DLI Set Different from the Results Framework	19
Box 2.3. Vietnam—DLIs in the Results-Based National Urban Development Program in the Northern Mountains Region.....	20
Box 3.1. Good Practice in Costing: Ethiopia Health Millennium Development Goals PforR	27
Box 4.1. Compensation for Land Acquisition—Vietnam Northern Mountains Urban Development Program.....	43
Box 5.1. Support to Municipal Urban Development in Uganda	46
Box 5.2. Mozambique Public Financial Management Capacity Support.....	47
Box 5.3. Morocco National Initiative for Human Development	52
Box 6.1. Key Aspects for PforR Ex-Post Completion Reports	56
Tables	
Table 1.1. Complementary Lending Instruments	5
Table 2.1. IBRD/IDA Lending, Total and for PforR Operations, Fiscal 2012–16.....	11
Table 4.1. Risk Ratings in PforR PADs.....	37
Figures	
Figure 2.1. The PforR Portfolio	12
Figure 2.2. PforR Commitment versus Disbursement.....	14
Appendixes	
APPENDIX A. METHODOLOGY NOTE FOR REVIEW OF THE BANK'S PFORR PORTFOLIO	63
APPENDIX B. RESULTS-BASED PROGRAMS BEFORE PFORR	71
APPENDIX C. PFORR PORTFOLIO	89
APPENDIX D. DISBURSEMENT-LINKED INDICATORS.....	91
APPENDIX E. FIDUCIARY ASSESSMENTS.....	109
APPENDIX F. ENVIRONMENTAL AND SOCIAL SYSTEMS.....	116
BIBLIOGRAPHY	127

Abbreviations

ACG	Anti-Corruption Guidelines (World Bank)
ADB	Asian Development Bank
APL	Adaptable Program Loan
BP	Bank Procedure
CBID	capacity building and institutional development
CCT	conditional cash transfer
CPAR	Country Procurement Assessment Report
DFID	U.K. Department for International Development
DLI	disbursement-linked indicator
DPF	development policy finance
DPO	development policy operation
EEP	eligible expenditure program
EMWF	East Meets West Foundation
ESP	Education Sector Plan
ESPIG	Education Sector Program Implementation Grant
ESSA	environmental and social systems assessment
EU	European Union
GPE	Global Partnership for Education
GPOBA	Global Partnership on Output-Based Aid
HMIS	health management information system
HNP	Health, Nutrition, and Population
IBRD	International Bank for Reconstruction and Development
IDA	International Development Association
IDB	Inter-American Development Bank
IEG	Independent Evaluation Group
IFSA	Integrated Fiduciary Systems Assessments
INDH	National Initiative for Human Development
IPF	investment policy finance
ISR	Implementation Status and Results Report
M&E	monitoring and evaluation
MCA	maximum country allocation
MDG	Millennium Development Goal
OBA	output-based aid
OBD	output-based disbursement
OECD	Organisation for Economic Co-operation and Development
OP	Operational Policy
OPCS	Operations Policy and Country Services
PAD	program appraisal document
PAP	Program Action Plan
PAR	Performance Assessment Report
PDL	performance-driven loans
PDO	program development objective
PEFA	Public Expenditure and Financial Accountability
PFM	public financial management
PforR	Program-for-Results
PFSA	Pharmaceutical Fund and Supply Agency
PMCHIP	Provincial Maternal-Child Health Investment Project

ABBREVIATIONS

RBA	results-based aid
RBF	results-based financing
RBL	results-based lending
RWSS	rural water supply and sanitation
SIL	Sector Investment Loan
SWAP	sectorwide approach
USAID	U.S. Agency for International Development
VAT	value added tax
WSS	water supply and sanitation

All monetary amounts are in U.S. dollars unless otherwise indicated.

Acknowledgments

This report of the Independent Evaluation Group (IEG) was prepared by a team led by Ismail Arslan, task team leader, with major contributions and background papers provided by Nils Fostvedt, Monika Huppi, Xue Li, Andres Liebenthal, Albert Martinez, David Mead, and Clay Wescott. Aimee Niane and Carla F. Chacaltana were responsible for all administrative aspects of the evaluation. William Hurlbut and Caroline McEuen provided editorial support.

The team gratefully acknowledges the support of task teams and task team leaders for Program-for-Results operations throughout the World Bank. Colleagues in Operations Policy and Country Services (OPCS) at the World Bank were also very helpful, as were the many IEG staff and managers who offered ideas and perspectives at various points during the process. The team also acknowledges the advice received from Victoria Alexeeva, Brenda Barbour, Joy Behrens, Giuseppe Iarossi, Javier Lanza, Midori Makino, Bahar Salimova, Mark Sundberg, and Fang Xu.

Peer reviewers were Anis Dani (former IEG lead evaluation officer), Alison Evans (chief commissioner of the Independent Commission for Aid Impact, U.K.), Louise Fox (Woodrow Wilson International Center; visiting professor at the University of California at Berkeley; and former World Bank staff member), Jeff Gutman (Brookings Institution and former OPCS vice-president), and Joanne Salop (former OPCS vice-president).

The evaluation was conducted under the guidance of Nick York, director; Marie Gaarder, manager; and Caroline Heider, director-general, evaluation.

Overview

The World Bank introduced the Program-for-Results (PforR) lending instrument in 2012. As foreseen at the time, the Bank's Operations Policy and Country Services department reviewed the first two years of operation in 2015. The Board of Executive Directors subsequently requested an early Independent Evaluation Group (IEG) evaluation of the instrument. This process evaluation responds to that request and has two objectives: to assess experience with design and implementation of PforR operations and to identify lessons and recommendations to strengthen this new lending instrument.

In addition to using the findings of the *Two-Year Review* (World Bank 2015e), the evaluation is based on reviews of all PforR projects approved through the end of fiscal 2015, a total of 27 projects. This was supplemented by field visits to 11 countries.

At this point, none of the operations has been completed, so there are no IEG-validated data on outcomes, and only partial data on disbursements, outputs, and results. The evaluation assesses program design and implementation experience against the stated expectations for the PforR instrument, particularly for institutional capacity building, results, management of fiduciary risks, environmental and social aspects, and evaluability. Since this is a new instrument, the evaluation has also paid attention to potential risks.

The PforR instrument meets the demand of client countries for financing and expertise to improve the performance and effectiveness of their own development programs. PforR operations disburse upon achievement of program results, provide support for the use of a government's own

systems, provide assurance that Bank financing is used appropriately, and ensure that the environmental and social impacts of the programs are adequately addressed.

For each PforR operation, the Bank carries out a process of identification, preparation/assessment, appraisal, and implementation support. The appraisal is informed by three assessments: a technical assessment (with a focus on strategic relevance and technical soundness of the program and its expenditure framework), a fiduciary assessment (with a focus on the procurement and financial management arrangements), and an environmental and social systems assessment (with a focus on the potential environmental and social impacts and risks).

These assessments identify measures to enhance performance, build capacity, and mitigate risks, which are reflected in an integrated risk assessment and in the resulting Program Action Plan (PAP). Preparation also includes the identification of disbursement-linked indicators (DLIs), each with a verification protocol to ensure that a credible mechanism is in place to monitor and verify its achievement.

The PforR Portfolio

After Board approval of the instrument, the Bank rolled out the PforR cautiously. The Board had limited aggregate PforR operations to 5 percent of total International Bank for Reconstruction and Development (IBRD) and International Development Association (IDA) commitments in the first two years. Following the *Two-Year Review*, the Board increased the limit to 15 percent of the three-year average of total IBRD and IDA commitments. If the PforR commitments

could exceed 15 percent of the three-year average of total IBRD and IDA commitments, Bank management would inform the Board and discuss how best to address the future demand for the instrument.

The PforR portfolio has grown rapidly, with significant design coherence that is broadly in line with original expectations. As of March 31, 2016, the Board had approved 39 operations, providing \$9.5 billion of financing to support \$49.9 billion in government programs, with a pipeline of 21 operations under preparation, totaling \$5.4 billion in expected Bank financing. The approved operations are largely on track.

All Bank Regions have PforR operations, with the Africa Region leading the way with 15. The introduction of PforR to different Regions appears to be influenced by a range of factors. In Africa, it has been influenced by the idea that PforR provides useful support for the regional agenda of building stronger institutions and delivering better services.

In other Regions, the instrument is regarded as a good fit for efforts to increase the emphasis on results and institutional capacity building. Some borrowers also focus on the predictability and general nature of disbursements. PforR operations now also cover most of the key sectors in which the Bank provides financing. Early approvals were largely for IDA countries, with a recent increase for IBRD clients.

Implementation Experience

Overall. The PforR instrument is an increasingly important lending vehicle for the Bank. Overall, the structure of the Bank's assessments for the PforRs—technical, fiduciary, and environmental and social—has proven to be appropriate, and the assessments have generally been credible

and comprehensive. The results frameworks, DLIs, and PAPs are often reasonably coherent, and risks related to PforR operations have generally been well identified and assessed. Nevertheless, there are areas in need of improvement when it comes to designing the programs to achieve results and to the monitoring and reporting systems.

The PforR was envisioned to help focus more on results than other existing instruments, as well as to help strengthen country systems and induce further alignment/harmonization among donors. Because none of the PforR programs has yet closed, it is too early to draw definite conclusions about whether the instrument is doing a better job of achieving these objectives than alternative approaches. Nevertheless, some insights can be derived from the early design and implementation experience.

While the programs do focus on results more explicitly than other instruments through the introduction of DLIs, these indicators are often—but not always—well integrated with the results frameworks. And while the results frameworks are often reasonably coherent, the program development objectives (PDOs) are rarely at the outcome level, and explanations of how the PforR objectives relate to the longer-term objectives of the supported government program are mostly absent from the program appraisal documents (PADs). To ensure a higher likelihood of achieving the ultimately desired developmental results, more consistent linking of the DLIs to the results frameworks and a clearer line of sight to the longer-term objectives of the program will be required.

The DLIs are designed with two main objectives: to be triggers for disbursements, with an inherent need for predictability, and

to provide incentives for performance—stretch targets. In a number of cases, the DLIs are linked to relatively small shares of total program disbursements, or to routine and repetitive actions, rather than key activities necessary to achieve the PDOs. This points to an inherent tendency to shift the balance toward the disbursement objective.

IEG finds that both ownership and partnership are well addressed in the Bank's program documents, and the field visits found a considerable degree of government ownership of the programs under implementation. However, there is no evidence yet that the instrument has encouraged much additional financing by other donors, let alone any broader use of the strengthened country systems. Capacity building is an important part of the programs, but specific goals could have been defined more clearly in some cases, and the implementation of capacity-building programs frequently has been delayed.

Both the Bank teams and government counterparts have moved well up the learning curve for this new instrument, and countries have been eager to rely on their own financial management systems and procedures. So far, the Bank's average costs for the preparation of new programs have been similar to those of other investment policy financing (IPF) operations, with significant variations among programs, while average Bank implementation costs have been significantly higher than for IPF operations. There may be increased positive externalities/public good aspects from strengthened country systems. Overall, however, there is not as yet sufficient evidence to derive any conclusions about the overall efficiency of PforRs.

Ownership and partnerships. The program appraisals appropriately discuss both ownership and partnerships, but only 2

of the 27 operations approved through fiscal 2015 involved formal cofinancing of the PforR instrument using the DLIs: Rwanda Agriculture, with the U.K. Department for International Development (DFID), and Tanzania Education, with the United States Agency for International Development (USAID). However, many PforR operations involve parallel financing with development partners through various means, including partners' financing of technical assistance requirements, or the Bank and other partners supporting the same government programs.

Technical assistance. Virtually all PforRs depend on the availability of technical assistance. This is provided through trust funds, organized financing by other donors, or parallel financing. Discussion of this assistance in PADs has sometimes been limited, though it would be appropriate to set out this assistance clearly at appraisal, including any linkages to actions in the PAPs.

Results frameworks. There is normally reasonably good coherence among the components of the PforR results frameworks, which consist of a PDO, a few PDO indicators, and a larger number of intermediate results indicators. However, in some cases the PDOs are not fully addressed by the indicators, or the PDO itself is incomplete in relation to the PAD discussion or the PDO indicators.

The PforR objectives are mostly institutional or focused on outputs/early outcomes. Objectives that lead to clear results at the outcome level are rare, and the PADs could benefit from pushing the envelope further in terms of concentrating on outcomes that are achievable during the program period. The focus in the PADs is largely on the program period, with not much discussion of the longer-term prospects and objectives of the programs, and thus there is also relatively little attention to the sustainability of the

targeted improvements beyond the program periods.

Disbursement-linked indicators. DLIs are given strong attention by all PforR parties, since their achievement determines the pace of disbursements. It is therefore important to keep in mind that the DLIs are not the main results of the PforR, but an instrument to contribute to the achievement of final outcomes. For some PforRs, the DLI sets are well structured and explained in relation to both the results chains and the sets of indicators in the results framework. However, there are other cases where the DLIs, with their attention to intermediate steps, may, subject to relevant PAD discussions, seem less well linked to program objectives. The role of DLIs to incentivize performance is also not always clear, and it sometimes seems that the DLIs have been designed to ensure disbursements and are not very demanding.

Program Action Plans. PAPs consist of actions agreed with the government, based on the findings of the different assessments. The implementation of all actions in the PAP is a contractual obligation incorporated in the Financing Agreements. However, a general obligation of that type can have little impact on a government's interest in executing individual PAP items.

IEG noted a tendency toward an increasing number of PAP actions—with as many as 49 actions in one case—which is likely to dilute the possible impact of the individual PAP items. It is also not clear why it seems necessary to include some items in the PAPs that are also in the DLIs, or actions also included separately in the Financing Agreements. Also, there are PADs that do not provide a clear explanation for the inclusion of actions in the PAPs, specifically with respect to the articulation of trade-offs among the actions recommended in the supporting PforR assessments.

Generally, the quality of PAPs at entry has been satisfactory, but implementation performance has been more uneven. One reason may be that there is, in practice, no penalty for slow or poor implementation of some PAP items, except for the actions linked to DLIs and effectiveness conditions. Another reason may be the inclusion of actions that are not critical to the achievement of PDOs. The reporting on individual PAP actions is also not systematic. A good practice in half of the reviewed PforR aide-memoires is the use of a table or annex that summarizes the status of each PAP action and the next steps. There may be room for more selectivity and focus in designing the PAPs. If they were to focus on (non-DLI) critical actions, there may also need to be greater clarity on the consequences of slow progress.

Political economy issues. Political economy constraints are behind many of the issues addressed by DLIs and PAPs, but those constraints are rarely discussed directly. Critical analysis tends to address only low-level issues. Interviews with team members during IEG missions confirmed their awareness of issues. The lack of published analysis is reportedly because of borrower-government sensitivity.

Use of country systems. The Bank has been committed to increasing the use of country systems for some time, including designation of the project as on-budget, aligning with the fiscal calendar, and relying on the range of national budget preparation and execution procedures, as well as intergovernmental transfers, audit, and national competitive bidding procedures for procurement. PforRs are part of this long-term attempt to design operations that better fit country contexts.

The PforR approach to country systems is different from the Bank's earlier country systems pilots. First, it uses the countries'

systems that are responsible for the program being supported, which may be at a higher standard than those for other programs in a country. Second, it does not insist that these program systems be at the same standard as the Bank's own policies and procedures for investment lending operations.

Technical assessments. Overall, the technical assessments are credible and comprehensive. They have mainly been prepared in collaboration with clients and, in some cases, with other development partners. The operations supported government programs that are well defined and clearly bounded, with adequate borrower capacity, borrower participation in design and implementation, defined program management mechanisms, and coordination processes with other development partners.

Expenditure frameworks are covered reasonably well. IEG rated the frameworks of 89 percent of the operations as moderately satisfactory or better. A common shortcoming is that there are few details in the PADs on the costing methodology used by the government to prepare the expenditure frameworks for the programs being supported. Without accurate costing, budgeting, procurement, and performance measurement can be compromised. The PADs include broadly reasonable economic justifications, but they rarely estimate economic or internal rates of return, and instead use other or partial measures.

Fiduciary systems. The systems used for PforRs are broadly the same as for other areas of the government programs. Overall, the fiduciary assessments have been comprehensive, and cover most of the key aspects that one would expect to find. The assessments are generally thorough in identifying relevant transparency measures such as timely provision of information to

stakeholders, disclosure of tender notices and award decisions, and parliamentary oversight. But some contextual or political economy issues may not be addressed, and the assessments are generally stronger on the formal systems, such as procurement laws and regulations, than on the practical realities, such as actual fraud and corruption.

Procurement systems. PforRs assess most program procurement systems as consistent with good public procurement principles, although implementation is often uneven because of a lack of political commitment, rigorous follow-up, trained staff, and effective demand-side processes. Both national and international competitive bidding systems are assessed. Challenges are sometimes pointed out without corresponding proposed actions, such as possible entry barriers for contractors created by their mandatory registration.

High-value contracts exclusion. PforR financing cannot normally be used for procurement packages above thresholds based on the type of procurement and the level of fiduciary risk. Following the *Two-Year Review* (World Bank 2015e), a provision was added that such contracts may be financed if they are important to the integrity of the program, and their cost is less than 25 percent of the overall program. Exceptions need to be approved by the Bank's managing director and chief operations officer. Excepted procurements may be handled outside the programs, and thus outside Bank oversight. This exclusion has not been a major obstacle in the PforRs reviewed, because excluded procurements can be financed from other sources, in which case the fiduciary risk is transformed into a funding risk to the part of the program not funded by the Bank, which could then become a risk to the development objective supported under the program.

Anti-Corruption Guidelines. Some borrowers expressed concern about the Bank's Anti-Corruption Guidelines. Under PforR financing, the borrower is responsible for taking action against possible fraud and corruption, and the Bank can investigate allegations and sanction parties if appropriate, which has raised some concerns. In light of the *Two-Year Review* (World Bank 2015e), new language has now been added to the guidelines, emphasizing that the country has the sovereign right to take actions in this area, and that Bank actions are solely for the purpose of determining compliance with its policies.

Environmental and social assessments. During the preparation of a PforR, the Bank assesses the degree to which systems manage and mitigate the environmental and social impacts of the program. The assessment also identifies and excludes high-risk activities—those that pose a risk of potentially significant adverse impact on the environment or affected people (activities classified as Category A under the IPF safeguards).

All reviewed assessments have been of reasonable quality, but with some shortcomings, most frequently related to the extent to which the capacity-building measures recommended in the assessments were followed up in the PAPs and the technical assistance provisions; the coverage of social issues; the adequacy of outreach and consultation with poor and vulnerable beneficiary groups; and the absence of monitoring indicators and reporting on safeguards implementation in the results frameworks of the operation.

Identification of program risks. The integrated risk assessment in the PAD identifies and consolidates the assessed risks. It also includes an overall risk rating for the project, based on the consolidation of risks associated with the operating environment—

country risk and stakeholder risk—and the main program-level risks—technical risk, fiduciary risk, environmental and social risk, and DLI risk. Based on the integrated risk assessments in the PADs of the 27 PforR operations approved by the end of fiscal 2015, the overall risks associated with these programs were rated as substantial in the majority of cases, as moderate in about a quarter of cases, and as high in four cases. Of the four major types of risk, fiduciary risks appear to be the main driver of overall program risk, followed by technical risks, DLI risks, and environmental and social risks.

Assessment of risk identification. Overall, IEG concurs with the conclusion of the *Two-Year Review* (World Bank 2015e) that the risks associated with PforRs have been identified and assessed reasonably well, with a few exceptions. Most of these exceptions involved the underrating of risks, particularly those related to technical, fiduciary, and DLI issues. Environmental and social risks, in contrast, had a slight tendency to be overrated.

Application of high environmental and social risk exclusion. IEG's review has found that the allowable exclusion of high environmental and social risk activities has sometimes been interpreted in an overly cautious manner. This has led to the avoidance in PforRs of investments (akin to Category B) that would normally be integral to the supported programs. As also noted in the *Two-Year Review*, this exclusion has significantly reduced the scope of several PforRs in relation to that of the supported government programs, which raises a concern of how such investments will be handled if they are not subject to the oversight associated with Bank involvement.

Strengthening of country systems. The potential for strengthening the institutional capacity of national systems is a key feature

of PforRs. Hence, it is expected that a priority area of both preparation and implementation support will be to strengthen the capacity of the national institutions to implement the program. And indeed, the assessments of technical, fiduciary, and environmental and social systems have usually identified numerous measures to strengthen the performance of program systems to ensure that they achieve the expected results.

Capacity-building measures. The identification and assessment of measures to strengthen the capacity and performance of program systems has been mostly satisfactory. In many cases, the PforR approach offered greater scope and flexibility compared with earlier Bank operations, stimulating the design of innovative features. In other cases, the measures to strengthen program systems drew from earlier projects and the support of donor partners. In some cases, the capacity and other system gaps were avoided or mitigated through the purposive bounding of PforRs within the government programs. In a few cases, the system capacity gaps were identified in the assessments, but the measures to address them were not clearly defined or were left undefined.

Implementation and monitoring. At the implementation stage, the most important issues relate to the progress of PAPs and the adequacy of monitoring of program results. PforR operations give more attention than other Bank operations to monitoring and evaluation of results. Disbursements are based on achievement of monitorable indicators, rather than inputs, and through specific and transparent verification protocols. Over 40 percent of the operations include program support for impact evaluations. In addition, about three-quarters of the operations provide for greater disclosure of the results of monitoring and evaluation processes and

greater engagement of civil society in this work.

PAP implementation experience. Based on the findings from operations visited by IEG, the implementation of PAP actions has often been substantially delayed, sometimes with important implications for the timely achievement of the programs' results. Both in the country monitoring processes and in some aide-memoires, there is a tendency to focus on the DLIs, with less systematic attention to the PAP items and the indicators in the results frameworks.

Oversight and reporting of environmental and social aspects. In projects visited by IEG, the stakeholder engagement associated with the programs' social systems was being fully implemented and supervised by the Bank teams. However, preexisting capacity and budgetary constraints continue to limit the effectiveness of the program systems. Planned technical assistance and PAP actions—mostly involving the preparation of guidelines and training for program staff—were being implemented, but had been unable to address the systemic constraints. Also, the systems were being implemented, and country teams were up to speed, but there was virtually no reporting of environmental and social effects, even at the program-management level. It would thus seem that environmental and social effects have not normally been seen as integral to PforR reporting.

Risk management experience. The management of risks is progressing well, albeit the reporting system is inadequate. PforRs appear to be making progress with the management of all kinds of program risks. The Implementation Status and Results Reports (ISRs), however, do not adequately reflect the frequency of delays, and they provide no information on the environmental and social effects of the

operations. This is surprising, since environmental and social risks have been rated moderate or substantial in most PforRs, and the monitoring of impacts is an essential element of the environmental and social risk management framework.

Capacity-building experience. The implementation of capacity-building measures is progressing reasonably well, although with frequent delays. In most cases, the measures focused on the program systems, as appropriate. In a few cases, the strengthening of program systems is being expanded to the whole country.

Recommendations

Strengthen the design of the results framework and DLIs to ensure that the PAD presents a clear line of sight to developmental results. It is essential to ensure that all programs produce a well-structured results framework that responds to borrower priorities through a set of logical, achievable steps, reflecting the PDOs and the critical role of DLIs. The DLIs serve partially offsetting objectives, including as triggers for disbursements (requiring predictability) and as incentives for performance (requiring stretch targets). The trade-offs between these considerations should be revisited. This could include steps, such as guidance to staff, to ensure that:

- The results frameworks reflect the development expectations and logic (theory of change) of the PforR within the context of the supported government program.
- The PDOs reflect measurable, monitorable developmental results (that is, at the outcome level, which could include institutional strengthening).

- There is a clear explanation of the rationale for DLI selection and how the DLIs achieve the PDOs.
- The financing associated with each DLI should reflect, among other things, the possible incentive effect of the size of such financing.
- The individual DLI targets need to strike the right balance between predictability of disbursement and achievement of results.

Strengthen the design and monitoring of the PAPs. The following is recommended to improve the effectiveness of the PAPs, including mitigation of a tendency for the proliferation of PAP actions and to ensure their adequate consideration during program implementation, when much attention has tended to focus on the DLIs:

- Focus the PAP on a few key areas where actions are important to enhance the capacity and performance of the implementing agencies and to mitigate risks.
- Devote greater attention to the monitoring of PAP implementation, including more systematic coverage in the ISRs.
- Provide clearer guidance to task teams on how to address poor implementation of PAPs.

Minimize the overly cautious interpretation of the high environmental and social risks exclusion. To address the overly cautious interpretation of the high environmental and social risk exclusion:

- Strengthen awareness of the guidance that this exclusion is only intended for high-risk activities, and is not intended to exclude substantial- and moderate-risk activities, such as small- and medium-scale investments that are integral to the supported programs. These projects are likely to cause

mostly local and short-term negative environmental and social impacts, for which effective mitigation measures are readily available.

Strengthen the monitoring and reporting of results to cover the entire results framework systematically, as well as the environmental and social performance of the projects. To address the tendency in the country monitoring processes and in some aide-memoires to focus mainly on the DLIs, with less systematic attention to results indicators, PAP items, and environmental

and social impacts, all of which should be important for the achievement of the PDOs, it will be important for the Bank to:

- Ensure that the ISRs and supervision reports systematically cover the entire results framework and its supporting elements.
- Ensure that program systems adequately report on the environmental and social effects associated with the implementation of the PforRs, and that the Bank has timely access to these reports.

Management Response

Introduction

The Management of the World Bank Group institutions welcomes the Independent Evaluation Group (IEG) report entitled *Program-for-Results: An Early-Stage Assessment of the Process and Effects of a New Lending Instrument*; and we thank IEG for the good insights into the first set of Program-for-Results (PforR) operations and for the opportunity to comment on the report. Management appreciates IEG's recognition that the first set of PforR operations indicates that the new instrument is performing well overall, and that the assessment supports the overall findings of Management's *Two-Year Review* discussed with Executive Directors on April 9, 2015 (World Bank 2015e). Management agrees with the majority of the review's recommendations, and concurs with most of its observations.

Context. PforR is the first new World Bank instrument in 30 years. It was carefully crafted to enable the World Bank to better respond to changing development needs, meet demand from client countries, and enhance development effectiveness by supporting a government program of expenditures, building institutional capacity, and tying financing to achievement of results. Client countries are increasingly implementing their own programs for development and poverty reduction that are rooted in their own legal, policy, regulatory, and institutional environments, and they are asking development partners for financing and expertise to improve the programs' effectiveness and efficiency in achieving results. There is a consensus, including in the IEG review, that PforR covers an important need in meeting this client demand. The design, key features, and initial implementation were based on intensive consultations with clients, staff, and development partners. The uptake of the instrument and its positive reception by clients has also led other donors to adopt it as a model. As the instrument matures, Management remains committed to taking a learning approach, including making policy and process adjustments, if needed. The IEG report is a valuable input into the learning process.

PforR Portfolio Growth. IEG's work was based on a review of a relatively young PforR portfolio – the 27 operations approved by end-FY15. At that point, the average age of an operation was 1.3 years; of the 27 operations, 23 were effective and 18 had started to disburse. The portfolio continues to grow. By the end of September 2016, the number of approved operations has increased to 52, Bank financing has increased from \$4.9 billion to \$12.9 billion, and government programs leveraged have grown from \$14.7 billion to \$60 billion. The number of countries has also increased: 15 new countries have approved or pipeline PforR operations. As the portfolio evolves, Management expects

to learn more about the effectiveness of the instrument and to use this evaluation and continued feedback from clients to help improve it.

World Bank Management Comments

Although Management welcomes and is in general agreement with the analysis and findings of the review, we have some comments on specific aspects.

Results Framework and Nature of Results. The report discusses PforR results frameworks and suggests reflecting the “theory of change” in those frameworks. Management partially concurs with this view. While Management will provide guidance on improving clarity regarding the drivers for change and the results chain, the World Bank’s role is to assist governments in improving, and financing a slice of, their own programs. Such results frameworks are part of the government overall program and the ownership of that is critical for the instrument. Moreover, when financing a slice of a program, part of the challenge is for the clients and World Bank team to work together to develop an operation that is of manageable size and within the implementation capacity of the government, and that can lay the groundwork for sustainable and substantial change. A single operation may not be able to include the complete set of the changes envisioned by the theory of change; moreover, the changes might be a set of institutional actions that do not easily map into a theory-of-change results framework.

Disbursement-Linked Indicators. The report comments that in the first set of operations, “in some cases the disbursement-linked indicators (DLIs) are not well integrated with the results frameworks.” Management notes that the PforR instrument has two key features: helping build institutions and focusing on results. DLIs, by design of the instrument, can be derived from the results framework or the Program Action Plan (PAP). Moreover, DLIs related to institutional strengthening are important in many cases to improve the systems that support results – for example, DLIs for strengthening the program’s environmental and social systems, monitoring and evaluation systems, or grievance redress systems. The use of such DLIs has been seen to be critical and effective in many sectors, including health, education, water and sanitation, and public sector reform. These “institutional strengthening” DLIs are therefore key to achieving the outcomes of the program.

Partnership. The report notes “there is no evidence that the instrument has encouraged much additional financing by other donors.” Management would like to note that, as the PforR policy paper emphasized, partnership goes beyond financing. It is about enhancing the World Bank’s ability to partner with the government as well as other

development partners. The partnership could be in the form of joint preparation or implementation support through parallel financing. The evaluation notes that virtually all PforRs rely on the availability of technical assistance for capacity building – assistance that is mainly provided through other development partner funds.

Sustainability. The report raises concerns about the sustainability of operations, indicating “there is relatively little attention to the sustainability of the targeted improvements.” Management would like to note that despite PforR’s broader program orientation, the World Bank’s financing remains time-bound and discrete (just as it does for development policy and investment project operations). There is no assumption on the part of teams and the government that World Bank financing or support will continue indefinitely, and, while factors contributing to sustainability are built into programs, there can be no guarantee. The expectation is that World Bank financing will help the government program consolidate and accelerate progress that can then be sustained independently.

Costing Methodology. The report assumes that the theory of change rests on the assumption that the “costs of achieving the intended results have been accurately estimated and presented by the borrower and can be tracked.” It also says “unclear costing methodology [was] used to prepare expenditure frameworks that support programs.” Management notes that PforR policy requires the World Bank to assess the expenditure framework supporting the program and to look at efficiency and effectiveness issues. It is through this analysis that PforR can help focus on the totality of the expenditures of the program and identify measures to enhance efficiency and effectiveness.

Recommendations

Management welcomes the report’s approach in formulating a limited number of focused recommendations, with some specific actions listed as possible options or suggestions for Management’s consideration. In principle, Management welcomes the recommendations and is planning to incorporate the majority of the proposed actions through Operations Policy and Country Services advisory support for PforR operations and in updated guidance notes and outreach, including e-learning modules, Global Practice thematic workshops, and client events.

Recommendation 1: Strengthen the design of the results framework and DLIs to ensure that the PAD presents a clear line of sight to development results.

Management is generally in agreement with this recommendation. Specifically, Management agrees on the principles that there should be a clear explanation of why the DLIs were selected and how they will help achieve the desired outcomes of the

program and project development objectives (PDOs); that the financing associated with each DLI should reflect, among other things, the possible incentive effect of the size of such financing; and that individual DLI targets need to have the right balance between predictability of disbursement and achievement of results. Management will update PforR guidance and training content to strengthen these messages. Management cannot, however, commit to reflecting the “theory of change” in all operations, and linking each DLI to PDOs. As noted above, clients and the World Bank team work together to develop an operation that is of manageable size and within the implementation capacity of the government, and that can lay the groundwork for sustainable and substantial change. A single operation may not be able to include the complete set of changes and outcomes envisioned by a results framework based on a theory of change. As for linking each DLI to the PDO of an operation, some DLIs may be used for strengthening the program’s systems, and as such, may not necessarily be included in the results framework of the operation.

Recommendation 2: Strengthen the design and monitoring of the PAPs. Management welcomes this recommendation. It is fully consistent with current guidance and training material that clearly point out that the PAP should focus on being strategic in selecting actions, and should be a main element of implementation support. Most PforR operations are already aligned with the guidance. The Implementation Status and Results Report (ISR) for PforRs is being revised, partly to address such improvements as better monitoring of PAPs. Management will work on reflecting lessons learned and best practices to improve the implementation of PforR – both in the context of the revised guidance notes that are being prepared as part of the follow-up to Management’s *Two-Year Review*, and in the PforR Academy for task team leads that will be delivered this year.

Recommendation 3: Minimize the overly cautious interpretation of the “high” environmental and social risks exclusion. Management agrees that some of the earlier operations may have taken an overly cautious approach. Management will continue to raise awareness among task team leads and environmental and social staff and consultants on how to consistently address environmental and social risks under PforR operations.

Recommendation 4: Strengthen the monitoring and reporting of results to cover systematically the entire results framework as well as the environmental and social effects of the projects. Management agrees with this recommendation to ensure that the reporting of results covers the entire results framework. The ISR already includes the results framework, and Management has recently revised the framework following the *Two-Year Review*. The ISR also includes the risk framework (now revised to use the

Systematic Operations Risk-rating Tool, to be consistent with other instruments) and the PAP. Hence, all recommended requirements are in place.

Conclusions

Management sees this review as an important resource that adds to its own Two-Year Review and will be a good input to future assessments. Management is committed to a learning-by-doing process and will use the lessons of experience to refine the PforR instrument. This review and its analysis and recommendations are already contributing to ongoing and planned PforR learning and outreach, including improved e-learning modules, other training activities in Washington, D.C., and in field offices, and events in which PforR clients can share their experiences. Detailed responses to the recommendations are set out in the Management Action Record.

Management Action Record

IEG Findings and Conclusions	IEG Recommendations	Acceptance by Management	Management Response
<p>Program-for-Results (PforR) results frameworks and disbursement-linked indicator (DLIs) are often reasonably coherent, but some improvements are needed. Objectives in the examined operations are mostly institutional or focused on early outcomes. Objectives that lead to clear results at the outcome level are rare, and the project appraisal documents (PADs) could benefit from pushing the envelope further in terms of focusing on outcomes that are achievable during the program period. Furthermore, while the project development objectives (PDOs) should be outcomes that are achievable during the program period, the PADs should be expected to go beyond the program period itself and discuss the longer-term objectives of the government program being supported and the sustainability of program improvements</p> <p>All programs need a well-structured results framework that responds to borrower priorities through a set of</p>	<p>Recommendation 1: Strengthen the design of the results framework and DLIs to ensure that the PAD presents a clear line of sight to developmental results. This could include steps, such as guidance to staff, to ensure that:</p> <ul style="list-style-type: none"> ▪ The results frameworks reflect the development expectations and logic (theory of change) of the PforR within the context of the supported government program. ▪ The PDOs reflect measurable, monitorable developmental results (that is, at the outcome level which could include institutional strengthening). ▪ There is a clear explanation of the rationale for the DLI selection and how the DLIs help achieve the PDOs. ▪ The financing associated with each DLI should reflect, among other things, the possible 	<p>World Bank: Partially Agrees</p>	<p>Management is generally in agreement with the recommendation to strengthen the design of the results framework and DLIs. Management also agrees with the recommendation for the PAD to present a clear line of sight to development results.</p> <p>Specifically, Management agrees on the principles that (i) there should be an explanation of why the DLIs were selected and how they will help achieve the results of the Program being supported; (ii) the financing associated with each DLI should reflect, among other things, the possible incentive effect of the size of such financing; and (iii) the DLI targets need to have the right balance between predictability of disbursement and achievement of results.</p>

IEG Findings and Conclusions	IEG Recommendations	Acceptance by Management	Management Response
<p>logical, achievable steps, reflecting the theory of change to achieve the stated PDOs, and development-linked indicators (DLIs) that are clearly linked to these. However, in some cases the DLIs are not well integrated with the results frameworks, some DLIs are linked to very small shares of total program disbursements, or are linked to routine and repetitive actions.</p> <p>The DLIs serve partially offsetting objectives, including as triggers for disbursements (requiring predictability) and as incentives for performance (requiring stretch targets). The trade-offs between these considerations should be revisited.</p>	<p>incentive effect of the size of such financing.</p> <ul style="list-style-type: none"> ▪ The individual DLI targets need to find right balance between predictability of disbursement and achievements of results. 		<p>Emphasis on these principles will be strengthened in updated guidance notes, through training workshops and outreach events, and when Operations Policy and Country Services provides advice to PforR teams.</p> <p>Management disagrees with the recommendation to reflect the “theory of change” in all operations and linking each DLI to PDOs. Results frameworks are part of the government overall program and the ownership of that is critical for the instrument. Moreover, when financing a slice of a program, part of the challenge is for the clients and World Bank team to work together to develop an operation that is of manageable size and within the implementation capacity of the government, and that can lay the groundwork for sustainable and substantial change. A single operation may not be able to include the complete set of changes and outcomes envisioned by the</p>

IEG Findings and Conclusions	IEG Recommendations	Acceptance by Management	Management Response
			theory of change; moreover, those changes might be a set of institutional actions that do not easily map into the theory of change.
<p>▪ The PforR guidelines state that a Program Action Plan (PAP) should be selective. However, some PAPs have too many actions and some program documents do not clearly justify the inclusion of actions in a PAP or articulate the trade-offs among the long lists of potential PAPs emerging from the assessment process. During implementation, many PAP actions have been substantially delayed, and monitoring documents give less systematic attention to PAP items.</p> <p>The incentives among staff and government counterparts are mostly for achieving the DLIs, with the risk mitigation activities captured in the PAP supervision receiving less attention. There is also room for greater selectivity and focus in designing the PAPs. Finally, PAP implementation performance has been</p>	<p>Recommendation 2: Strengthen the design and monitoring of the PAPs. This could include steps to:</p> <ul style="list-style-type: none"> ▪ Focus the PAP on a few key areas where actions are important to enhance the capacity and performance of the implementing agencies and to mitigate risks. ▪ Devote greater attention to monitoring of PAP implementation, including more systematic coverage in the Implementation Status and Results Reports (ISRs). ▪ Provide clearer guidance to task teams on how to address poor implementation of PAPs. 	<p>World Bank: Agrees</p>	<p>Management agrees with the recommendation.</p> <p>Management will enhance the guidance to staff on this issue to emphasize the message about the design of the PAP and how best to address issues that may arise during design. This enhancement will be provided in the context of revised guidance notes and training events.</p>

IEG Findings and Conclusions	IEG Recommendations	Acceptance by Management	Management Response
<p>somewhat uneven, and the reporting on individual PAP actions has not been systematic.</p>			
<p>The overly cautious application of the “high” environmental and social risks exclusion has unnecessarily reduced the scope of several PforRs in relation to government programs. All reviewed operations have appropriately excluded high-risk environmental and social risk activities. But in some cases, the interpretation of this policy requirement appears to have been excessively cautious, to the extent that it has significantly reduced the scope of the PforRs in relation to the supported government programs, and the opportunities associated with strengthening the programs’ systems. For example, the India Maharashtra Rural Water Supply and Sanitation Program excludes schemes involving highly polluted water sources, the Uganda Municipal Development Program excludes water treatment plants and sanitary landfills, and the Ethiopia Local Government Program excludes roadways outside existing rights-of-way. This raises a concern</p>	<p>Recommendation 3: Minimize the overly cautious interpretation of the “high” environmental and social risks exclusion. To this end:</p> <ul style="list-style-type: none"> ▪ Strengthen awareness of the guidance that this exclusion is only intended for high-risk activities and is not intended to exclude substantial- and moderate-risk activities such as small- and medium-scale investments that are integral to the supported programs and are likely to cause mostly local and short-term negative environmental and social impacts for which effective mitigation measures are readily available. 	<p>World Bank: Agrees</p>	<p>Management agrees with this recommendation and believes that some of the earlier operations may have taken an overly cautious approach. However, the most recently approved set of operations show that this trend is now changing.</p> <p>Using guidance, training, and targeted workshops, Management will continue to raise awareness among task team leads and environmental and social staff on how to consistently address environmental and social risks under PforR operations.</p>

IEG Findings and Conclusions	IEG Recommendations	Acceptance by Management	Management Response
<p>about how such small- and medium-scale investments that could cause mostly local and short-term impacts for which effective mitigation measures are readily available will be handled if they are not subject to the oversight associated with Bank involvement.</p>			
<p>There is an observed tendency in the PforR country monitoring processes and in some supervision reports to focus mainly on the DLIs. Less systematic attention is given to the indicators in the results framework, PAP items, and environmental and social impacts, all of which should be important for the overall progress of the programs and for the achievement of the PDOs.</p> <p>The ISRs generally report that the environmental and social risk management measures are “in progress,” but provide little information on actual environmental and social effects. This is surprising, since environmental and social risks have been rated as moderate or substantial in most PforRs, and the monitoring of impacts is an essential</p>	<p>Recommendation 4: Strengthen the monitoring and reporting of results to cover systematically the entire results framework as well as the environmental and social effects of the projects. To this end improve staff guidance and awareness to:</p> <ul style="list-style-type: none"> ▪ Ensure the ISRs and supervision reports systematically cover the entire results framework and its supporting elements. ▪ Ensure that program systems adequately report on the environmental and social effects associated with the implementation of the PforRs, and that the World Bank has timely access to these reports. 	<p>World Bank: Agrees</p>	<p>Management agrees that the monitoring and reporting of results should cover the whole results framework. This is reflected in current guidance notes and learning materials. The ISR includes the results framework, the revised risk framework, and the PAP. Management will further clarify these messages in training material and guidance on preparation of and implementation support for PforR operations, and will monitor implementation.</p>

MANAGEMENT ACTION RECORD

IEG Findings and Conclusions	IEG Recommendations	Acceptance by Management	Management Response
<p>element of the environmental and social risk management framework. While it is appropriate for PforR operations to rely on the country systems for implementing the environmental and social principles as expected under the policy, it is important for the World Bank to be able to oversee, review, and evaluate implementation based on timely and accessible monitoring information in its own project records.</p>			

Chairperson's Summary: Committee on Development Effectiveness

The Committee on Development Effectiveness discussed the report of the Independent Evaluation Group (IEG) entitled *Program-for-Results: An Early-Stage Assessment of the Process and Effects of a New Lending Instrument* and the draft response of Management of the World Bank Group institutions.

Summary

The committee welcomed the early-stage assessment of the Program-for-Results (PforR) instrument, which was approved in January 2012, and endorsed the findings and recommendations. Members expressed robust support for PforR, and were pleased that IEG found that the instrument is performing well to date. They agreed the assessment will strengthen and contribute to PforR's further development. Members saw PforR as an important addition to the World Bank's menu of instruments, and valued the main objectives of PforR, including responding to client demand for financing.

Members underscored the importance of focusing on fine-tuning implementation to lead to even better results, emphasizing quality assurance and risk management; they appreciated the report's focus on these aspects. In this context, members acknowledged the complexity of linking results frameworks, Disbursement-linked indicators (DLIs), the Program Action Plans (PAPs), and considerations when using the program environmental and social systems. They welcomed the findings that there is a considerable degree of government ownership and commitment to work towards results, and underscored that part of the challenge is developing PforR operations that are of manageable size, are within the implementation capacity of the government, and contribute to the long-term sustainability of results. Members noted that while there are donor partnerships in conjunction with PforR operations, this support has not been reflected in official project cofinancing. Donors' enthusiasm for the instrument could be improved if the World Bank were to share the success of its initial experience using PforR.

Members urged the World Bank to consider the complementarity of PforR to the other World Bank lending instruments, namely investment policy financing (IPF) and development policy financing (DPF). Some members commented that a review may be warranted of the consistency of application of environmental and social

safeguards/considerations across all instruments. Members underscored that capacity building measures have been satisfactorily identified but there was a need to strengthen capacity and performance of programs and the implementation of capacity building programs.

Members agreed with the need to strengthen the monitoring and reporting of results to cover the entire results framework systematically. They also highlighted the importance of aligning the expected results of PforR operations with the respective country's development agenda and with the World Bank's overall objectives, e.g., through Country Partnership Frameworks (CPFs).

Recommendation and Next Steps

Members generally agreed with the report's findings and recommendations. They encouraged Management to continue with PforR reviews and impact evaluations once operations are completed and looked forward to future reporting through the quarterly Operations Update.

Issues Discussed

Recommendation 1. Management noted its general agreement with recommendation 1 and the need to strengthen the design of the results framework and DLIs. Management however disagreed with the recommendation that the Results Framework reflect PforR's "line of sight" or "theory of change" and that the DLIs should clearly link to this "line of sight." Management noted that a single operation may not be able to include the complete set of changes and outcomes envisaged by a "line of sight"-based Results Framework and, with respect to linking each DLI to the project development objectives (PDOs) of an operation, added that some DLIs cannot be directly mapped into a Results Framework that relates directly to the "line of sight."

Some members agreed with Management that it is unrealistic to expect that a perfect link can/will be made between output and development outcomes; but at the same time, indicators should not be solely focused on institutional aspects. They generally agreed with the spirit of the recommendation, that the design of the results framework and DLIs could be strengthened to ensure that the project appraisal document (PAD) presents a line of sight to development results. A few commented that the weaknesses identified by IEG in establishing DLIs were not idiosyncratic to PforR operations, and that the linkages likely depend on the type of sector and financing of a slice of a program. While acknowledging the difficulties in designing

DLIs, they concurred that there should be a clear rationale for the selection of DLIs and how they contribute to achieving results that support the PDOs. They encouraged Management to explore how the Results Frameworks might reflect the development expectations and logic of the PforR.

Implementation. Members queried how different country conditions, including institutional frameworks, affect the implementation of PforR. Members observed that the average implementation support costs have been significantly higher than for IPF operations, and commented that consideration might be given to exploring how to increase efficiency in implementation. In this regard, they suggested that one viable approach may be to increase collaboration through harmonization of operations and technical assistance with other donors and development partners, including those who provide parallel financing. Management informed that the initial costs for preparation are higher than IPF due to the newness of the instrument, including World Bank teams conducting technical, fiduciary, and environmental and social assessments of the program, but if there is a subsequent PforR in the same country, the costs tend to go down. For implementation support, the increased costs are mainly driven by the greater role the Bank is carrying in supporting the selected program, rather than a more focused scope in a traditional project.

Members welcomed IEG and Management's agreement on the need to strengthen the design and monitoring of the Program Action Plans (PAPs). They asked how clients are incentivized to implement institutional improvements when they are not linked to disbursements. They noted that there was a mandate to reporting the implementation of all PAPs in the Implementation and Status Reports (ISRs), and some emphasized the benefit of public ISRs to ensure accountability and improve implementation. Management agreed going forward to including PAP implementation information in the publicly available ISRs.

1. The Program-for-Results Instrument and the Evaluation Approach

The PforR Instrument

The World Bank introduced the Program-for-Results (PforR) instrument in 2012 (World Bank 2011a). In 2015, as mandated by the Bank's Board of Executive Directors, Operations Policy and Country Services (OPCS) reviewed the first two years of PforR operation (World Bank 2015e). This review assessed the early experience with the design and implementation of operations and the challenges faced by borrowers, development partners, and Bank staff. The Board of Executive Directors requested an early Independent Evaluation Group (IEG) evaluation of the instrument as well.¹ This process evaluation responds to that request.

This new instrument is designed to enable the Bank to respond to changing development needs, meet demand from country clients, and enhance development effectiveness. By directly supporting government programs, PforR operations are expected to help countries strengthen institutions, build capacity, and enhance partnerships with stakeholders to achieve lasting impact. They link disbursements to achievement of results that are tangible, transparent, and verifiable. The design of the instrument is intended to accommodate a broad range of countries, sectors, and programs. It is also expected to enable the Bank to leverage its own financing and to partner with other development organizations in supporting country programs (see World Bank 2011a, para. 88).

PforR responds to country demand for financing and expertise to support their efforts in implementing their own programs for development and in improving program effectiveness and efficiency in achieving results. The Bank's ability to meet this demand has been constrained by the limitations of its long-standing investment policy financing (IPF) and development policy financing (DPF) instruments. As stated in the Board paper (World Bank 2011a, para. 14), neither instrument fully allows support to a government program of expenditure, building institutional capacity, and tying financing to achievement of results.² With PforR, the Bank now has three complementary lending instruments to offer country clients: policy support (DPF), project support (IPF), and program support (PforR).

The design of the PforR instrument benefited from the experience the Bank has gained in designing and implementing IPF and DPF operations: specifically, with technical and design issues, results definition, fiduciary systems, environmental and

social impacts and risk management, and improving policy environments. It also built on decade-long experience – both inside and outside the Bank – with sectorwide, program-based, and results-focused operations (see appendix B). Some key lessons incorporated into the design include the importance of using the institutional setup of the government program, ownership of the results framework, and transparency and accountability, such as having credible verification processes.

PforR operations, as described in the original Board paper (World Bank 2011a), are expected to do the following (these points are discussed as relevant later in this report and are summarized in chapter 6):

- **Finance and support borrowers' programs**, which can be ongoing or new, sectoral or subsectoral, national or subnational, as well as community development programs.
- **Disburse upon achievement of program results**, as determined by the achievement of indicators that can be monitored and verified rather than disbursing for inputs. Advances of up to 25 percent of outstanding commitments are allowed. Together with funds from other sources, Bank disbursements will finance a borrower's expenditure program rather than being linked to individual transactions.
- **Provide support for the use of a government's own systems to implement the program**, including for financing planning, procurement, anti-corruption, and environmental and social standards.
- **Provide assurance that Bank financing is used appropriately and that the environmental and social impacts of the programs are adequately addressed**. To this end, the Bank will assess a program's fiduciary and environmental and social management systems, and agree as necessary with a borrower on any additional measures to provide assurance that potential impacts to the environment and affected people are adequately addressed.
- **Focus on strengthening the institutional capacity needed for programs to achieve their desired results**, thereby enhancing development impact and sustainability. The strengthening of capacity to implement a program will be a priority area for both preparation and implementation support.
- **Support improvements in governance and transparency** by making program information publicly available and monitoring the achievement of results, including through enhancing the role of beneficiaries and civil society organizations.
- **Help strengthen partnerships with governments and development partners, and increase efficiency by reducing transaction costs for the government and development partners**. While the Board paper (World Bank 2011a) found it difficult to predict the budget implications for the Bank, experience with

IPF projects with PforR features suggested that costs might be within the norm for IPF projects.

THE THEORY OF CHANGE

The theory of change of the PforR instrument is that PforR operations, working alongside the other two instruments, will enable the Bank to assist country clients in delivering priority results more efficiently by working through their own country systems, leveraging Bank financing with that of partners and other development organizations, and strengthening their own systems. The theory of change rests on three assumptions. The first is that working with country program systems will help strengthen them. This is a reasonable assumption for three reasons. First, sector ministries engage directly in the budget process and are less likely to work off-budget because of their close relationships with donors. Second, donors have developed heightened concern about public financial management (PFM) and combating corruption, because of fiduciary concerns about their resources passing through national PFM systems, and because of the key role of these systems in linking policy and implementation. Third, it reduces transaction costs by avoiding multiple donor procedures and adopting the government standard.

A second assumption is that the costs of achieving the intended results have been accurately estimated and presented by the borrower and that the borrower can track these costs. This is important to provide reasonable assurance that program expenditures are used with due attention to the efficient use of resources. A third assumption is that it is possible to carry out technical, fiduciary, and environmental and social assessments, and in each case come to a clearly formulated, reliable conclusion. When systems are not adequate, they can be made so through targeted strengthening.³

PforR operations focus on the behavioral and institutional changes that are required to realize this targeted strengthening, and in turn achieve results and manage associated risks. Hence it is expected that many will require some level of capacity-building activities, which will be informed by the technical, fiduciary, and environmental and social systems assessments. Capacity-building support, where needed, can be provided through different modalities, from direct technical assistance and training to specific actions or indicators that will strengthen performance.

The PforR instrument is intended to complement, not replace, the Bank's two existing lending instruments. While all Bank instruments focus on development results, borrowers are now able to choose from a wider range of instruments to suit their objectives, desired results, and risks. The description below, largely taken from

the Board paper (World Bank 2011a), summarizes the differences and complementarities among the instruments (see also table 1.1).

- **Development policy financing.** DPFs remain the primary Bank instrument to support policy and institutional actions. They focus on discrete policy actions that are under the direct control of governments, and they link disbursements to evidence that such actions have been adopted. As stated in the Board paper (World Bank 2011a), Bank management sees the DPFs as a practical and effective way of supporting policy actions that help create the enabling conditions for improving results – for example, when new regulations are required for the better functioning of markets or new policy frameworks are necessary to improve government efficiency. DPFs provide general budget support and do not earmark loan proceeds for specific programs. Moreover, DPFs disburse against specific policy and institutional actions and not against the results or outputs and outcomes associated with sector and program expenditures.
- **Investment project financing.** IPFs support a heterogeneous range of operations, but their common characteristic is typically the financing of specific investment activities that involve a set of expenditure transactions, most of which are used for the purchase of works, goods, and services. As stated in the Board paper (World Bank 2011a), Bank management sees the IPFs as a practical and effective way of supporting the achievement of results when risk management and controls are needed on the inputs side (such as construction or technology) and when technical design and implementation challenges are critical bottlenecks to achieving results. Often such situations involve discrete, one-off activities (for example, the construction of a large infrastructure project or the purchase of expensive and technically complex equipment). By focusing on the proper implementation and risk management of individual transactions, IPF operations put the emphasis of Bank-client relations on making sure that the right inputs and technology are in place and the operation is implemented as planned.
- **Program-for-Results.** Many of the development challenges countries face cannot be addressed solely through discrete policy actions or the proper technical implementation of a program. For example, improving service delivery (such as better-maintained roads, functioning schools and health clinics, effective agricultural extension services) may require both policy actions (such as a decentralization law) and some discrete investment activity (such as constructing new schools or contracting works for road maintenance). But in many cases, these are not sufficient for the achievement of results. Schools can be built, but teachers may remain absent; health clinics

may have new equipment, but essential drugs may not be available at the point of service; and rural roads may remain unmaintained, despite the existence of contracts. Addressing such bottlenecks involves improvements in the governance of institutions and systems, including capacity building and changes in management practices and behaviors by service providers and users alike. PforR can be the instrument of choice when the objective is to support the performance of a government program using the government’s own systems, when the results require expenditures, and when the risks to achieving the program’s objectives relate to the capacity of the systems – for example, the monitoring and evaluation (M&E), fiduciary, and environmental and social systems – to achieve better results.

Table 1.1. Complementary Lending Instruments

Category	Project support lending (IPF)	Program-for-Results	Policy support lending (DPF)	Category
Purpose	Supports specific investment operations	Supports government programs or subprograms	Supports policy and institutional actions	Purpose
Disbursement mechanism	Disburses against specific expenditures that support the operation	Disburses upon achievement of results and performance indicators	Disburses against policy and institutional actions	Disbursement mechanism
Implementation mechanisms	Bank IPF rules and procedures Funds for specific expenditures	Program systems Funds for specific expenditure program	Country policy processes Non-earmarked funds for general budget support	Implementation mechanisms

Source: World Bank 2011a.

SELECTION, APPRAISAL, AND IMPLEMENTATION PROCESSES

The Bank will determine the choice of lending instrument for specific countries, sectors, and programs in the context of its Country Partnership Framework and its assessment of the country’s policies, programs, and institutional capacity. PforR has the potential for significant development impacts, though it also has risks. For that reason, certain high-risk activities have been excluded from PforR operations: activities that pose a risk of potentially significant and irreversible adverse impacts on the environment or affected people (activities classified as Category A under IPF) and activities that involve procurement of works, goods, and services under contracts whose estimated value exceeds specified monetary amounts.

For each PforR operation, the Bank carries out a process of identification, preparation and assessment, appraisal, and implementation support. Appraisal of each operation is informed by assessments in three areas that are then applied to the overall program and its expenditures. The *technical assessment* focuses on the strategic relevance and technical soundness of the program and its expenditure framework, the results framework, and the M&E arrangements. The *fiduciary assessment*, covering the procurement and financial management arrangements, seeks to make sure that program funds are used appropriately. Both national and international competitive bidding systems are assessed. The *environmental and social systems assessment* seeks to make sure that the potential environmental and social impacts and risks are adequately addressed. These assessments are expected to identify measures to enhance performance, build capacity, and mitigate key risks, and are reflected in an *integrated risk assessment*. The resulting Program Action Plan (PAP) is then reflected in the legal agreement between the Bank and the government. A central focus of preparation is the identification of disbursement-linked indicators (DLIs), each with a verification protocol to ensure that a credible mechanism is in place for monitoring and verifying its achievement.

During implementation, the Bank task teams are expected to monitor overall program progress, associated expenditures, and the achievement of results (including the DLIs). Task teams monitor progress in implementing the PAP, changes in the program's risks, and compliance with the provisions of the legal agreements. Technical support from the Bank team focuses on improving systems performance and resolving implementation issues. Operations are subject to the same corporate oversight functions as other Bank lending instruments, and the Bank retains the right to carry out investigations that it deems necessary and to sanction entities that are found to have engaged in fraud or corruption. The Bank's debarment list applies to PforR operations.

The monitoring and verification of results is an essential feature of the instrument, and the DLIs require a credible verification process that is acceptable to the Bank and is agreed at the time of appraisal. Where appropriate, the verification process may also include independent or third-party monitoring. DLIs are public information and their progress is supposed to be reported in the implementation reports. PforR documents are available to the public, giving stakeholders access to information about the performance of the public institutions and programs.

MANAGEMENT'S TWO-YEAR REVIEW

The OPCS *Two-Year Review* (World Bank 2015e) addressed the initial experience of PforR operations from the management perspective. It summarized the Bank's experience, drawing on a literature and desk review organized thematically and

with participation of task teams and managers who had worked on the instrument, internal and external surveys, structured interviews with government officials and senior Bank managers, and consultations with key stakeholders.

The review concluded that the instrument had been successfully rolled out across a broad range of countries and sectors, policy requirements had been met, and implementation was broadly on track for all but one of the approved operations. On this basis, the paper found that no major changes in the design of the instrument appeared necessary, but it identified some adjustments and areas for improvement. In this regard, the report recommended that the 5 percent cap on PforR commitments (as a percentage of total Bank commitments) be increased to 15 percent, although the volume of approved operations had not yet reached the 5 percent level, and that limited exceptions should be permitted to the procurement exclusions. If the 15 percent cap should be exceeded, management is expected to inform the Board and discuss how best to address the future demand. The paper also suggested other editorial and technical changes to Operational Policy (OP) 9.00, the Anti-Corruption Guidelines (ACG), and Bank Procedure (BP) 9.00. (OPs and BPs have now been replaced by Policy, Directive, and Procedure.⁴) The suggested changes were endorsed by the Board and put into implementation as of July 1, 2015.

The Evaluation

An independent assessment, as requested by the Board. This IEG evaluation provides an independent assessment of the early experience with the PforR instrument. It uses the findings of the *Two-Year Review* (World Bank 2015e), but goes beyond it in several respects, most importantly through a desk review of all approved PforR projects (until the end of fiscal 2015), supplemented by field visits to 11 countries (IEG 2015a).

Conceptual aspects of the evaluation. This is a real-time process evaluation of an instrument that has no completed operations, and thus no available IEG-validated data on final outcomes and only partial data on disbursements, results, and outputs. Through a review of all the PforR operations approved by the end of fiscal 2015 (27 operations), the evaluation has assessed project designs and the early implementation experience against the stated expectations in the 2011 Board paper (World Bank 2011a) and, in particular, the key issues highlighted in discussions at the Committee on Development Effectiveness and the Board: institutional capacity building, expected results, management of fiduciary risks, environmental and social aspects, and eventual evaluability.⁵

Evaluation objectives. This evaluation provides early feedback on how the PforR instrument is working and its adherence to the stated objectives and intentions in the Board paper (World Bank 2011a). It has two objectives: (i) to assess the early experience with the design and implementation of PforR operations and (ii) to provide lessons and recommendations relevant for the use and possible improvement of this instrument. Since this is a new instrument, the evaluation has also paid significant attention to possible risks – particularly the fiduciary, environmental, and social risks that have been in the forefront during discussions so far.

Specific audiences are the Board, which is guiding the rollout of this new instrument; Bank senior management, which is implementing the instrument; and country-client stakeholders, who would stand to benefit from the appropriate use of the instrument. The evaluation is also expected to be helpful in informing future Board decisions on the further mainstreaming of this instrument. In addition, the evaluation should be useful for program task teams engaged in the identification, preparation, and supervision of PforR operations.

Evaluation questions. The report seeks to answer the following questions:

- What has been the overall experience to date with the design, preparation, and early implementation of PforR operations and the associated opportunities and challenges?
- What is the quality of the program assessments, including the technical, fiduciary systems, and environmental and social assessments?
- How effective has the PforR instrument (including policies, procedures, and guidelines, and their application) been in identifying, assessing, and mitigating critical risks?
- To what extent is the PforR instrument being used to strengthen national systems for financial management, procurement, environmental and social safeguards, and M&E?

Analytical framework. The evaluation consists of the following activities:

- **A literature survey of relevant material concerning the results focus for development activities, the use of country systems, and political economy issues.** This included a review of literature from or about relevant program initiatives of other multilaterals (in particular, the Inter-American Development Bank [IDB] and the Asian Development Bank [ADB]) and leading bilateral aid agencies (including the United Kingdom’s Department for International Development [DFID]).

- **Systematic assessment of approved PforR projects.** This covered all 27 operations approved by the end of fiscal 2015 – a review of the full universe of operations was possible and desirable because of the small number of programs. The assessment was done as a desk review by the evaluation team using a standardized approach to focus on the new aspects of PforR operations and their implications. (Appendix A describes this standardized approach.)
- **Field visits carried out in eight countries with PforR operations approved in fiscal 2012–14 (Bangladesh, Brazil, Croatia, Ethiopia, Morocco, Mozambique, Rwanda, and Vietnam).** There were also brief, abbreviated visits to three countries (Kenya, Nepal, and Uganda), undertaken in the context of other tasks. The main purpose of these field visits was to seek the views of government counterparts and other relevant donors on various aspects of their PforRs – what works well or not so well, and how they see advantages and disadvantages compared with other Bank lending instruments.

Limitations of the evaluation. The newness of the PforR program and its short implementation experience set clear limitations for the evaluation. A few operations have had significant results during the evaluation period, but none has been completed.⁶ This process evaluation, therefore, cannot discuss final results. At the same time, this limitation enables the evaluation to focus on the assessment of the objectives, structures, and risks identified and addressed during preparation and the early performance of the approved operations (without the possibility of biased judgment of design based on knowledge of the eventual results). The evaluation also devotes special attention to the quality, relevance, and use of results frameworks and DLIs, including verification protocols. In addition, the evaluation looks at the strengths and risks of the PforR programs and reviews the extent to which PforR operations are helping to build national systems for the monitoring of results.

Organization of the report. The rest of the report is organized into five chapters. Chapter 2 provides a response to the first evaluation question: What has been the overall experience to date with the early implementation of PforR operations? Chapter 3 summarizes the findings on quality of program assessments. Chapter 4 provides a response to the third question: How effective has the PforR instrument been in identifying, assessing, and mitigating critical risks? Chapter 5 looks at the extent to which the PforR instrument is being used to strengthen national systems for financial management, procurement, environmental and social safeguards, and M&E. The concluding chapter summarizes the key findings and provides recommendations for strengthening the PforR instrument.

¹ The request was made at the meeting on October 24, 2014, and reiterated at the discussion of the paper by the full Board on April 9, 2015.

² The Board paper (World Bank 2011a) gave several examples of how it was possible to use the IPF instrument for programmatic purposes, but this required modification of counterpart procedures, exclusions of significant classes of expenditures, drawing significantly on preparation resources to find acceptable ways of adapting to prescriptive IPF requirements, and high expected costs of implementation support.

³ An additional implicit assumption is that Bank staff will not succumb to the usual disbursement pressures, and hence conclude that inadequate systems are “adequate.” See IEG 2015b.

⁴ In August 2013, the Executive Directors approved a Policy and Procedure Framework to replace the OPs/BPs that governed existing Bank operations. (“Bank Procedure: Policy and Procedure Framework” [R2013-0156; IDA/R2013-0214] (World Bank 2013a), approved by the Executive Directors on August 9, 2013.) With respect to PforR operations, OP and BP 9.00 have been replaced by “Bank Policy, Program-for-Results Financing” (World Bank 2015b), issued and effective July 10, 2015, Catalogue Number OPCS5.04-POL.01; “Bank Directive, Program-for-Results Financing” (World Bank 2015a), issued and effective July 10, 2015, Catalogue Number OPCS5.04-DIR.01; and “Bank Procedure, Program-for-Results Financing” (World Bank 2016), issued December 31, 2015, effective January 1, 2016, Catalogue Number OPSVP5.04-PROC.62. These new requirements are supported by the *Program-for-Results Financing: Interim Guidance Notes to Staff on Assessment* (World Bank 2012c).

⁵ In accordance with the Board paper, all PforR operations will be evaluated upon completion (separately from the present study). These self-evaluations by Bank teams (to be followed by IEG validations and selective Project Performance Assessment Reports) will analyze the results of each operation as well as efforts to build institutional capacity (building on the various assessments and their outcomes). The self-evaluations will also look at the performance of the Bank and the borrower in carrying out their respective roles.

⁶ Only two PforR operations will be closed by June 30, 2016: the Morocco Human Development Project and the Uruguay Road Maintenance Program. Implementation Completion Reports for these operations will be available by December 2016.

2. Experience with the PforR Program

After its approval by the Board, the Bank rolled out the PforR program cautiously, seeking a balance between responding to the potentially strong client demand and learning from implementation experience. There was also a limit to aggregate PforR operations, determined by the Board, of 5 percent of total International Bank for Reconstruction and Development (IBRD) and International Development Association (IDA) commitments for the first two years from Board approval. The Board reviewed the progress in rolling out the instrument as part of the *Two-Year Review* in 2015 (World Bank 2015e) and increased this limit to 15 percent of the three-year average of total IBRD and IDA commitments. The PforR portfolio has grown rapidly, and with significant design coherence that is broadly in line with the original expectations.

Growth and Diversification of the PforR Portfolio

From a modest start, the Bank's PforR portfolio has grown rapidly. As of March 31, 2016, the Board had approved a total of 39 PforR operations, providing \$9.5 billion of Bank financing to support a total of \$49.9 billion in government programs (table 2.1), with an additional 21 operations under preparation (having completed the concept stage), totaling \$5.4 billion in expected Bank financing.

Table 2.1. IBRD/IDA Lending, Total and for PforR Operations, Fiscal 2012–16 (US\$, billion)

Commitments	FY12	FY13	FY14	FY15	FY16 ^a
PforR–IBRD lending	0.3	0.1	0.5	0.9	3.3
PforR–IDA lending	0.1	0.7	1.2	1.3	1.1
Total PforR–IBRD/IDA lending	0.4	0.8	1.7	2.2	4.4
IBRD lending	20.4	14.8	18.2	23.1	23.6
IDA lending	14.3	16.2	21.3	18.4	6.8
Total IBRD/IDA lending	34.7	31.0	39.5	41.5	30.4
Total PforR as percentage of total IBRD/IDA lending	1.2	2.6	4.3	5.3	14.5

Source: Business Intelligence as of April 28, 2016.

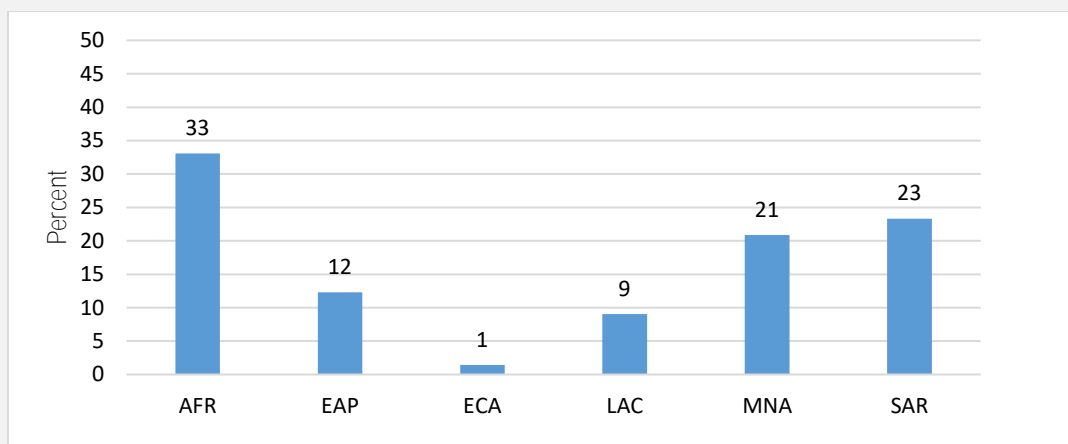
a. Commitment amount for fiscal 2016 is as of March 31, 2016.

As of March 31, 2016, all Bank Regions had at least two approved PforR operations (figure 2.1A). The introduction of PforR to different Regions appears to be influenced by a range of factors. In Africa, for example, it has been influenced by the

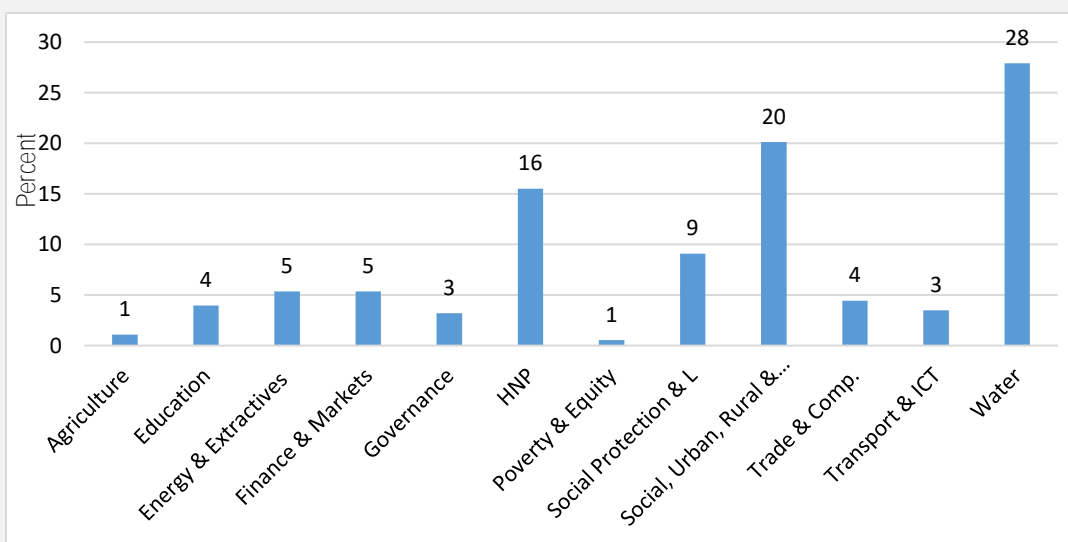
perspective that the PforR is a good instrument for supporting the regional agenda of building stronger institutions and delivering better services. In the Middle East and North Africa and East Asia and Pacific Regions, the instrument is regarded as a good fit for efforts to increase the emphasis on results and institutional capacity building. PforR operations cover most of the sectors in which the Bank traditionally provides financing. In terms of global practices (sectors), Water; Social, Urban, and Rural; and Health, Nutrition, and Population are the leading practices (figure 2.1B).

Figure 2.1. The PforR Portfolio (percent of operations)

A. BY REGION (BY AMOUNT AS OF MARCH 31, 2016)



B. BY GLOBAL PRACTICE (BY AMOUNT AS OF MARCH 31, 2016)



Source: World Bank Data Warehouse.

Note: By amount as of March 31, 2016. AFR = Sub-Saharan Africa, EAP = East Asia and the Pacific, ECA = Europe and Central Asia, LAC = Latin America and the Caribbean, MNA = Middle East and North Africa, SAR = South Asia. HNP = Health, Nutrition, and Population; ICT = Information and Communication Technology.

Early approvals were largely for IDA countries, with a recent increase for IBRD countries. During consultations on the design of the instrument, many observers expected the strongest uptake to come in IBRD countries. However, of the 27 operations approved by end of fiscal 2015, IBRD countries accounted for only 8 operations (35 percent of total PforR Bank financing), while 19 operations (65 percent by amount) were under way in IDA countries (see appendix C, table C.1). In the first three quarters of fiscal 2016, there was a stronger uptake of the instrument for IBRD lending, with 6 out of 12 approvals, accounting for around 75 percent of total Bank PforR financing (see appendix C, table C.2), and with the IBRD accounting for a majority of the operations in the pipeline (17 out of 21 operations).

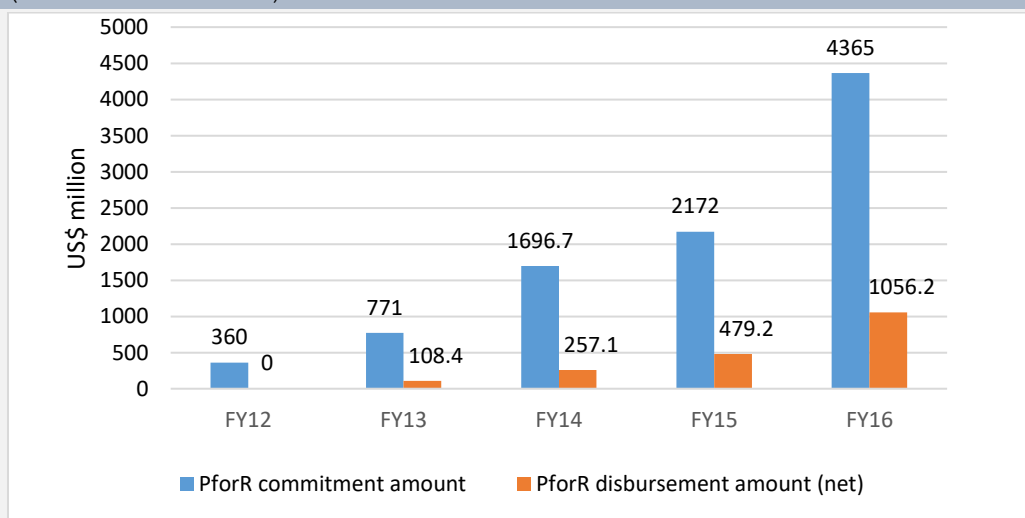
Some PforRs were combined with IPF lending. The PforR policy framework provides flexibility through the possible combination with other Bank instruments. Thus 3 of the 27 operations approved through the end of fiscal 2015 included IPF components for technical assistance activities.¹ During interviews, task teams indicated that they like having this hybrid option available. This approach, however, can create some complexities, because it requires the use of a separate set of policies and procedures. (This is consistent with the Bank's earlier experience with hybrid operations, when using IPF and DPF for the same operation.) For example, in Brazil, the technical assistance IPF incurred a multiplicity of procurement processes for capacity building and institutional development programs, with 14 of the agencies participating in the 4 programs supported by the PforR. This overwhelmed the capacity of both the project management unit and the Bank team, with attendant delays, including a delay in the achievement of three of the DLIs.

The programs do a good job of covering both partnerships and ownership. The IEG team rated 82 percent of the operations approved through the end of fiscal 2015 as moderately satisfactory or better on partnership relationships with donors – for example, that previous relationships have been continued under the PforR instrument, and/or that donors are working in parallel with the Bank's instrument, when relevant, and all but one was so rated on borrower ownership. (See appendix A for methodology.) In one case – Rwanda Agriculture – to enhance ownership by all stakeholders, the government signed a memorandum of understanding with the private sector, civil society, and development partners supporting the principles and objectives of the program, building on previous work. This program also includes cofinancing by DFID through a multidonor trust fund, with common results and DLIs. Mozambique PFM had an extensive participatory process, building wide understanding and clarity of goals, and enabling greater commitment from officials and other stakeholders that have been part of the process of setting goals and program design. The processing of this PforR took a long time – time that was

needed to create a reasonable internal understanding and acceptance of the concept and the proposed modalities.

There is little formal cofinancing with other development partners. Only 2 of the 27 approved operations involved formal cofinancing of the PforR instrument using the DLIs – Rwanda Agriculture with DFID, which for formal reasons was very difficult to put together (see para 2.6) and Tanzania Education with the U.S. Agency for International Development (USAID). However, many PforR operations involve some sort of parallel financing with development partners, through various means, including partners’ financing of technical assistance requirements, or the Bank and other partners supporting the same government programs. For example, in Ethiopia, the Bank supports the Millennium Development Goals (MDGs) Performance Fund for the health sector together with nine other multilateral and bilateral partners. Based on estimates in the project appraisal documents (PADs), the share of Bank financing varies from 13 to 100 percent of the program in approved operations (appendix C, table C.2), but there are no indications that the PforR instrument has encouraged any other donor financing. Bank financing and development-partner financing as a share of these overall government programs averages 15 and 24 percent, respectively. The participation of other donors seems to have increased for operations approved in fiscal 2016 – six out of eight operations have financing from other donors (appendix C, table C.2).

Figure 2.2. PforR Commitment versus Disbursement
(as of March 31, 2016)



Source: World Bank Data Warehouse.

Approved PforR operations largely follow PAD expectations. As of March 31, 2016, 32 of the 39 approved operations were effective. Bank disbursements have been on track relative to expectations, with annual disbursement amounts increasing from

\$108 million in fiscal 2013, to \$479 million in fiscal 2015, and to \$1,056 million for the first three quarters in fiscal 2016 (figure 2.2). Of the 32 effective operations, 21 have been effective for over a year, and have disbursed 35 percent on average. Most of the effective operations have shown a steady pattern of disbursement, but a few either disbursed a lot in a relatively short period (the Tunisia Urban Development Project disbursed 0.1 million in fiscal 2015 as front-end fee payments, but \$65.5 million in fiscal 2016) or experienced difficulties in disbursing (the India Maharashtra Rural Water Supply and Sanitation Program [RWSSP], effective in August 2014, had 4 percent disbursement in early 2016). Overall implementation of PforR operations is also broadly on track, with performance ratings satisfactory or moderately satisfactory and risk ratings stable through early implementation, except the Moldova Health Transformation Project, where ratings for progress toward achievement of project development objectives (PDOs) and implementation progress have both been downgraded, from moderately satisfactory to moderately unsatisfactory, and the associated overall risk from moderate to substantial. Domestic political problems underlie many of these issues.

Thus far, PforR operations have shown themselves to be no less expensive for the Bank than IPF projects. Expenses directly related to the preparation of PforR operations average \$554,000, with a median of \$547,000, both somewhat higher than for IPF operations (with Track II regular processing requirements) in fiscal 2013 and fiscal 2014 (\$548,000² average and \$466,000 median). Average supervision costs for the 27 programs were \$261,000 for fiscal 2015 and \$180,000 for fiscal 2016 (through April for the first nine months of the fiscal year), but with wide differences between high- and low-cost programs (depending, in part, on the amount of continued Bank engagement with institution building during implementation), and thus with the median numbers a bit lower. These supervision costs are significantly above the Bank-wide average for IPFs of about \$110,000, although it is possible that the average PforR preparation and supervision costs will decline somewhat over time as programs mature. It is likely that introducing the instrument to clients and the learning process for the clients and Bank staff contributed to relatively high early costs for preparation and implementation support. In addition, Bank PforR work includes the preparation of the fiduciary and environmental and social assessments. The Bank's work during both PforR preparation and implementation can thus involve an amount of public goods work, with potential benefits beyond the specific programs and participating institutions. Overall, however, there is not yet sufficient evidence to derive any conclusions about the overall efficiency of PforRs.

The average Bank financing of a PforR operation has been about \$241 million, larger than the average IPF operation. There may be potential for further increases, since there may be potential for supporting larger programs in some countries, and since

some borrowers welcome what they see as the predictability and general nature of disbursements under PforRs. Operations approved in fiscal 2016 have an average of \$363.8 million in Bank financing, which is significantly larger than those approved between fiscal 2012 and 2015 (\$185.2 million, on average), an average increased, in particular, by a \$1.5 billion operation for India.

As reported in the *Two-Year Review* (World Bank 2015e), the average preparation time for a PforR was 13.6 months, which is slightly shorter than the 15 months for Track II IPF operations in fiscal 2014, although actual preparation time ranged from 6.9 to 28.7 months. Short preparation times were feasible primarily when working with established government programs and where governments themselves focused on results, as was the case for the first two PforRs in Rwanda.

PforRs account for a growing share of the Bank's operations. Table 2.1 shows that since their inception, there has been a steady increase in the share of PforR commitments in Bank approvals, from \$0.4 billion in fiscal 2012 (1.1 percent of overall Bank approvals) to \$2.2 billion in fiscal 2015 (5.3 percent). In addition, as of March 31, 2016, 12 fiscal 2016 operations had been approved by the Board, totaling \$4.4 billion in commitments, or 14.5 percent of the total Bank lending for this period. There is also a marked tendency that when a country has one PforR operation, more are likely to follow. As of end fiscal 2015, almost half (13 of the 27 operations) were in six countries with two PforRs, and, in one case, with three operations (appendix C, table C.1), and these countries also accounted for one-third (4 out of 12 operations) of the approvals in the first three quarters of fiscal 2016, while three countries accounted for their second PforRs in this period. Based also on the findings of the field missions, IEG expects that PforRs will – over time – account for substantial shares of the overall lending programs in a number of countries.

Design Coherence

RESULTS FRAMEWORKS AND DLIS

The 2011 Board paper (World Bank 2011a) on PforRs makes clear that the programs are aimed at achieving results, but otherwise the paper does not provide much discussion of what should be understood by results, beyond a brief discussion of the wide range of possible DLIs. The *Guidance Notes* stipulate that results are “the output, outcome, or impact of a development intervention. In general, the Bank seeks to encourage results that support sustainable improvements in country outcomes – that is, evident changes in peoples' lives, and/or the behaviors of targeted organizations, households, or firms” (World Bank 2012c, p. 16). During program preparation and, in particular, implementation, the DLIs may get the most

attention since their progress will determine the pace of disbursements. However, as discussed in chapter 6, a comprehensive assessment of the eventual results of a PforR must cover several dimensions, with the PforR results framework at the core.

The PforR frameworks consist of an overall PDO, a few PDO indicators, and a larger number of intermediate results indicators. Together, these three components constitute the results framework for a PforR. Although there is usually good coherence among the three, in some cases the indicators are not coherent with the PDO, or the PDO is not fully aligned with PAD discussion or the indicators.

The PforR objectives are expressed in the PDOs, supplemented by the few PDO indicators for each PforR. For the PforRs approved by the end of fiscal 2015, the targeted results are mostly institutional, such as “To improve transparency and resource management of targeted departments of the Province of Punjab,” in the Pakistan Punjab Public Management Program, or they represent intermediate outcomes, such as “To improve the delivery and use of a comprehensive package of maternal and child health services,” in the Morocco Improving Primary Health Care in Rural Areas PforR. Objectives that lead to clear results at the outcome level are rare. PADs could benefit from pushing the envelope further in terms of concentrating on outcomes that are achievable during the program period (such as increased immunization coverage or improved reading levels). There is seldom a discussion of the longer-term prospects and objectives for the programs, and thus also relatively little attention to the sustainability of the targeted improvements beyond the program periods.

Issues with the results frameworks are not limited to PforRs, but are also found in other instruments. Thus, as noted in *The Quality of Results Frameworks in Development Policy Operations* (IEG 2015d), some development policy operations (DPOs) suffer from a lack of clear statements of objectives and outcomes. Their results frameworks lack explicitly stated outcomes, while results often fall short of meaningfully measuring the impact of a DPO. In some cases, the identified prior actions lack significant additionality, because they have little or no tangible implications for overall policy, involve recurrent government functions, or pilot actions without a clearly defined scaling-up strategy. In other cases, the results frameworks lack essential triggers or accept partially met triggers that do not capture the true essence of intended reforms and may substantially undermine the quality of the DPO.

The PforR instrument finances the implementation of a program, and the support for program implementation should, in turn, determine the choice and design of the DLIs, but the DLIs should not be considered the results of a PforR. Unlike output-based aid, which directly links financing to outputs or outcomes (that is, program or

project implementation is financed elsewhere), the design of a PforR must provide a way to ensure that there is a sufficient flow of funds during program implementation through the DLI disbursements. In some PforRs, IEG found quite close and clear integration between the results framework and the DLIs (and with the PAP). One example is the Ethiopia Health MDGs PforR (box 2.1).

Box 2.1. Scope and Results Supported by the Ethiopia Health MDGs Program-for-Results			
Priority area	Activities	Intermediate results	Outcomes
Accelerating progress toward maternal health MDG	<ul style="list-style-type: none"> • Supplying equipment and commodities for providing emergency obstetric care • Supplying contraceptives • Providing ambulances to all <i>woredas</i> • In-service training of midwives and training of health officers in emergency surgical and obstetric skills • Capacity building of health extension workers in clean and safe delivery 	<ul style="list-style-type: none"> • Health centers offer basic emergency obstetric care • <i>Woredas</i> have functional ambulance services • Midwives receive in-service training • Health officers trained in emergency surgical and obstetric care 	Increase in: <ul style="list-style-type: none"> • Skilled care at childbirth • Antenatal care • Contraceptive prevalence
Sustain the gains made in child health MDG	<ul style="list-style-type: none"> • Strengthening of cold chain systems • Supplying vaccines • Holding immunization campaigns • Supplying bed nets 	<ul style="list-style-type: none"> • Health centers have functional cold chain equipment • Outreach campaigns held • Long-lasting insecticidal nets distributed 	<ul style="list-style-type: none"> • Increased immunization coverage
Strengthen health systems	<ul style="list-style-type: none"> • Constructing health centers • Supplying essential medical products and equipment • Validating health management information system (HMIS) semi-annually • Undertaking surveys and studies 	<ul style="list-style-type: none"> • Health centers built • Health facilities report HMIS information in time • Annual facility readiness assessment undertaken 	<ul style="list-style-type: none"> • Improved HMIS • Rollout of balanced scorecard and institutional performance incentives • Improved facility readiness

Source: World Bank 2013b.
Note. The indicators in bold are DLIs in this operation. *Woreda* = district.

However, overall the DLI sets may be more or less representative of the results frameworks. In two cases (among the 27 in the IEG sample) the DLI set consists of all the indicators in the respective results frameworks, and they are thus fully representative of those frameworks. For some PforRs, the DLI sets are well structured and explained in relation to the results framework. However, there are other cases where the DLIs are focused on intermediate steps for which the

relevance and the linkage to the program objectives and results framework appear weak, or at least have not been made apparent. Box 2.2 describes one example.

Box 2.2. Example of a DLI Set Different from the Results Framework

In the Tanzania Big Results in Education Now Program, there is a clear difference between the results framework (PAD annex 2) and the set of DLIs (PAD annex 3):

- The PDO is to improve the quality of education in Tanzanian primary and secondary schools, and the results framework supports this through appropriate PDO indicators that include reading levels, subtraction levels, teachers found in classrooms, and knowledge levels of teachers.
- But the six DLIs are overwhelmingly (80 percent by amounts) institutional – completion of foundational activities, resource flows, annual reporting, and deployment of teachers across districts, and schools receiving incentive grants. Only one DLI addresses education results – for reading improvements.
- It is thus hypothetically possible for all DLIs to be met in full, and on time, with 100 percent disbursements, but with only one of the four PDO indicators (reading levels) being met. (However, management believes in this case that the DLIs will support the achievement of the PDO indicators).

Source: World Bank 2014d.

The extent to which DLIs incentivize performance may depend on the financing attached. The role of DLIs to incentivize performance is mentioned regularly in PforR PADs, in particular in the risk assessments. There can be two aspects to incentives. A well-structured set of DLIs responding to borrower priorities will help to encourage performance through a set of logical, achievable steps, and this has been observed by IEG field missions. But it is a separate question whether the prospects for an amount of financing will encourage prompt and timely actions – and the missions saw examples where such prospects did not appear to do so.³ A related question is whether small DLIs (in dollar terms) can have much of an incentive effect. As one example, for the relatively recent Vietnam National Urban Development Program, one DLI accounts for 1 percent of total financing, one for 3 percent, and two others for 4 percent each – at the other end of the scale, one DLI is expected to account for 62 percent of overall disbursements.

It is unclear whether DLIs represent stretch targets. The *Two-Year Review* did not discuss the degree of difficulty or stretch in the DLIs, but noted that the formulation of the DLIs (and the associated disbursement mechanisms) had been one of the most challenging aspects of the preparation of PforR operations (World Bank 2015e, para. 37). This evaluation has not been able to determine the degree of difficulty in the DLIs – more appropriately assessed ex post, but field missions noted several cases

where DLIs seemed relatively easy. As illustration, Box 2.3 offers one example of an apparently unambitious set of DLIs.

Box 2.3. Vietnam—DLIs in the Results-Based National Urban Development Program in the Northern Mountains Region

The program has eight DLI indicators, divided among four DLIs:

- **DLI 1.1: Enhanced annual city plans approved and disclosed to the public.** These plans to be produced annually. Presumably, once a plan has been produced it will be easy for the cities to modify it as appropriate in the following years.
- **DLI 1.2: Professionally staffed management units in place within each participating city People’s Committee.** Presumably, once a unit is in place it will be straightforward to maintain it.
- **DLI 2: Local urban infrastructure investments delivered as per each participating city’s approved enhanced annual city plan.** One of two DLI indicators concerning outputs on the ground.
- **DLI 3.1: Asset management plan adopted and local urban infrastructure subprojects in full service after completion.** This indicator also concerns outputs on the ground. For the first two years, the indicator concerns asset condition assessment and asset management plans. From the third year, condition that at least 80 percent of completed local infrastructure subprojects are free of physical damage and operational.
- **DLI 3.2: Increased annual own-sources revenue in participating cities.** The requirement (12 percent increase in nominal terms) may not appear difficult considering an inflation rate of 6-7 percent and a GDP growth rate of 7 percent in 2014, but at least one participating city is concerned because of limitations on their revenue-generation possibilities by central government policies.
- **DLI 4.1: Implementation strategy for National Urban Development Program adopted with annual milestones.** A set of gradual milestones, with approval of a policy note in year one and management capacity in place in year four.
- **DLI 4.2: Professionally staffed unit in place in the Ministry of Construction, preparation of annual capacity development plans, and capacity-building support provided to cities in accordance with such plans.** So the degree of progression under this DLI will depend on the degree of progression of the annual capacity building. No apparent progression in requirements after year one, but annual disbursements based on continuation of the requirement.
- **DLI 4.3: Completed program report.** An annual report.

Of the eight indicators, six are institutional (and therefore under the control of the central or city authorities), and two concern the investments on the ground, with a modest degree of apparent progression over the years.

Source: World Bank 2014e.

PROGRAM ACTION PLANS

The PAP is a key component of the PforR instrument. It contains a set of actions agreed with the government, based on the outcomes of the different assessments. Among the examples of improvements that the Board paper (World Bank 2011a) mentions for inclusion in the PAP are the following: actions to improve the technical dimensions of the program and formal rules and procedures governing the organization and management of the systems used to implement the program; actions to enhance the capacity and performance of the agencies involved; and risk-mitigating measures to increase the potential for the program to achieve its results and to address fiduciary, social, and environmental concerns. The implementation of all actions in the PAP is a contractual obligation incorporated in the Financing Agreements. However, such a general obligation can have little impact on a government's interest in executing individual PAP items.

In general, there is room to improve the discussion of PAP design in the PADs of the projects examined by IEG. The PforR guidelines state that a PAP should be selective regarding areas of focus (four or five areas), with a few specific actions in each. This implies a need to make trade-offs among recommended actions from the assessments. However, in some recent cases, IEG noted a tendency toward an increasing number of PAP actions, with 49 actions in one case (Ethiopia – Shared Prosperity Operation). It is likely that this will dilute the impact of the individual PAP items. In contrast, the 13 PforR programs visited had a total of 190 actions, of which 35 percent were fiduciary, 40 percent technical, and 25 percent environment and social. About 40 percent of the actions in the PAPs reviewed were actions to enable the achievement of DLIs or were incorporated in the DLIs – it is not clear why it is considered necessary to include such actions (or actions also included separately in the Financing Agreements). Most PADs identified the PAP actions as “critical” or “key” to the achievement of program objectives or to mitigating risks. However, in some cases, the PAPs included “opportunities for strengthening” institutional capacity, where the actions may not be critical to the program. Also, there are PADs that do not provide a clear basis for inclusion of actions in the PAPs, specifically the articulation of trade-offs among the actions recommended in the assessments.

IEG found the overall quality of PAPs at entry to be satisfactory. IEG rated 6 of the 13 programs visited as satisfactory, and rated four moderately satisfactory for “appropriateness of PAP and conditionalities.” A related quality measure in the IEG template is “appropriateness of risk-mitigation measures,” many of which were supported by the actions in the PAPs – IEG rated nine programs satisfactory and four moderately satisfactory using this measure. (See appendix A for methodology.)

Implementation performance has been more uneven, however. Based on the IEG program assessments, the implementation performance of PAPs has not matched the quality at entry. Of the eight reviewed PAPs with implementation ratings, two were rated below the line, while the majority (five programs, accounting for about 30 percent of total actions) of the PAPs were rated moderately satisfactory for implementation. One reason may be that, in practice, there seems to be no penalty for slow or poor implementation of a number of PAP items, except for the actions linked to DLIs and effectiveness conditions. Another reason may be the inclusion of actions that are not critical to the achievement of program objectives. If the PAPs were to focus on (non-DLI) critical actions, there may need to be greater clarity on the consequences of slow progress.

Reporting on individual PAP actions is not systematic. The Implementation Status and Results Reports (ISRs) typically do not provide information on the status of PAPs, mainly because the format does not allow for it except through notes and comments. In the aide-memoires, there is variability in the quality of the reporting on PAPs – discussion of PAP actions is sometimes spread across the various topics covered in the text. A good practice found in half of the reviewed PforR aide-memoires is the use of a table or annex that summarizes the status of each of the PAP actions and the next steps and consolidates the information on PAP actions discussed in various sections of the aide-memoire.

There may be room for greater selectivity and focus in designing the PAPs. Generally, the PADs should do a better job of articulating the rationale for the choice of actions included in the PAPs. This would result in a smaller number of more critical actions. For example, the PAPs could give attention primarily to areas with substantial and moderate risks. In addition, it may not be necessary to include DLI-related actions, which are already being tracked through the DLIs.

Emerging Messages

The PforR instrument is an increasingly important lending vehicle for the Bank. The PforR portfolio has been growing steadily and is becoming an increasing share of the Bank's operations. There are now PforR operations in all Bank Regions and in most of the key sectors. Approved operations are largely on track. Early approvals were largely for IDA countries, but there has been a recent increase in IBRD country participation, with IBRD countries accounting for a strong majority of the operations in the pipeline. The programs do a good job of covering ownership by counterpart stakeholders, which is essential for development effectiveness. There has been a pronounced tendency for a country with one PforR to be interested in more such operations.

The PforRs also cover partnerships well, but have resulted in little formal cofinancing with other development partners. While many PforR operations involve some sort of parallel financing with development partners, there is little evidence that the PforR instrument has been conducive to more or stronger partnerships – and the theory of change may thus not have been fully met. Only 1 of the 27 operations approved through the end of fiscal 2015 has involved cofinancing of the PforR instrument using the DLIs, and this was very difficult to put together for formal reasons. Partners who cofinance and participate in the disbursement aspects of a PforR operation may face challenges because their internal procedures and practices relating to the flow of funds may need to be modified to participate efficiently and effectively in the DLI aspects of an operation. Each potential partner will know best how to adjust its own systems, but for the blending of the systems to work smoothly, it is important that the Bank team be able to explain clearly and accurately the details of the DLI procedures at the outset. Such clear information will alleviate any lack of understanding and will allow a partner not only to tweak its own funding mechanisms, but also to keep its expectations in line with the realities of the new instrument.

The results frameworks, DLIs, and PAPs are often reasonably coherent, but there is room for improvement in several areas. PforR objectives in the projects examined are mostly institutional or focused on intermediate outcomes. Objectives in terms of outcomes are rare, and the focus in the PADs on objectives and supporting discussions could go beyond the program period itself to consider longer-term objectives and sustainability of program improvements. Some DLI sets do not seem to represent adequately the results frameworks that they are supporting, while others may be unduly modest, probably to ensure steady progress and disbursements. The incentives among staff and government counterparts are largely for achieving the DLIs; the risk-mitigation activities captured in the PAP supervision receive less attention. There is also room for greater selectivity and focus in designing the PAPs, including a reversal of the trend in some PforRs toward an increasing number of PAP actions. Also, the PAP implementation performance has been somewhat uneven, and the reporting on individual PAP actions has not been systematic.

¹ Brazil Strengthening Service Delivery for Growth, Poverty Reduction, and Environmental Sustainability in the State of Ceará Program, Mexico Oaxaca Water Supply and Sanitation Sector Modernization Program, and Moldova Health Transformation Program.

² Some operations were converted from other instruments; for those for which teams did not change the code, the costs were broken down, and the numbers reflect those related to PforR aspects only.

³ One reason can be that timely disbursements may have different effects at different government levels (for example, at a Ministry of Finance compared to a Ministry of Health). There may also be circumstances where the timing of disbursements may not be all that important, as long as the funds are likely to be disbursed during the life of the program.

3. The Quality of Program Assessments

The Bank carries out three kinds of assessments for each PforR. These are technical (including expenditures and results), fiduciary, and environmental and social assessments. Each of these should consider the overall quality of a program (including its expenditures) and its systems in its area – the framework and rules, implementation capacity, and actual performance under the program – and should identify improvements, including measures to mitigate major risks that the program will not achieve its expected results. The program assessments for the 27 operations approved by the end of fiscal 2015 have been comprehensive in scope and particularly strong with respect to addressing each program’s technical soundness, implementation capacity, fiduciary, and environmental and social impacts.

Technical Assessments

The technical assessments are a critical step for the processing of PforR operations. They consider strategic relevance, appropriateness of program structure and implementation arrangements, the expenditure framework, and the economic justification. They also cover the quality of the results framework for M&E (see chapter 2). The assessment is meant to identify strengths in the government programs and how they will be used for leverage, and weaknesses and risks in the government programs, and how they will be addressed or mitigated as needed by the PforR in an integrated and coherent manner (see World Bank 2011a).

Overall, these assessments are credible and comprehensive. IEG’s review of the PforRs approved through fiscal 2015 found that they were mainly prepared in collaboration with clients and, in some cases, with other development partners. They place the PforRs in the context of the Bank’s country programs and demonstrate the importance of the sectors for the countries’ development. Strategic relevance is generally good. The IEG team rated only 3 out of 27 PforRs as moderately unsatisfactory in this regard (with the other 24 rated from highly satisfactory to moderately satisfactory). Even these three operations were generally consistent with country and country assistance strategy priorities, but were found to have weaknesses such as the balance of program priorities or design.

The IEG team rated all but 1 of the 27 PforR operations moderately satisfactory or better on program structure and implementation arrangements. Operations supported government programs, which are well defined and clearly bounded, with adequate borrower capacity, borrower participation in design and implementation,

defined program management mechanisms, and coordination processes with other development partners.

Expenditure frameworks are often covered reasonably well. The IEG team rated the frameworks of 89 percent of the operations as moderately satisfactory or better, with only three operations rated below the line. Assessments look at the efficiency and effectiveness of a government program, along with its sustainability and linkage with priorities and results. For example, Kenya's Medium-Term Expenditure and Financing Framework supports the expansion of safety nets, with government resources gradually increasing the country's share of total program financing relative to donors. In the Rwanda PFM, the PforR would finance 58 percent of the overall program, with contributions from government and other development partners. However, here and elsewhere, tracking government expenditures can be difficult because of budget classifications and coding.

The role of PforRs can be unclear where they finance modest shares of overall programs. As one example, in Tanzania Education, the program expenditures seem reasonable, with a results focus incentivized by DLIs, and a useful typology of costs provided in the PAD (although the financing of technical assistance could have been broken out more clearly). However, the program will account for less than 10 percent of total basic education expenditures. The more than 90 percent of spending outside the program, depending on circumstances, may have greater influence over educational results. In this case, is it also not clear whether the program will be sustained after its completion, or even whether the program (if successful) ought to be sustained for the long haul, separate from the overall education programs.

Some operations may be too ambitious. Thus, the Mexico Water and Sanitation Program aims to triple the annual investment level of the sector in Oaxaca State, which the PAD itself deems "very challenging in a context of limited counterpart financing and limited capacity to absorb such an increased amount of investments" (World Bank 2014c, p. 6). While extensive technical assistance will be provided, such a rapid increase seems overambitious when considering the more rigorous controls required by the program and the sector's poor past performance record.

Some operations may fail to address key sectoral issues. Thus, the Tunisia Urban Program supports reforms in the public expenditure framework to make the capital investment allocations from the central budget to local governments more transparent, predictable, and efficient, and the government has passed several legal changes in that regard. However, the trajectory of growth of allocations to local governments outlined in the program is consistent with historical trends, which means that levels of investment in municipal infrastructure will remain modest and

below expected requirements. In addition, the program does not address the modest growth in municipal revenues, which fund operations and maintenance of infrastructure and are critical to sustainability and effectiveness. It is also unclear how the PforR program supports the local governments that are in financial distress; however, the program and the underlying policy reforms seek to address the causes for such distress.

Costing methodologies in many programs are unclear. A common shortcoming is that there are few details in the PADs on the costing methodology used by the government to prepare the expenditure frameworks for the programs being supported. Without accurate costing, budgeting, procurement, and performance measurement can be compromised. Most PforRs seem to base their costing estimates on the costing of the underlying government program, but with few or no details on the methodology used by the program. An exception and example of good practice is the Ethiopia Health Millennium Development Goals PforR, where costing was done using Marginal Budgeting for Bottlenecks, a tool created by development partners and the national ministries of health from several countries (box 3.1).

Box 3.1. Good Practice in Costing: Ethiopia Health Millennium Development Goals PforR

Costing for the Ethiopia Health PforR was done by the government using the Marginal Budgeting for Bottlenecks tool, which was created by development partners and the national ministries of health from several countries. The tool contributes to the removal of health system bottlenecks by helping to prepare strategic plans and expenditure programs to increase the quality of high-impact health, nutrition, HIV/AIDS, tuberculosis, and malaria interventions. The approach focuses on three service-delivery modes: family-oriented, community-based services; population-oriented schedulable services; and individual-oriented clinical services. It helps to assess current performance of health services and to identify bottlenecks in both supply and demand, and it enables a tailored approach to each country's specific situation. This helps to facilitate selection of the types, quantities, and costs of salaries, drugs, training, and other inputs that are needed to overcome bottlenecks and achieve optimal results.

Sources: UNICEF, World Bank, and African Development Bank 2011; World Bank 2013b.

The economic justifications in the PADs are broadly reasonable. The IEG team rated 89 percent of the economic justifications moderately satisfactory or better. Program teams looked at the rationale for public financing, the economic impact of the program with and without PforR support, and the added value of the latter. For example, the economic evaluation of Croatia Health estimated the economic benefits from seven key interventions supported by the PforR. The technical assessment used

evidence from other country experiences in estimating cost savings from a variety of activities, such as the outpatient versus inpatient costs of various procedures.

The PforR programs often do not estimate economic or internal rates of return, instead using other or partial measures. Thus, for Rwanda PFM, the PAD cited research by the Overseas Development Institute pointing to general economic benefits of PFM reforms: macroeconomic stability, timely and reliable fiscal and financial information, service delivery enhanced by regular payment of salaries, and state building. (However, the PforR results framework is not set up to measure most of these aspects.) The economic evaluation of Pakistan (Punjab) Public Management is generally reasonable. Expected benefits during the program's lifetime – increases in property tax collection and labor productivity and reductions in transaction cost for accessing services – are valued at 1.7 times the present value of the cost of intervention. However, the property tax estimates assume a level of enforcement for ensuring that tax bills are paid for new properties added to the registry, despite the huge challenge of nonpayment of taxes in Pakistan. The economic evaluation gives a range of estimated benefits (low, medium, high), but it is not clear to what extent property tax payment compliance has been factored into these estimates, and the benefits achieved will be, in part, the result of other initiatives, so they cannot be attributed entirely to the program. In addition, the PAD points out that it underestimates the net present value, since it does not take into account all the benefits that program interventions will bring about through subsequent systems improvements.

Fiduciary Assessments

The Bank has for some time been committed to increasing the use of country systems. Country systems would include, for instance, designation of the project as on-budget, aligning with the fiscal calendar, relying on the range of national budget preparation and execution procedures, as well as intergovernmental transfers, audit, and national competitive bidding procedures for procurement. PforRs are part of this long-term attempt to design operations that are a better fit for country contexts. The PforR approach to country systems is different from the Bank's earlier country systems pilots. First, it uses the systems that are responsible for the program being supported, which may be at a higher standard than those for other programs. Second, it does not insist that these program systems be at the same standard as the Bank's policies and procedures for investment lending operations.

Overall, IEG finds that the fiduciary assessments have been comprehensive, and that they cover most of the key aspects that one would expect to find. Each area of

discussion typically begins with an overview of the laws and processes applied in the relevant areas of the government’s program, the shortcomings in practice, and how the shortcomings will be addressed and monitored. Most PforRs have also drawn from extensive previous analytical work on fiduciary systems, including Public Expenditure and Financial Accountability (PEFA) Assessments, Country Procurement Assessment Reports (CPARs), and a range of fiduciary assessments done by development partners in connection with budget support operations. The assessments are generally thorough in identifying relevant transparency measures, such as timely provision of information to stakeholders, disclosure of tender notices and award decisions, and parliamentary oversight. But some contextual or political economy issues may not be addressed, as discussed below, and the assessments are generally stronger for the formal systems (such as procurement laws and regulations) than for the practical realities (such as actual fraud and corruption).

PforRs assess most program procurement systems as consistent with good public procurement principles, although implementation is often uneven because of a lack of political commitment, rigorous follow-up, trained staff, and effective demand-side processes. Challenges are sometimes pointed out without corresponding actions being proposed, such as possible entry barriers for contractors due to their mandatory registration.

The fiduciary systems used for PforRs are broadly the same as for other areas of the government programs. One exception is the PforR for Brazil – Strengthening Service Delivery for Growth, Poverty Reduction and Environmental Sustainability in the State of Ceará, where technical assistance procurement (10 percent of Bank-funded project costs) uses Bank rules because local rules were not thought to give enough attention to quality. However, IEG’s field mission found that the technical assistance component had become unduly complex and had turned into a source of implementation delays. In hindsight, it might have been more efficient for the Bank to have provided some initial capacity building and training for the government’s own procurement staff in participating agencies to implement their own “technical quality + price” methodology. Although rarely used, this methodology had been legally available from the start.

Another exception is the high-value procurement exclusion, which has reportedly excluded some financing in areas supported by the PforRs. PforR financing cannot normally be used for procurement packages of high-value contracts above thresholds based on the type of procurement and the level of fiduciary risk. Following the *Two-Year Review* (World Bank 2015e), a provision was added that such contracts may be financed if they are important to the integrity of the program and their cost is less than 25 percent of the overall program cost. This exception needs to

be approved by the Bank's managing director and chief operating officer. In the absence of such exceptions, high-value procurements may be handled outside the programs, and thus outside Bank oversight. IEG finds that this exclusion has not been a major obstacle in the PforRs reviewed, in part because these areas can be financed from other sources. The PforR fiduciary risk is then transformed into a funding risk to the non-Bank-funded part of the program, which could, in turn, become a risk to the development objective supported under the program.

It can be problematic to include high-value contracts under the PforR instrument. High-value procurements necessarily impose special risks, and they are likely to create more work for the clients as well as the Bank. One example is the Bangladesh value added tax (VAT) Improvement Program, which included financing for a large information technology system. The procurement of this system had been prepared carefully and in great detail. As agreed, the hardware and software components were initially bid as separate procurements, using international competitive bidding procedures. However, the Cabinet Committee on Government Purchase decided not to accept the successful bidders from the initial procurement, as recommended by the Technical Committee, to cancel the initial tender, and to retender, merging the two components. The resulting new tender was for \$29 million, above the \$20 million threshold allowable under the program. As a result, and following lengthy discussions, the full component (hardware and software) will now be financed by the government, while the PforR will support the implementation of the system in other ways. In this case, a hybrid approach may have been better, with the IPF instrument financing this component.

Some borrowers have raised concerns about the ACG. Under PforR financing, the borrower is responsible for taking actions regarding possible fraud and corruption, and the Bank can investigate allegations and sanction parties if appropriate, which has raised some borrower concerns. In Brazil, some activities were taken out of the scope of the PforR to avoid inconsistency at the local-government level. In Morocco, government counterparts repeatedly raised this issue with the IEG mission. They questioned why the Bank agreed to the use of all other aspects of national systems, but insisted on the application of its own ACG, without even assessing the quality of the country's own anti-corruption system.

Following the *Two-Year Review* (World Bank 2015e), the above rights and responsibilities did not change, but a provision was added that the Bank and the borrower will formulate a program-specific protocol on how to take actions. The protocol may be a legal agreement, a memorandum of understanding, or part of the negotiation minutes. In addition, new language was added, emphasizing that the country has the sovereign right to take actions in this area, and that Bank actions are

administrative, for the purpose of determining compliance with Bank policies. If any actions taken by the borrower conflict with the laws and regulations of the country, the Bank and borrower will consult to agree on alternative actions that will avoid such a conflict.

Political Economy Aspects

Political economy issues are behind many of the constraints raised in fiduciary assessments, but are rarely discussed directly. Any critical analysis tends to address only low-level issues. For example, in one case, “The membership of the district council internal audit committee cannot be easily controlled since they are elected politicians. This can be addressed by co-opting competent members based on an agreed sitting fee.” Through simple literature searches, the IEG evaluation found several examples where outside parties, such as researchers, had raised critical views about the actual state of procurement and other fiduciary practices that do not seem to be reflected in the PforR assessments, where the focus instead is typically on institutional reforms to address corruption, including program-specific measures. There are seldom any convincing arguments that such measures will work in the case of the new programs. Several interviews with team members confirmed their awareness of such information and issues; the lack of published analysis is reportedly because of borrower-government sensitivity in this area.

An example of good practice is the Mozambique PFM PforR. The PAD includes a good summary of the political economy challenges, the identification of winners and losers, and design features that address these challenges: strong, senior-level ownership across key ministries; a shift in incentives that creates greater benefits for cooperation; flexibility to account for unexpected impacts; the use of change agents at the deconcentrated levels to drive change; and sharing of risk between the Bank and the client. The design also addresses the complex political economy of medicine supply-chain reform, including interdepartmental coordination and competing policy aims.

Environmental and Social Systems

PforRs seek to ensure that the environmental and social effects of the program are adequately addressed, as stated in the fiscal 2012 Board paper. During the preparation of a PforR operation, the Bank is expected to assess, against the requirements of OP/BP 9.00, the degree to which a program’s systems manage and mitigate the environmental and social impacts of the overall program. The assessment will also identify and exclude high-risk activities, specifically those that

pose a risk of potentially significant adverse impact on the environment or affected people (activities classified as Category A under the IPF safeguards).

OP statement 9.00—Program-for-Results Financing, sets out requirements for assessing a program’s environmental and social systems. An assessment is expected to consider to what degree the program’s systems:

- Promote environmental and social sustainability by avoiding, minimizing, or mitigating adverse impacts and promoting informed decision making.
- Avoid, minimize, or mitigate adverse impacts on natural habitats and physical cultural resources.
- Protect public and worker safety against potential risks associated with construction activities, exposure to hazardous materials, and natural hazards.
- Manage land acquisition and loss of access to natural resources in a way that avoids or minimizes displacement, and assist the affected people in improving, or at the minimum restoring, their livelihoods and living standards.
- Give due consideration to the cultural appropriateness of, and equitable access to, program benefits, giving special attention to the rights and interests of the indigenous peoples and to the needs or concerns of vulnerable groups.
- Avoid exacerbating social conflict, especially in fragile states, post-conflict areas, or areas subject to territorial disputes.

In contrast with the practice for IPFs, the Bank itself is expected to undertake the environmental and social systems assessment (ESSA) for a PforR, and an ESSA is required for every operation. It is to identify any actions needed to enhance the systems during program preparation and implementation (the latter to be included in the PAP). The Bank is also expected to consult with program stakeholders and disclose the results and recommendations of its ESSA before the appraisal of the PforR. In addition, the environmental and social risks—and the related mitigation measures—are expected to be inputs to the integrated risk assessment of the PforR. During implementation, the Bank is expected to pay particular attention to reviewing the borrower’s compliance with its contractual undertakings in the environmental and social management areas, including those related to the PAP; monitor the evolution of risks; and make adjustments as appropriate.

All reviewed ESSAs have been of reasonable quality. IEG’s desk review assessed the 27 ESSAs (for operations approved through fiscal 2015) against the above criteria. On this basis, IEG rated all of these as satisfactory or moderately satisfactory.

But there are some shortcomings in the assessments. The most frequent shortcomings related to the extent to which the capacity-building measures recommended in the ESSAs were followed up in the PAPs and the technical assistance provisions; the coverage of social issues; the adequacy of outreach and consultation with poor and vulnerable beneficiary groups; and the absence of monitoring indicators and reporting on safeguards implementation in the results frameworks of the operation. These issues are discussed further below.

Stakeholder engagement associated with the programs' social systems was being fully implemented and supervised by the Bank teams, as IEG found for the projects visited. Thus, the Brazil Ceará Strengthening Service Delivery Program had established a publicly accessible website with all the program's documents, as well as the results framework, targets, and status. In addition, the level and frequency of stakeholder engagement seemed appropriate for the type and implementation status of the program. The skills development program involved consultations with private sector organizations, and the family assistance program involved consultations with local community leaders for the development of an outreach strategy for indigenous peoples and other vulnerable minorities.

Preexisting capacity and budgetary constraints continue to limit the effectiveness of the program systems. IEG's field visits found that the planned technical assistance and PAP actions – mostly involving the preparation of specific guidelines and training for program staff – were being implemented, but had been unable to address the existing constraints. For example, the PAD of the Vietnam Rural Water Supply and Sanitation Project appropriately mentions the Ministry of Natural Resources and Environment – and, by implication, the provincial Departments of Natural Resources and Environment – as responsible for monitoring and evaluating the environmental aspects of program implementation. During IEG's visit, the provincial departments confirmed that they were responsible for monitoring and enforcing environmental requirements, but they had been unable to carry out these functions as regularly and promptly as expected under the law due to continuing staff and budgetary shortages that had already been identified in the PAD.

Reporting on the monitoring of environmental and social effects of the PforRs has been sparse. While most (78 percent) of the PforRs reviewed have been rated as facing moderate or substantial environmental and social risks, and two-thirds (63 percent) of the ESSAs point to the need for monitoring and reporting, only 10 (37 percent) of the PADs discuss specific provisions for the monitoring of environmental and social effects, and five (18 percent) of the ISRs provide any information on their implementation. During its country visits, IEG found that the systems were being implemented, but with virtually no reporting of environmental and social effects,

even at the program-management level. The country teams were adequately up to speed on the status and extent of land acquisition, involuntary resettlement, community consultations, extent of compliance with environmental permitting regulations, and similar issues, but little of this information was (with some exceptions) reflected in the progress reports, and even less in the ISRs.

It would seem that environmental and social effects have not normally been seen as integral to PforR reporting. This is indicated by the evidence from the review of program documents and field visits. This is surprising, because environmental and social risks have been rated as moderate or substantial in all but six (77 percent) of the sample portfolios and, as already noted, these risk ratings were raised during supervision in five (18 percent) cases. Thus, IEG found no evidence that environmental and social risks are not being managed, but the very limited reporting of environmental and social effects available in the Bank's systems is a major concern, given the essential role of monitoring for the environmental and social management framework and the important place of environmental and social risks within the PforRs' integrated risk management framework.

Emerging Messages

Overall, the structure of the Bank's assessments for the PforRs – technical, fiduciary, and environmental and social – has proven to be appropriate, and the assessments have been credible and comprehensive, but there is some room for improvement. In particular, costing methodologies used to prepare the expenditure frameworks for the supported programs are often unclear. In addition, some contextual or political economy issues may not be addressed in the fiduciary assessments, which are generally stronger on the formal systems (such as procurement laws and regulations) than on the practical realities (such as actual fraud and corruption). The most frequent shortcomings for the environmental and social assessments relate to the extent to which the capacity-building measures recommended in the assessments were followed up in the PAPs and the technical assistance provisions; the coverage of social issues; the adequacy of outreach and consultation to poor and vulnerable beneficiary groups; and the absence of monitoring indicators and reporting on safeguards implementation in the results frameworks of the operation.

There are issues surrounding high-value procurements. The high-value procurement exclusion has reportedly prevented some financing in areas supported by the PforRs. Following the *Two-Year Review* (World Bank 2015e), a provision was added that such contracts may be financed if they are important to the integrity of the PforR program and their cost is less than 25 percent of the cost of the overall

program. In the absence of such exceptions, these procurements may be handled outside the programs, and thus outside Bank oversight. IEG finds that this exclusion has not been a major obstacle in the PforRs reviewed. Experience has demonstrated that it can also be problematic to include high-value contracts under the PforR instrument; this is likely to impose special risks and more work on the clients as well as the Bank.

Some borrowers raised concerns that the ACG limited the scope of PforR operations. Clients wondered, for instance, why the Bank agreed to the use of all other aspects of national systems, but insisted on the application of its own ACG, without even assessing the quality of the country's own anti-corruption system. The Integrity Vice Presidency has received only 5 complaints alleging corruption in 3 of the 27 PforR operations assessed by IEG. The complaints were judged to not warrant an investigation and no investigations were carried out. The guidelines have been modified in accordance with the findings of the *Two-Year Review* (World Bank 2015e) to address these concerns and to be clearer about their applications to PforR operations, but it is too early to determine the result.

4. Identification, Assessment, and Mitigation of Risks

The identification, assessment, and mitigation of risks is expected to play a central role in the appraisal and implementation of PforR operations. As stated in *Bank Policy – Program-for-Results Financing*, “the Bank’s assessment of a proposed program (will) evaluate the relevant risks and the scope for improvement and managing such risks, including proposed institution strengthening activities to be undertaken before, if deemed appropriate, and during program implementation” (World Bank 2015b, para. 5). The ability of PforRs to rely on and support the strengthening of the government programs’ own risk management systems during program implementation constitutes one of the defining features of the instrument.

The integrated risk assessment in the PAD identifies and consolidates the risk-related findings for each PforR. As provided in the *Bank Directive – Program-for-Results Financing* (World Bank 2015a), the programs’ integrated risk assessment is expected to consolidate the risk-related findings of the technical, fiduciary, and environmental and social systems assessments and provide a key input into the Bank’s decision to provide PforR financing (World Bank 2015a, para. 31). In line with the *Interim Guidance Notes to Staff on Assessments for PforR Financing*, the integrated risk assessment will arrive at an overall risk rating for the project based on the consolidation of risks associated with the operating environment – country risk and stakeholder risk – and the following program-level risks (World Bank 2012c):

- *Technical risk*: related to the programs’ economic rationale, technical soundness, institutional capacity, implementation arrangements, and M&E arrangements.
- *Fiduciary risk*: related to the programs’ fiduciary systems arrangements and capacity, including fiduciary-related integrity issues.
- *Environmental and social risk*: related to the potential environmental and social impacts of the program and capacity for avoiding, mitigating, or managing those impacts and risks.
- *DLI risk*: related to the programs’ results framework, the selection and level of DLIs, and the risks associated with the verification arrangement for DLIs.

The *Two-Year Review* of the implementation of the PforRs concluded that “in general terms, technical, fiduciary, environmental and social, and DLI risks seem to have been reasonably well identified” (World Bank 2015e, p. 30). It also found that the

risk assessments had influenced the design of PforRs through the selection of DLIs and certain exclusions, and that risk aversion might have resulted in limiting the scope and content of particular programs. The review also found that, in general, the risk-mitigation measures embedded in the design of PforR operations and usually included in PAPs seemed to be relevant and achievable.

Appropriateness of the Identification and Assessment of Program Risks

Fiduciary risks appear to be the main drivers of overall program risk, followed by technical risk (table 4.1). The integrated risk assessments in the PADs of the 27 approved PforR operations indicate that Bank teams have rated the overall risks associated with these programs as substantial in the majority of cases, as moderate in about a quarter of cases, and as high in four cases. To assess the effectiveness of the PforR preparation process in identifying and assessing critical risks, IEG has reviewed the appropriateness of these ratings.

Table 4.1. Risk Ratings in PforR PADs

Risk ratings in PforR PADs	High	Substantial	Moderate	Low	Total
Technical risk	2	12	13	0	27
Fiduciary risk	4	17	6	0	27
Environmental and social risk	0	2	20	5	27
DLI risk	1	8	18	0	27
Overall risk	4	16	7	0	27

Technical risk ratings: IEG found most of the technical risk ratings to be satisfactory, but underrated in some cases. The most significant differences were in relation to the Nepal Bridges and Rwanda Agriculture Programs. In both cases the PAD rated the technical risk as moderate, but it should – in IEG’s view – have been rated high. In the Nepal Bridges Program, this was due to a lack of commitment to the budget reforms supported by the PforR and the weak and untested capacity of the implementing agency. In the Rwanda Agriculture Program, it was because of the weak capacity of the private sector to implement a critical component and the lack of clarity concerning such a response.

The challenges to the implementation of reforms were underestimated in some cases. For the Tunisia Urban PforR, a high rather than substantial rating would have been more appropriate, given the implementation challenges, including ownership risks – the program involves major changes in the roles of central and local governments and an active role of the citizens, who had previously been excluded

from the decision-making and monitoring processes. For the Vietnam Rural Water Supply and Sanitation Operation, the technical risk rating of moderate appears optimistic in light of the substantial risk to the financial sustainability of capital investments, considering the borrower's historic unwillingness to raise tariffs to cost-recovery levels.

Fiduciary risk ratings: The fiduciary risks associated with the PforRs have been satisfactorily identified and rated in most cases, though a few PADs underestimated the risk. Thus, for Nepal Bridges, the fiduciary risk should have been rated high, rather than substantial, in light of the high level of accountability and governance issues, the prevailing weaknesses of a dysfunctional budget process, and vacancies in key anti-corruption positions.

Environmental and social risk ratings: Environmental and social risks were satisfactorily rated in most cases. Given the exclusion of high-risk components, as mandated by the Bank Directive (World Bank 2015a), the actual environmental and social risks faced by the sample PforRs ranged from low to substantial, but seemed too cautious for two PforRs. For the Kenya National Safety Net Program, which has no direct environmental effects and only a very limited range of potentially adverse social effects, the risk would have been appropriately rated as low rather than moderate. Similarly, a risk rating of moderate would seem more appropriate than the rating of substantial for the Vietnam Northern Mountains Urban Development Program, since it has excluded investments in environmentally sensitive areas, avoided investments in ethnic minority areas, and is addressing identified social risks through appropriate PAP actions.

DLI risk ratings: The DLI risks have been appropriately rated in over half of the cases. While DLI risks are generally rated a bit lower than overall program risks, indicating that they are not normally seen as risk drivers, there is a substantial share of underratings. The most frequent source of underratings has been related to the risk that some of the DLIs could not be achieved. Thus, for the Tanzania Health Program, the risk was rated as moderate in the PAD, although concerns were raised about timely availability of resources, including technical assistance to support program implementation. The formulation of the DLIs was also quite complex. Risk would have been more appropriately rated as substantial. For the Moldova Health Program, a major risk relates to the timing of two DLIs (smoking prevalence among adults and adults with hypertension whose blood pressure is under control), given the possible lag between reforms and the achievement of such outcomes. On this basis, IEG would rate the DLI risk as high, compared with the moderate rating in the PAD. Finally, while it is important to have accurate ratings for DLI risk, it is also

important to note that DLI ratings are an indicator of the level of expectation set by the PforR, so they do not necessarily need to be low.

Overall risk ratings: Overall risk ratings have been appropriate in most cases. In line with the policy mandate and guidelines, program risk ratings were integrated into an overall risk rating for each PforR, which also reflected the risks to the operating environment – country risk and stakeholder risk. Based on IEG’s desk review of the 27-project portfolio and its field missions to 8 countries, most of these risks had been satisfactorily identified, and the overall risk ratings were appropriate. In a few cases, the overall risk rating was lower than called for in light of the underlying information, as illustrated by the examples below.

- For the Brazil Ceará Service Delivery Program, the overall PAD risk rating of moderate looks optimistic in light of the substantial risks related to the ability of the skill development program to place its trained technicians in productive sector jobs during the ongoing recession; the ability of the family assistance program to reach its intended beneficiaries among the most vulnerable household groups, given the exclusion of municipal-level components from the scope of the PforR; and the ability of the water quality program to improve water quality in the target reservoirs, given the increasing length and frequency of droughts brought about by climate change.
- For the Croatia Health Program, the integrated risk rating of substantial also seems optimistic, considering the complexity of the program, which involves wide-ranging reforms. Program implementation would be challenging even with technical assistance, which the PforR does not include. It would have been more appropriate to rate the program risk as high, and then lower it if warranted by the quality of implementation progress.

Appropriateness of Risk-Based Exclusions

Risk-based exclusions have been built into the PforR design. *Bank Policy – Program-for-Results Financing* (World Bank 2015b) provides for two types of exclusions intended to reduce the risks associated with the instrument: activities that involve procurement of goods, works, and services under high-value contracts above certain thresholds based on the type of procurement and the level of fiduciary risk, and those with potentially significant and irreversible adverse impacts on the environment or affected people.

HIGH-VALUE PROCUREMENTS

The exclusion of high-value procurement has not been a major obstacle in the PforRs reviewed, though it has led to the exclusion of certain areas supported by PforRs in Bangladesh, Ethiopia, Uruguay, and Vietnam. The ensuing financing gaps have stimulated some creative solutions. Thus, in Ethiopia Health, to ensure that there was no financing of contracts exceeding the PforR thresholds, the IDA credit (and the associated Health Results Innovation Trust Fund grant) went into a subaccount within the MDG Performance Fund. This still allowed IDA funds to be harmonized with the rest of the Performance Fund activities, excluding high-value contracts. The processes for procurement, reporting, audits, and anti-corruption are the same for all donors. (The Bangladesh case is described in chapter 3.)

HIGH-RISK ENVIRONMENTAL AND SOCIAL ACTIVITIES

All of the PforRs reviewed have appropriately excluded high-risk environmental and social risk activities. In one case, however, the interpretation of this policy requirement appears to have been overly cautious, to the extent of having an impact on the achievement of the PDO. For the India Maharashtra Rural Water Supply and Sanitation Program (World Bank 2014a), the ex-ante exclusion of “schemes involving highly polluted water sources” (p. 6) without comparison of treatment options with alternative raw water sources may needlessly foreclose the most feasible option in specific locations. It thus appears to be at variance with the operation’s PDO “to improve access to quality and sustainable services in peri-urban villages and in water-stressed and water-quality affected areas” (p. vii).

In a few other cases, an excessively risk-averse application of this exclusion has significantly reduced the scope of the PforRs in relation to the supported government programs and the attendant challenges and opportunities associated with strengthening the programs’ systems. For example, the Uganda Municipal Development Program excludes water treatment plants and sanitary landfills, and the Ethiopia Local Government Program excludes roadways outside existing rights-of-way. All projects exclude any investments with potential impacts on natural habitats or cultural areas. This raises the concern of how such small- and medium-scale investments, which are integral to the investment programs of a variety of sectors, will be handled if they are not subject to the oversight associated with Bank involvement. In addition, as already noted in the *Two-Year Review*, the exclusion of potentially significant environmental and social impacts causes differences between PforRs and the government programs they support (World Bank 2015e, p 15).

The observed overly cautious interpretation of the high environmental and social risk exclusion detracts from the integrity of the PforRs and may represent a missed

opportunity to strengthen country systems. While the exclusion of high risks is in line with the Bank *Policy on Program-for-Results Financing* (World Bank 2015b), IEG's review has found that the substantial and moderate risks encountered throughout the PforR portfolio were being diligently supervised by the Bank teams and were reasonably well managed by the country systems. On this basis, it would seem appropriate for PforRs to include the full range of what would be classified as Category B projects under the IPF safeguards, rather than excluding subprojects perceived to be at the riskier end of the Category B spectrum. This would mean the inclusion of all small- and medium-scale investments that are integral to the supported programs and are likely to cause mostly local and short-term negative environmental and social impacts for which effective mitigation measures are readily available. An implication would be that the PforRs would need to strengthen, where necessary, the capacity of the country systems to enable them to handle all the investments in the government program in an environmentally and socially sound manner, rather than creating a two-track system, which can undermine the integrity of the program and of the country systems.

Risk Management Implementation Experience

Based on the ISRs, the integrated risk management framework used for the PforR portfolio appears to be working reasonably well. A comparison of the five risk ratings in the PAD – technical, fiduciary, environment and social, DLI, and overall – with those in the latest ISRs for the 12 projects visited by IEG (in 8 countries) found that the ratings had been changed in only 3 projects, involving 6 of a total of 60 ratings. These ratings had been lowered in three cases and raised in another three. This impression of stability is consistent with the ISR's reporting that the implementation of the numerous program risk management measures associated with these PforRs – between 5 and 13 for every project – was “in progress” in every case, except for 2 that had been delayed and 4 that had been completed.

MANAGEMENT OF FIDUCIARY RISKS

Overall, IEG country visits found the implementation of fiduciary risk management to have been “in progress,” but often subject to delays. While Bank supervision missions have been actively dealing with these delays and other shortcomings, and their actions are promising, it is still too early to establish their final results. Thus, for the Ethiopia Health Project, the ISRs have reported an improvement in the transparency of the Pharmaceutical Fund and Supply Agency (PFSA), as demonstrated by the launch of the PFSA website, as well its disclosure of agreed procurement information and the completion of pending audits that had been delayed since December 31, 2013. Given the multiple requests for extensions, the

task team has diligently followed up with the PFSA management to identify the root causes of the delays, and jointly identified areas of weakness in the capacity of the PFSA to undertake sound financial management and procurement processes. For the Rwanda Agriculture Project, the first ISR raised a number of issues about the adequacy of the agricultural expenditure framework (level, allocation priorities, management arrangements, and effectiveness at both the national and subnational levels), and the appropriateness and effectiveness of the funding modality to help ensure strong alignment and harmonization of the funding being provided by a multiplicity of development partners and to help leverage expanded private sector financing. The Bank had requested that the borrower prepare an operational action plan to address these issues, but its implementation is still under way.

MANAGEMENT OF ENVIRONMENTAL AND SOCIAL RISKS

The ISRs generally report that the environmental and social risk management measures are “in progress,” but provide little information on actual environmental and social effects. This is surprising, since environmental and social risks have been rated as moderate or substantial in most PforRs, and the monitoring of impacts is an essential element of the environmental and social risk management framework. This finding parallels that of IEG’s recent review, *Managing Environmental and Social Risks in Development Policy Financing* (IEG 2015c), which also found that there is no formal system in place in the Bank to monitor and report on the actual environmental and social effects of DPFs.

However, during its country visits, IEG found that the environmental and social risks were being diligently supervised by the Bank teams and generally implemented as expected. No serious issues were anticipated, especially since the programs had excluded investments in environmentally sensitive areas and generally minimized land acquisition. In the few instances (Vietnam and Brazil) where the programs involved indigenous peoples, adequate provisions for their culturally sensitive participation had been addressed through the PAP and supported by technical assistance.

Even so, IEG’s field visits provided an understanding of important challenges faced in the adequate implementation of the PforR policy requirements. One recurrent challenge relates to the adequacy of the program systems, which most ESSAs have described as adequate from a legal and policy perspective, but as stretched and uneven in terms of implementation capacity. IEG found that this situation continued for the programs visited. While technical assistance has been provided in every case to prepare environmental and social guidelines and train program staff to implement them, the missions found no indications that the resource and staffing shortages diagnosed in the ESSA had been addressed. Thus, the enforcement of

environmental requirements and monitoring of environmental performance continued to be sporadic and uneven.

Another recurrent challenge relates to the adequacy of compensation for land acquisition. Overall, IEG missions found that the programs that required land acquisition had supported the preparation of the land acquisition and resettlement guidelines and trained staff to implement them in conformity with Bank requirements. Verifying the adequacy of land compensation, however, remains a difficult challenge, as illustrated by the Vietnam Northern Mountains Urban Development Program (box 4.1).

Box 4.1. Compensation for Land Acquisition—Vietnam Northern Mountains Urban Development Program

In this program, the PAP supports “independent market-based assessments” of land prices. IEG was informed that such assessments had arrived at prices that were 2–5 percent higher than established government guidelines, but that it was difficult to establish market prices because of the common practice of underreporting the actual transaction prices in the official records to minimize land transfer taxes. A possible indication of the distortion created by this underreporting and underassessment is that out of the 40+ households that needed to be relocated for the first investment project in a major city, all had accepted the land compensation option in an outlying resettlement area, and not a single household chose the cash compensation option that was based on the “independent market-based assessment.”

Source: World Bank 2014e.

Emerging Findings

The PforR risk management framework has worked largely as intended. Risks related to PforR operations have generally been well identified and assessed. Overall, IEG’s findings concur with those of the *Two-Year Review* (World Bank 2015e), which concluded that the risks associated with PforRs have been reasonably well identified and assessed, with only a few exceptions. IEG found that most of these exceptions involved the underrating of risks, particularly those related to technical, fiduciary, and DLI risks. Environmental and social risks, in contrast, had a slight tendency to be overrated.

The exclusion of high-risk environmental and social activities has unnecessarily reduced the scope of several PforRs in relation to government programs. IEG has concluded that the high-value procurement exclusion has not been a major obstacle for the achievement of the PDOs. The exclusion of high-risk environmental and

social activities, however, has reduced the scope of the PforRs in relation to the supported government programs in several cases. This is consistent with the finding of the *Two-Year Review* that this exclusion was one of the most important factors accounting for differences between the scope of the PforRs and the government programs. In one case (the India Maharashtra Rural Water Supply and Sanitation Program), the exclusion of “schemes involving highly polluted water sources” appears to be at variance with the operation’s PDO to improve access to quality and sustainable services (World Bank 2014a).

The management of risks is progressing well, though the reporting system is inadequate. The PforRs appear to be making good progress with the management of all kinds of program risks, although the ISRs do not adequately reflect the frequency of delays, and they provide no information on the environmental and social effects of the operations. This is surprising, because environmental and social risks have been rated as moderate or substantial in most PforRs, and the monitoring of impacts is an essential element of the environmental and social risk management framework.

5. Strengthening National Systems

The potential for strengthening the institutional capacity of national systems is a key feature of PforRs. As stated in the Board paper (World Bank 2011a), the PforR focuses on the behavioral and institutional changes that are required to achieve results and manage associated risks. Hence, strengthening the capacity of national institutions to implement the program is expected to be a priority for both preparation and implementation support. This chapter discusses the use of the PforR instrument to strengthen the countries' systems for the financial management, procurement, environmental and social safeguards, and M&E of the supported programs.

The Identification and Assessment of Capacity-Building Measures

Assessments of technical, fiduciary, and environmental and social systems have identified numerous measures to strengthen the performance of program systems to ensure that they achieve the expected results. IEG has reviewed these measures for the 27 PforR operations in the portfolio and assessed their appropriateness in light of the gaps identified.

MEASURES FOR STRENGTHENING TECHNICAL CAPACITY

Measures for addressing gaps in technical capacity are largely adequate. This was the case for all but 1 of the 27 PforR operations assessed. Operations supported government programs that are well defined and clearly bounded, with adequate borrower capacity, borrower participation in design and implementation, defined program management mechanisms, and coordination processes with other development partners.

Countries are eager to rely on their own financial management systems, including designation of aid funds as on-budget, aligning with the fiscal calendar, and relying on existing procedures for budget preparation and execution, intergovernmental transfers, and audit. Among the advantages to governments of the use of the PforR instrument, the ability to use and strengthen their own country systems instead of employing Bank processes is an important one. For example, under the Ethiopia Local Government Program, urban local governments have achieved targeted budgetary improvements in capital investment plans and value-for-money audits as measured in annual performance assessments linked to DLIs. The Kenya Safety Net PforR has implemented program information systems using agreed standards for internal

payroll controls, making payments electronically using two-factor authentication, and disbursing to payment-service providers in a timely manner.

In many cases, the PforR approach offered greater accountability, scope, and flexibility when compared with the earlier Bank operations. This is illustrated by the example of the Uganda Urban PforR (box 5.1).

Box 5.1. Support to Municipal Urban Development in Uganda

The Uganda PforR built on a 10-year period during which three operations implemented performance-based disbursements: Local Government 1 and 2 and Municipal Government 1. In contrast to these IPFs, the PforR linked specific disbursements to DLIs, instead of channeling local development grants based on local government performance assessments. Another difference is that the scale of the infrastructure investments supported by the PforR was much larger than in the previous operations. A final difference is the use of an independent assessment firm to verify performance, rather than using the government's own monitoring processes, which in the previous operations proved unreliable and subject to excessive delay.

Capacity building under the PforR focuses on providing training, essential tools, and systems, such as renovated offices and internet access, and the resulting capacity improvement is monitored through DLIs. Previous capacity building had focused on skills training and certification, but was not monitored through the results frameworks, and it was found that the best-trained people would leave. Retention is reportedly much better under the new approach, and is enforced by an annual assessment of municipalities to ensure that strategic positions are filled. A program support team provides seven professional staff to the Ministry of Lands, Housing, and Urban Development to help manage the program. This was requested by the ministry because of its staffing constraints. Staffing levels in the ministry show that only 50 percent of positions are filled, and the project budget is bigger than the budget of the ministry itself. This challenge of unfilled positions is the same in other institutions.

Source: World Bank 2013c.

In some cases, the capacity and other system gaps were avoided or mitigated through the purposive bounding of the PforR versus the government program. The portion of the program supported by the PforR can be defined in different ways. Thus, the Croatia Health PforR covers a slice of the government program based on duration, priorities, and institutions. The Moldova Health PforR covers 5 of the 17 subprograms in the government program that are expected to have the greatest impact on the achievement of the 2 PforR objectives. In two of the five subprograms, the PforR further limited or focused its coverage to some specific performance-based incentives.

In a few cases, the institutional capacity measures could have been more clearly defined. For example, the PAD for the Ethiopia Health MDG's operation (World Bank 2013b) points to the weak audit capacity of the Pharmaceutical Fund and Supply Agency and aims to have adequate professionally qualified and trained staff for internal audit. It also targeted completion of pending audits at that agency by December 2013. As of February 2016, these issues still had not been addressed, although progress was being made. The India Teacher Effectiveness PAD (World Bank 2015d) does not discuss the capacity of the various entities, except to note that the academic authority for all teacher education programs and the nodal agencies for training delivery are in a "deplorable state of disrepair" (p. 27). The specific gaps and appropriate mitigation measures could have been elaborated.

Box 5.2. Mozambique Public Financial Management Capacity Support

The Mozambique PforR has a number of innovative features in capacity support. The program has an allocation of \$8 million for capacity development. This support is implemented by the government, not a separate component executed by the Bank or other donors, as in some other PforRs. The program includes a substantial emphasis on change management, which is appropriate. An innovative approach, for Mozambique, uses a system of 5 coaches (recruited internationally) and 22 national facilitators. This is only mentioned briefly in the PAD, though it is discussed in detail in the aide-memoires of February and June 2015. A related innovation is the use of the rapid results initiative, not mentioned at all in the PAD, but briefly discussed in the aide-memoires, drawing on examples from Kenya and Madagascar. Both innovations are supported by a Bank team member with special background and interest in such efforts. A World Bank-managed trust fund has also provided support for the coaches and facilitators to participate in a training workshop in Accra on rapid results initiatives. It will be interesting to see how this will work out—clearly there are teething problems and a widespread expectation by many stakeholders, particularly in the provinces, that coaching should be the same as technical assistance by experts. A challenge for coaching is also that government health technical staff are stretched thin and face rapid turnover, because government staff are recruited by private health clinics and nongovernmental organizations. The recent introduction of the rapid results initiative concept seems to represent a hope that results can be fast-tracked after a long start-up delay.

It is not clear on what basis it was determined that the coaching and technical assistance provided under the program and by other donors will be sufficient. The concept for coaches and facilitators originated from the design facilitation supported by a former Bank staff member and was taken onboard by the government.

MEASURES FOR STRENGTHENING FIDUCIARY SYSTEMS

In almost all cases, the measures for strengthening fiduciary systems were adequate. The exception is Nepal, which Transparency International has ranked as the second-most corrupt country in South Asia, and 139th out of 183 countries surveyed

worldwide. This would seem to call for intensive anti-corruption measures, such as arrangements for random post-audits, on-site field verification of the quality and quantity of outputs and appropriateness of unit costs, and provision for rigorous follow-up of audit reports and management letters, including audit committees with external stakeholders, such as civil society, professional organizations, and government representatives from outside the project entity. There is not enough attention to such measures in the PAD or the Institutional Fiduciary Assessment.

In many cases, addressing fiduciary capacity gaps required the design of innovative features, as illustrated by the Mozambique PFM operation in box 5.2.

In several cases, the measures to strengthen fiduciary systems drew from earlier projects and other donor support. This was the case in the Rwanda PFM PforR, where a key constraint is the high turnover of procurement officers. The Bank's earlier Public Sector Capacity Building Project had worked to address this issue, which is pervasive across professional positions in the public sector. But despite extensive support provided by that operation, and related capacity-building operations funded by other donors, the problem has persisted. The PAD correctly points out that better incentives for attracting, promoting, and retaining staff are needed to reduce high (by regional standards) vacancy and turnover rates. It was apparently awkward to address this issue through the PforR, but the PFM fund donors are addressing it through other support.

Institutional strengthening may be supported by DLIs, as in Tanzania Education, where the six DLIs are institutional (five in number, 80 percent by amounts) – completion of foundational activities, resource flows, annual reporting, deployment of teachers across districts, and schools receiving incentive grants. The PDO in the Rwanda PFM is itself institutional – to enhance the country's public financial management and statistics systems to improve transparency and accountability in the use of public funds, revenue mobilization, and the quality and accessibility of development data for decision making.

In a few cases, fiduciary system gaps are identified in the assessments, but it is not clear how they are all being addressed. For example, the India Maharashtra Rural Water Supply and Sanitation Program Fiduciary Systems Assessment (World Bank 2014b) points out a challenge: lack of comprehensive instructions for all steps of the procurement process, with rules and guidelines spread across many documents, and inadequate rules for service procurement. To address these, it supports the action: "Compile the rules and procedures on procurement, including for selection of consultant services, into a comprehensive procurement manual for the RWSS [rural

water supply and sanitation] sector and adopt it” (World Bank 2014b, p. 15), but other challenges are pointed out without corresponding proposed actions, such as possible entry barriers for contractors created by mandatory registration requirements. Later information could indicate that such matters had been addressed in a separate document.

MEASURES FOR STRENGTHENING ENVIRONMENTAL AND SOCIAL SYSTEMS

The provisions for strengthening the programs’ environmental and social systems were largely reasonable. The most frequent gaps related to limited or vaguely specified inclusions in the PAPs of capacity-strengthening measures recommended in the assessments and, in some cases, inadequate support for the PAP in the technical assistance provisions of the programs. For example, the ESSA of the Uruguay Road Program identified a need for strengthened environmental management and supervision and recommended seven actions to address this gap, but only two of these were included in the PAP. Similarly, the PAP for the Mexico Oaxaca Rural Water and Sanitation Program did not mention any provision for several of the actions that the ESSA identified as required to address the substantial social risks of the program. It is possible in these and other cases that such actions have been covered through other means, but the lack of any explicit mention offers little assurance that they will be adequately addressed and implemented.

MEASURES FOR STRENGTHENING MONITORING AND EVALUATION

PforRs have a greater focus than other Bank operations on reviewing progress on achievement of results through M&E. This follows from the requirement to base PforR disbursements on the achievement of monitorable indicators, rather than inputs, confirmed through specific and transparent verification protocols. Strong M&E is also important in enabling the PforRs to adhere over time to the theory of change underpinning this instrument, as set out in chapter 2.

Over 40 percent of the PforR portfolio includes program support for what the PADs describe as impact evaluations. (IEG did not set out to review the quality of the proposed evaluation designs.) For example, the India Teacher Effectiveness PforR includes capacity-building and systems-strengthening activities. It also provides support for monitoring, reporting, and impact evaluation studies on teacher performance, which will track gains emerging from program interventions. Nigeria Health includes two relevant DLIs: lagging states will strengthen their maternal, neonatal, and child health as part of an impact evaluation. This evaluation will test the effectiveness of results-based disbursements to states and provide some evidence on whether improvements in performance can be attributed to the PforR. Another DLI is testing new approaches to improving the delivery of services by

nonstate actors. All these innovations would be subjected to rigorous evaluations (including impact evaluations where practical). For Tanzania Education, evaluations will be carried out for selected activities of the program. For selected innovative interventions, the government, with the support of the World Bank and the government of Norway, plans to carry out a rigorous impact evaluation of the results-based financing, using an experimental design.

M&E processes vary across the portfolio. Many PforRs use independent monitoring agents contracted to verify performance of DLIs prior to disbursement. For example, the Bangladesh VAT Improvement PforR allocates 4 percent of program expenditure for verification activities, monitoring and reporting work, and capacity-building efforts associated with strengthening government systems, including technical capacity support for the Anti-Corruption Commission, to be coordinated by the Economic Resources Division. The involvement of the Economic Resources Division, which is a separate organization from the client, the National Revenue Board, is thought to make the process more independent and objective than if the client were responsible. In addition, the project office has developed a tracking system for progress on each DLI and component.

Links with IPFs

Some PforRs have evolved from operations using other instruments. IEG found that four programs (Brazil Ceará Strengthening Service Delivery for Growth, Poverty Reduction, and Environmental Sustainability; Ethiopia Second Urban Local Government Development Program; Rwanda Transformation of Agriculture Sector Program Phase 3; and Vietnam Results-Based Rural Water and Sanitation under the National Target Program) had evolved from previous IPF operations that contributed to program design and institutional capacity. Hence, in Brazil, the PforR had many similarities with the predecessor sectorwide approaches (SWAPs), including the use of DLIs and country systems, with exceptions for procurement above threshold amounts. However, the PforR has stronger verification processes and greater focus on improving country systems through the use of the PAPs. The client also found greater opportunities for cross-sectoral coordination in the PforR, in part because the SWAPs had expenditure rules that led the sectors to focus on their own programs. The PforR was generally perceived to be a more flexible instrument.

Traditional IPF operations (Ethiopia) and Adaptable Program Loans (APLs) (Rwanda and Vietnam) contributed to the development of the programs supported by the PforRs. A common characteristic of these predecessor operations was the

introduction of new institutional arrangements, such as the establishment of RWSS enterprises in Vietnam and the use of performance-based intergovernmental transfer mechanisms in Ethiopia. Capacity building was also a major focus of the IPF operations, especially where the institutions were weak, as in Rwanda, where the first of three Rural Sector Support Project APLs focused on establishing basic institutional and technical capacities. The IPF instrument provided hands-on support in such cases, while APLs provided flexibility when introducing or testing new institutional arrangements and approaches. As the programs became better defined, institutional capacity was strengthened, and program/country systems were more firmly in place, which made the risks of the use of the PforR instrument manageable. It is unclear whether the PforR would have been an appropriate instrument for the programs supported by predecessor operations, with the possible exception of the later programs in the Rwanda APL series, which had a good record of meeting targeted results.

Links with Development Partners and Technical Assistance

Technical assistance provided by development partners made an essential contribution in several cases. For Morocco Human Development, the main external partner is the European Union (EU). Both the Bank and the EU supported the first phase of this program (which the Bank supported through an IPL-financed SWAP). The EU now finances targeted technical assistance for the program as well as direct support to the general budget against certain indicators (some are similar to or the same as those used by the Bank, others are different) with a focus on rural areas. There are no formal cofinancing agreements between the Bank and the EU, but there is nevertheless effective coordination between the two entities, with joint missions and partially identical monitoring and results indicators.

For the two PforR operations in Rwanda, technical assistance is provided by other donors through a variety of support interventions in parallel with the PforRs. This support would probably have gone ahead largely or fully, even without the PforRs, but several donors said that the Bank's operations help to provide an overall envelope that may also strengthen the impact of their support. Some complementarity was also found for the Bangladesh VAT PforR. In that case, a number of partners had important roles at an early stage, including in the preparation of the overall Tax Administration Modernization Plan. There was, however, a lack of consultation between the Bank (regarding the VAT) and the ADB (regarding income tax administration) that may have made it more difficult to achieve use of a common taxpayer numbering system that is linked to a national identification number across income tax, VAT, and customs.

Implementation Experience with Capacity-Building Measures

Overall, IEG country visits found the implementation of capacity-building measures to be progressing reasonably well, although frequently subject to delays. For example, under the Ethiopia Local Government Program, urban local governments have achieved targeted fiduciary improvements in capital investment plans and value-for-money audits as measured in annual performance assessments linked to DLIs. The Kenya Safety Net PforR has implemented program information systems using agreed standards for internal payroll controls, making payments electronically using two-factor authentication, and disbursing to payment-service providers in a timely manner. Eight of the 12 actions had satisfactory progress, but an important action for targeting and enrollment of beneficiaries was substantially delayed.

IEG found very few cases where PforR capacity-building impacts extended beyond the scope of the programs financed. This is not surprising, given the program focus of these operations and the pervasive budget and staff constraints on both the Bank and the borrower. An exception was the Morocco Human Development Program, which included significant investments in strengthening the fiduciary systems at the local level to address shortcomings identified during an earlier Bank project. Following up on the initial results of these measures, the government is embarking on a reform to introduce results-based budgeting in five additional ministries (box 5.3).

Box 5.3. Morocco National Initiative for Human Development (INDH2)

INDH2, with the support of the PforR, has made significant investments in strengthening the fiduciary system at the local level to address shortcomings identified during the first phase of INDH. Audit reports from the first phase of INDH (supported through an IPF) found considerable shortcomings in financial management and procurement in subprojects implemented by municipalities and by associations or nongovernmental organizations. This was ascribed to capacity constraints. As part of the INDH2 PforR, detailed fiduciary guidelines were developed and training in their use was conducted in all participating provinces and municipalities.

By all accounts, including those of the auditors, there were more shortcomings during INDH1, when municipalities were asked to follow a modified local system to adhere to Bank rules, which created substantial confusion and required double bookkeeping.

The same fiduciary system is applied to all INDH subprograms and subprojects, regardless of whether they receive PforR financing (PforR supports three out of five programs and its financial contribution of \$300 million is small compared with overall program costs of \$2.1 billion). Thus, any improvements introduced to this point have a significantly greater reach than they would have had were Bank procedures only applied to its financing of \$300 million and limited project counterpart funds under a traditional IPF project. Under INDH,

efforts also have been made to strengthen and expand the grievance system. It was not possible for the mission to ascertain how well this system is functioning.

Morocco is embarking on a reform to introduce results-based budgeting over the next several years, with a pilot starting in five ministries in 2016. While the PforR operation is a special initiative and has somewhat simplified procedures, officials in the Ministry of Finance and Ministry of General Affairs (the Bank's main counterpart) indicated that the work with the Bank under the INDH PforR was a useful prelude to the introduction of results-based budgeting in line agencies.

The strengthening of monitoring and verification has yielded improved verification. In the Uganda Urban PforR, DLI verification for previous non-PforR operations in the sector was done by the Inspection Office in Ministry. However, this verification was reportedly not rigorous. Under the PforR, DLI verification is done by an independent firm, and the procurement of the firm is carefully monitored by the Bank; in one case the bidding was redone because of questions raised. The verification report builds on findings of the auditor-general, and there are spot checks by the Bank team before the report is finalized. The result is rigorous verification of DLIs.

But there can also be side effects to strengthening monitoring and verification. The Mozambique PFM PforR uses the government's supreme audit authority, the Tribunal Administrativo, for the verification of most DLIs. Although the work of this unit is highly regarded, it is severely understaffed in relation to its many tasks. To address this, the unit has been awarded two capacity-building grants from the PforR. However, the use of a public institution in this case adds to the work burdens of an already stretched entity.

Emerging Findings

The identification and assessment of measures to strengthen the capacity and performance of program systems has been mostly satisfactory. In many cases, the PforR approach offered greater scope and flexibility compared with earlier Bank operations, stimulating the design of innovative features. In other cases, the measures to strengthen program systems drew from earlier projects and the support of donor partners. In some cases, the capacity and other system gaps were avoided or mitigated through the purposive focusing of PforRs within the government programs. In a few cases, the system capacity gaps were identified in the assessments, but the measures to address them were not clearly defined or were left undefined.

Based on IEG's country visits, it can be concluded that the implementation of capacity-building measures is progressing reasonably well, although it is subject to frequent delays. In most cases, the capacity-building measures were focused on the program systems, as appropriate. In a few cases, the strengthening of program systems is being expanded to the whole country. Substantial and sustained capacity improvements will take time, however, often exceeding the duration of a single PforR program.

6. Findings and Recommendations

Main Findings

Since its launch in 2012, the Bank has rolled out the PforR instrument gradually, seeking a balance between responding to potentially strong client demand (and interest among Bank staff) and learning from implementation experience. As of March 31, 2016, the Board had approved 39 PforR operations, providing \$9.5 billion of Bank financing to support a total of \$49.9 billion in government programs, with a pipeline of 21 operations under preparation. There are ongoing PforR operations in all Bank Regions, and they cover most of the sectors in which the Bank has traditionally been active.

The PforR instrument is an increasingly important lending vehicle for the Bank. Overall, the structure of the Bank's assessments of PforRs – technical, fiduciary, and environmental and social – has proven to be appropriate, and the assessments have generally been credible and comprehensive. The results frameworks, DLIs, and PAPs are often reasonably coherent, and risks related to PforR operations have generally been well identified and assessed. Nevertheless, there are areas in need of improvement when it comes to designing the programs to achieve results and to the monitoring and reporting systems.

The PforR instrument was envisioned to help focus more on results than other instruments, as well as to help strengthen country systems and induce further alignment/harmonization among donors. Because none of the PforR programs has yet closed, it is too early to draw any definite conclusions about whether the instrument is doing a better job of achieving these objectives than alternative approaches. Nevertheless, some insights can be derived from the early design and implementation experience.

While the programs do focus on results more explicitly than other instruments through the introduction of DLIs, these indicators are often – but not always – well integrated with the results frameworks. And while the results frameworks are often reasonably coherent, the PDOs are rarely at the outcome level, and explanations of how the PforR objectives relate to the longer-term objectives of the supported government programs are largely absent from the PADs. To ensure a greater likelihood of ultimately achieving the desired development results, more consistent linking of the DLIs to the results frameworks and the longer-term objectives of the program will be required. Box 6.1 sets out some key aspects that should be considered in PforR ex-post completion reports.

Box 6.1. Key Aspects for PforR Ex-Post Completion Reports

No PforR program has reached its closing date, but at such time it would be important for staff to have access to a format for the Implementation Completion Reports. Presumably, this format would follow the existing formats for IPF completion reports, but with emphasis on the special features of PforRs, as discussed in this evaluation report. Areas of special emphasis should include the following:

- The assessment of results against the stated objectives and all the specific indicators in the results frameworks
- Longer-term prospects for the supported programs, including the prospects for eventual outcomes and the sustainability of the results achieved during the PforR period
- Assessment of the DLIs and the progress of disbursements
- Applicability and progress of the PAP as a whole and of all of the individual PAP items
- The role of the PforR in donor cooperation and resource mobilization
- Aspects of the PforR operation that may have encouraged or discouraged the achievement of sustainable results
- Quality and usefulness of the Bank's work, including for the up-front assessments and during implementation
- Lessons going forward.

The DLIs are designed with the two main objectives: to be triggers for disbursements, with an inherent need for predictability, and to provide incentives for performance—stretch targets. In many cases the DLIs are linked to very small shares of total program disbursements, or to routine and repetitive actions, rather than key activities to achieve PDOs. This points to an inherent tendency to shift the balance toward the disbursement objective.

IEG finds that both ownership and partnership are addressed well in the Bank's program documents, and the field visits found a considerable degree of government ownership of the programs under implementation. It also found that the focus of PforRs on programs rather than investments has been helpful in promoting coordination among government agencies that may have had little incentive to work together. However, there is no evidence yet that the instrument has encouraged much additional financing by other donors, let alone any broader use of the strengthened country systems.

Capacity building is an important part of the programs, but specific goals could have been defined more clearly in some cases, and there have been frequent delays in the implementation of capacity-building programs.

The Bank's average costs for the preparation of the new programs have been similar to those of other IPF operations, with significant variations between programs, while average Bank implementation costs have been significantly higher than for IPF operations. There may be increased positive externalities/public good aspects from strengthened country systems. Overall, however, there is not yet sufficient evidence to derive any conclusions about the overall efficiency of PforRs.

Most of the participating borrowers appreciate the PforRs. For the early operations, Bank teams often were the driving force for the new instrument, but governments came to appreciate the PforR approach quite quickly. They particularly value the instrument's reliance on country systems. Among the advantages to governments, the ability to avoid the use of normal Bank procurement processes stands out. Counterparts also emphasized the importance of the focus on programmatic results that are aligned with the government's own strategy and longer-term programs. They spoke positively about the potential impact over time on programs supported by PforRs, including the impact of the almost constant interactions with the Bank teams. For some decentralized programs, as in Vietnam, participants also noted the motivational impact of the focus on results, including the DLIs. Finally, some governments (in particular, countries such as Ethiopia and Rwanda) are also positively disposed toward the PforR instrument, because they welcome what they see as the predictability and general nature of disbursements under PforRs, without policy conditionalities as prior actions.

One exception to the use of country systems is the requirement to use the Bank's ACG, which was viewed by borrowers in Brazil, India, Morocco, and Uruguay as limiting the scope of PforR operations. The guidelines have been modified in accordance with the findings of the *Two-Year Review* (World Bank 2015e) to be clearer about their application to PforR operations and to address these concerns, but it is too early to determine the result. These modifications do not change the Bank's substantive rights under the current ACG. Although referred to as "guidance," the ACG must be followed in accordance with the provisions in the legal documents.

The high-value procurement exclusion and adherence to the ACG are supported by IEG findings given the weaknesses in country systems documented in PforR fiduciary assessments. Corruption can be an issue in the PforR countries assessed, and the special provisions of the Bank's ACG are needed to provide assurance to key shareholders, while also acknowledging that combating corruption is primarily the sovereign responsibility of the borrower.

Approved PforR operations are largely on track in relation to the PAD expectations. As of March 31, 2016, 32 of the 39 approved operations had become effective. Bank disbursements were on track relative to expectations, with annual disbursements increasing from \$108 million in fiscal 2013 to \$479 million in fiscal 2015, and \$1,056 million in the first three quarters of fiscal 2016. Overall, implementation is also broadly on track, with ISR performance ratings of satisfactory or moderately satisfactory and stable risk ratings.

The Bank's PforR directive, procedures, and processes have been working reasonably well. This was the conclusion of IEG's review of the 27 PforRs approved through the end of fiscal 2015. IEG has also found that, for the most part, program systems have provided reasonable assurance that the funds will be used for their intended purpose. While the initial PforR operation in a country has generally taken time, and has been quite expensive to prepare, follow-up operations in the same country tend to have lower costs as the Bank and counterparts move up the learning curve.

While fiduciary assessments are largely well prepared, a shortcoming is the lack of details on the costing methodologies used by the government for the programs being supported. Most PforR operations seem to base their costing on that of the underlying government program, but with little or no detail on the methodology used.

At the same time, while IEG found during its country visits that the programs' fiduciary policies and environmental and social systems were being implemented largely as expected, there were few indications that they had yet been effective in addressing systemic weaknesses affecting national systems, such as budget and staff constraints. Such changes can take longer than these programs have been in place, which could explain why there are only a few examples of institutional improvements under PforRs that are having broader impacts within a government. It is also clear that this instrument's focus on programs may lessen the potential for addressing cross-cutting policy issues.

IEG found that the PAPs have tended to increase in scope. The PforR guidelines state that a PAP should be selective. It should focus on four or five areas, with a few actions in each, as needed, to improve the technical, organizational, and management dimensions of the program; enhance the capacity and performance of the implementing agencies; and mitigate risk. As discussed in chapter 2, IEG found that most of the PAPs were adequately designed, but some have too many actions—49 in one case. It is likely that such a proliferation of actions will dilute the possible impact of the individual PAP items. There are also PADs that do not provide a clear basis for

inclusion of actions in the PAPs, and no articulation of the inevitable trade-offs necessary to identify the most critical or key actions from the long lists emerging from the technical, fiduciary, and environmental and social assessments.

The implementation of PAP actions has been substantially delayed in many cases, which sometimes has had important implications for the timely achievement of the programs' results. Also, in both the country monitoring processes and some aide-memoires, there is a tendency to focus on the DLIs, with less systematic attention to the various PAP items and the indicators in the results frameworks. There is also very little information on the environmental and social effects of the operations in the Bank's portfolio monitoring system – this material is mainly found in the ISRs. This is surprising, since environmental and social risks have been rated as moderate or substantial in most PforRs, and given the essential role of monitoring for the environmental and social management framework and the important place of environmental and social risks within the PforRs' integrated risk management framework.

Another issue relates to the exclusion of high environmental and social risk activities. While IEG has concluded that the high-value procurement exclusion has not been a major obstacle for the achievement of the PDOs, the allowable exclusion of high environmental and social risk activities has at times been interpreted in an overly cautious manner that has led to the PforRs' avoidance of substantial and moderate risk investments (akin to Category B investments in the IPF safeguards) that would normally be integral to the supported programs. As discussed in chapter 4, and noted in the *Two-Year Review* (World Bank 2015e), this exclusion has significantly reduced the scope of several PforRs in relation to that of the supported government programs, and this raises a concern of how such investments will be handled if they are not subject to the oversight associated with Bank involvement.

Political economy constraints are behind many of the issues addressed by DLIs and PAPs, but are rarely discussed directly. Any critical analysis tends to address only low-level issues. However, it should be emphasized that interviews with team members during IEG missions confirm their awareness of such issues. The lack of published analysis is reportedly because of the sensitivity of this area for borrower governments. The Systematic Operations Risk-rating Tool, introduced for PforR operations in January 2016, measures political and governance risk as one of nine risk categories. Another category measures the related aspect of institutional-capacity risk. The use of this new tool is likely to address the political economy challenges raised in this report.

Investing resources up front to ensure that task teams fully understand the mechanisms of PforR can help overcome the initially steep learning curve. In Ethiopia, the first operation was the most difficult and costly to prepare. Both sides are now very comfortable with the instrument, and the cost of subsequent PforRs seems to be coming down, but the instrument is unlikely become advantageous for the Bank on cost grounds. Task team leaders of operations that are to be financed through a PforR need to clearly understand the instrument and be able to explain it to their counterparts (line agencies) and to partners early on, so that they can make an informed decision about whether they want to pursue a PforR operation or use another instrument. The Bank needs to invest sufficient resources up front to ensure that task teams preparing and supervising PforRs master the instrument.

A long-term program with an existing record is well suited to a PforR, especially if the Bank already has experience with the program. It is often an advantage if the programs to be supported have been operational for some time and are expected to be in operation for the long term. It is also an advantage for the Bank to have developed familiarity with the programs and their managements through other vehicles, such as prior operations.

Effective engagement of fiduciary and safeguards staff throughout the process can help with identifying and addressing bottlenecks and shortfalls in country systems. The nature of the PforRs requires intensive Bank staff support during the implementation stage, as the nature of their engagement changes from ensuring or checking that implementing agencies adhere to Bank fiduciary and safeguards requirements to working with the client on how to strengthen capacity to best employ the country's own system, to identify bottlenecks and potential shortfalls in the national system, and to work with counterparts to address these issues during implementation.

Recommendations

Strengthen the design of the results framework and the DLIs to ensure that the PAD presents a clear line of sight to developmental results. It is essential to ensure that all programs produce a well-structured results framework that responds to borrower priorities through a set of logical, achievable steps, reflecting the PDOs and the critical role of DLIs. The DLIs serve partially offsetting objectives, including as triggers for disbursements (requiring predictability) and as incentives for performance (requiring stretch targets). The trade-offs between these considerations should be revisited. This could include steps, such as guidance to staff, to ensure that:

- The results frameworks reflect the development expectations and logic (theory of change) of the PforR within the context of the supported government programs.
- The PDOs reflect measurable, monitorable development results (that is, at the outcome level, which could include institutional strengthening).
- There is a clear explanation of the rationale for the DLI selection and how the DLIs help achieve the PDOs.
- The financing associated with each DLI should reflect, among other things, the possible incentive effect of the size of such financing.
- The individual DLI targets need to strike the right balance between predictability of disbursement and achievement of results.

Strengthen the design and monitoring of the Program Action Plans. The following is recommended to improve the effectiveness of the PAPs, including to mitigate a tendency for the proliferation of PAP actions, which is likely to dilute their impact, and to ensure their adequate consideration during program implementation, when much attention has tended to focus on the DLIs:

- Focus the PAP on a few key areas where actions are important to enhance the capacity and performance of the implementing agencies and to mitigate risks.
- Devote greater attention to the monitoring of PAP implementation, including more systematic coverage in the ISRs.
- Provide clearer guidance to task teams on how to address poor implementation of PAPs.

Minimize the overly cautious interpretation of the high environmental and social risks exclusion. To address the overly cautious interpretation of the high environmental and social risk exclusion:

- Strengthen awareness of the guidance that this exclusion is only intended for high-risk activities, and is not intended to exclude substantial- and moderate-risk activities, such as small- and medium-scale investments that are integral to the supported programs and are likely to cause mostly local and short-term negative environmental and social impacts for which effective mitigation measures are readily available.

Strengthen the monitoring and reporting of results to cover systematically the entire results framework, as well as the environmental and social effects of the projects. To address the tendency in the country monitoring processes and in some aide-memoires to focus mainly on the DLIs, with less systematic attention to other

indicators, PAP items, and environmental and social impacts, all of which should be important for the achievement of the PDOs, it will be important for the Bank to:

- Ensure that the ISRs and supervision reports systematically cover the entire results framework and its supporting elements.
- Ensure that the program systems adequately report on the environmental and social effects associated with the implementation of the PforRs, and that the Bank has timely access to these reports.

Appendix A. Methodology Note for Review of the Bank's PforR Portfolio

Guidance Questionnaire

1. The design of the guidance questionnaire (template) is based on the 2011 PforR Board paper (World Bank 2011a), Bank Procedure/Operational Policy (BP/OP) 9.0 Program-for-Results Financing, and the guidance to staff (World Bank 2012c). It also takes into account the findings and issues raised in the 2015 PforR *Two-Year Review* (World Bank 2015e). The questionnaire will enable panelists to systematically assess PforR programs using the following measures:

- Likelihood of achieving development objectives (highly likely, likely, moderately likely, moderately unlikely, unlikely, highly unlikely)
- Integrated risk rating (high, substantial, medium, low).

2. In addition, the following key risk categories will be rated (low, medium, substantial, high):

- Operating environment risks
- Country risk (the report will identify which were the main factors contributing to country risk – for example, governance, macroeconomic, security, and the like)
- Stakeholder risk (such as vested interests).

3. Program risks:

- Technical risk (program design and governance, institutional capacity, sustainability, and monitoring and evaluation (M&E) arrangements and capacity)
- Fiduciary risk (program fiduciary system and performance monitoring, including audit arrangements and capacity)
- Environmental and social risk (potential environmental and social impacts and ability of program systems to manage these)
- Disbursement-linked indicator (DLI) risk (results framework, selection of DLIs, and verification protocols)
- Other risks (to be identified by reviewers).

4. The above ratings for likelihood of achieving development objectives and various risks will be based on information available at the time of the review and

will be informed by the a more detailed assessment of various program components across the following dimensions, with governance as a cross-cutting theme:

- Strategic relevance
- Technical aspects
- Fiduciary and fraud and corruption aspects
- Environment and social aspects
- Risk assessment
- Program implementation (for programs that are in the implementation stage).

5. An important part of the program review is an assessment of the interplay between the program and its operating environment – for example, the program's environment and social systems versus the national systems. The main dimensions or sections of the Guidance Questionnaire are discussed below. The questionnaire is attached.

6. **Strategic relevance.** This dimension assesses the strategic relevance of the program from the perspectives of the government, as well as Bank corporate and relevant Global Practice priorities. The case for government action or intervention is also examined in this dimension. The program contribution to partnership coordination is assessed as well. Borrower ownership is also reviewed. For programs under implementation, there is more evidence of the degree of borrower ownership and commitment (see also box 1.3 of the *PforR Guidance Notes* [World Bank 2012c] for illustrative indicators for assessing borrower commitment), based on implementation experience. Finally, the value added by the Bank will be reviewed.

7. **Technical aspects.** This dimension assesses the technical soundness of the program – its structure, implementation arrangements, results framework, M&E framework, and capacity to implement and monitor the program. This dimension also includes an assessment of the economic evaluation of the program; the quality of the DLIs and the associated verification protocols; the quality of the Program Action Plan (PAP); and the appropriateness of the program conditionalities and the technical risk rating. The main sources of information will be the technical assessment, as well as implementation experience and Implementation Status and Results Reports (ISRs) for programs under implementation.

8. **Fiduciary and fraud and corruption aspects.** This dimension assesses the fiduciary system, including procurement, financial management, compliance with the Bank's Anti-Corruption Guidelines (ACG), and capacity for implementation and monitoring. With respect to the ACG, reviewers will use the *Bank Guidelines on Preventing and Combating Fraud and Corruption in Program-for-Results Financing*

(World Bank 2012a), as well as attachment 1 of the PforR *Interim Guidance Notes* (World Bank 2012c). The reviewers will also use actual implementation experience, where applicable, in assessing this dimension.

9. **Environmental and social aspects.** This dimension assesses the program arrangements for managing environmental and social effects consistent with OP/BP 9.0. Unlike in an investment lending operation, where the client bears full responsibility for managing all environmental and social consultations with stakeholders, for PforR, the Bank is responsible for managing the consultation process for the program-specific environmental and social systems (see attachment 4.4 of the PforR *Interim Guidance Notes*). The reviewers will also use implementation experience, where applicable, in assessing this dimension.

10. **Risk assessment.** This dimension assesses the appropriateness of the integrated risk rating and the ratings for various operating environment and program risks. In addition, the impact on risk to development outcome of various fiduciary, environment, and social exclusions is assessed.

11. **Program implementation.** This dimension assesses the quality of program implementation for PforR operations that are in the implementation stage. Reviewers will use supervision documents, including ISRs, as main sources of information. In assessing program implementation, the reviewers will identify the various factors contributing to implementation issues, both program-related and those outside the scope of the program.

12. **Ratings.** The reviewers will use a six-point rating scale (highly satisfactory, satisfactory, moderately satisfactory, moderately unsatisfactory, unsatisfactory, and highly unsatisfactory) for the above dimensions. Risk ratings will follow the IEG scale (low, medium, substantial, high).

Panel of Reviewers

13. Each PforR operation will be reviewed by a panel consisting of representatives of the following areas of expertise: (i) the lead reviewer, with expertise in the relevant Global Practice area; (ii) one specialist to cover fiduciary aspects, including anti-corruption; and (iii) one specialist to cover environmental and social aspects. Governance will be a cross-cutting area that will be assessed by the reviewers.

14. The lead reviewer will be responsible for the assessment of the operation, with the specialists providing inputs in their specific areas. The choice of the lead reviewer for each operation will be dependent on the content of the operation.

However, the specialist reviewers will cover several operations to ensure consistency in the assessments.

15. An IEG staff member or consultant will oversee the portfolio review, including quality control of the individual assessments and consistency of the basis for ratings. The IEG staff member or consultant will consolidate the individual PforR program assessments into a report of findings that will include lessons learned and recommendations for improving various aspects of the PforR instrument.

16. The reviewers will have to be familiar with the PforR instrument, guidelines, and issues, in addition to having extensive operational experience. Reviewers will be provided with all relevant materials on the PforR instrument, including the PforR Board reports, BP/OP 9.0 Program-for-Results Financing, *Interim Guidance Notes to Staff on Assessments* (World Bank 2012c), and the Bank *Guidelines on Preventing and Combating Fraud and Corruption in Program-for-Results Financing* (World Bank 2012a).

Documentation

17. An IEG staff assistant will provide the relevant documents for the portfolio review. While most of these are in the Bank system, it may be necessary to get in touch with the task team for some documents, especially correspondence that may not have been included in the document system. The main documents are the following:

- Background documents:
 - Country Assistance Strategy (or similar instrument)
 - Sector Strategy (or similar Global Practice document if available)
 - Poverty Reduction Strategy Paper (for International Development Association [IDA] countries)
 - Relevant economic and sector work (the staff assistant will generate the list of economic and sector work done during the past five years, and the reviewers will determine which outputs are relevant – for example, Public Expenditure Reviews)
 - Operational documents for operations with links to the PforR operations (the staff assistant will generate the list of Bank lending operations during the past five years, and the reviewers will determine which operational documents are relevant)
 - Other documents (such as evaluations) that the reviewers may find relevant.
- PforR operational documents

- Program Concept Note and supporting documents
- Project Information Document
- Management Decision Meeting minutes
- Program Appraisal Document
- Legal agreements
- Negotiation minutes
- Technical Assessment
- Fiduciary Systems Assessment
- Environmental and Social Systems Assessment
- Program Action Plan
- Implementation Support Plan
- ISRs
- Audit reports
- Relevant communications within the Bank and between the Bank and borrowers.

Interviews with Management, Staff, Borrower, Stakeholders, and Partners

18. After a desk review and an initial assessment of the PforR program, the reviewers will identify the information gaps and determine whether interviewing specific persons is necessary, specifying the proposed agenda and questions. The interview can take the form of e-mails, phone calls, face-to-face meetings, or short questionnaires. A limited number of operational reviews will benefit from a country visit by a PforR team.

Quality Assurance of PforR Program Assessments

19. The IEG staff/consultant overseeing the portfolio review will be responsible for the quality of the assessments. There will be a peer review of the individual assessments; peer reviewers will be from the set of panelists used for the portfolio review.

Project Name and Number

Lead Reviewer

Fiduciary Reviewer

E&S Reviewer

Table 1. PforR Review Guidance Questionnaire

Summary Assessment	Comments	Rating
Likelihood of achieving DO		
Overall program risk		
Assessment Dimensions	Comments	Rating
1. Strategic relevance		
2. Technical aspects		
3. Fiduciary and fraud and corruption aspects		
4. Environmental and social aspects		
5. Risk assessment		
6. Program implementation		
7. Bank performance at entry		
8. Bank performance during supervision		
	Comments	Rating
1. Strategic Relevance		
1.1 Program rationale		
1.2 Clarity, realism, and scope of program objectives		
1.3 Appropriateness of partnership arrangements with donors		
1.4 Borrower ownership		
	Comments	Rating
2. Technical Aspects		
2.1 Appropriateness of program structure and implementation arrangements, including borrower capacity		
2.2 Appropriateness of the program expenditure framework, including financial sustainability and funding predictability		

2.3 Quality of the economic evaluation of the program		
2.4. Quality of the results framework and M&E arrangements, including capacity		
2.5 Quality of the disbursement-linked indicators and verification protocols		
2.6 Adequacy of Program Action Plan and conditionalities		
	Comments	Rating
3. Fiduciary and Fraud and Corruption Aspects		
3.1 Adequacy of program procurement system		
3.2 Adequacy of program financial management system		
3.3 Adequacy of program-specific system to handle issues relating to fraud and corruption, based on the Bank's Anti-Corruption Guidelines		
3.4 Adequacy of stakeholder involvement in program implementation and oversight, including verification of disbursement-linked indicators and establishment of a complaint and grievance mechanism		
3.5 Level of transparency in program decision-making processes and performance reporting		
3.6 Appropriateness of measures for strengthening program's fiduciary system and governance, including in the area of anti-corruption		
3.7 Appropriateness of exclusions		
	Comments	Rating
4. Environmental and Social Aspects		
4.1 Adequacy of program systems for managing environmental and social effects		
4.2 Appropriateness of measures for improving program systems for managing environmental and social effects		
4.3 Appropriateness of exclusions		

4.4 Adequacy of national systems for managing any environmental and social effects of government programs

	Comments	Rating
5. Risk Assessment		
5.1 Appropriateness of the program integrated risk rating		
5.2 Appropriateness of technical risk rating		
5.3 Appropriateness of fiduciary risk rating		
5.4 Appropriateness of environmental and social risk rating		
5.5 Appropriateness of DLI risk		
5.6 Appropriateness of other program risk(s), if any		
5.7 Appropriateness of country risk		
5.8 Appropriateness of stakeholder risk		
5.9 Appropriateness of risk-mitigation measures		
5.10 Adequacy of assessment of impact on risk to DO of fiduciary, environmental, and social exclusions, if applicable		
	Comments	Rating
6. Program Implementation		
6.1 Quality of overall program implementation		
6.2 Quality of implementation of fiduciary systems		
6.3 Quality of implementation of environmental and social systems		
6.4 Quality of implementation of Program Action Plans		
6.5 Quality of partnership engagement during program implementation		
6.6 Appropriateness of stakeholder engagement and program transparency during implementation		
Notes:		

Appendix B. Results-Based Programs before PforR

1. The Program-for-Results (PforR) is a new lending instrument introduced in 2012 to complement the development policy financing (DPF) and investment policy financing (IPF) instruments. The PforR instrument supports government programs and disburses against results, while the DPF supports policy and institutional reforms and provides general budget support, and the IPF supports specific projects and disburses against specific expenditures and transactions. More specifically, the PforR:

- Finances expenditures of specific borrower development programs
- Disburses on the basis of the achievement of key results under such programs
- Uses and strengthens the program systems to provide assurance that program funds are used appropriately and that environmental and social impacts are adequately addressed by such programs
- Strengthens the institutional capacity necessary for such programs to achieve their intended results
- Strengthens partnerships by providing a major opportunity to improve coordination among development partners in supporting government programs and provides a vehicle for pooling resources of the government, development partners, and the Bank.

2. Certain high-risk activities are excluded from PforR operations: activities that may have potentially significant and irreversible adverse impacts on the environment or affected people (activities classified as Category A under IPF) and activities that involve procurement contracts exceeding specified monetary amounts.

3. This appendix describes Bank experience with various (non-PforR) IPF operations that include one or more of the characteristics of the PforR instrument. The note also covers major initiatives in results-based financing (RBF), specifically output-based aid (OBA) and government use of RBF approaches in financing public health expenditures. Finally, the experiences of other development institutions (U.K. Department for International Development [DFID], the Inter-American Development Bank, and the Asian Development Bank [ADB]) are described.

Focus on Results and Development Effectiveness

4. The Bank has been embarking on various initiatives and reforms to improve the results focus of Bank strategies and operations. All projects now include results

and monitoring and evaluation (M&E) frameworks. Sector and country assistance strategies are now results-based, with monitorable indicators. Bank projects – in particular, IPF – traditionally have financed inputs, with results (outputs and outcomes) achieved after the bulk of the financing had been disbursed or at project completion. But over the past few years, several Bank (non-PforR) IPF operations have included features that enable or support disbursements against results – a key feature of RBF. Other development institutions and donor organizations have begun to use some form of RBF to establish a stronger link between development finance and results.

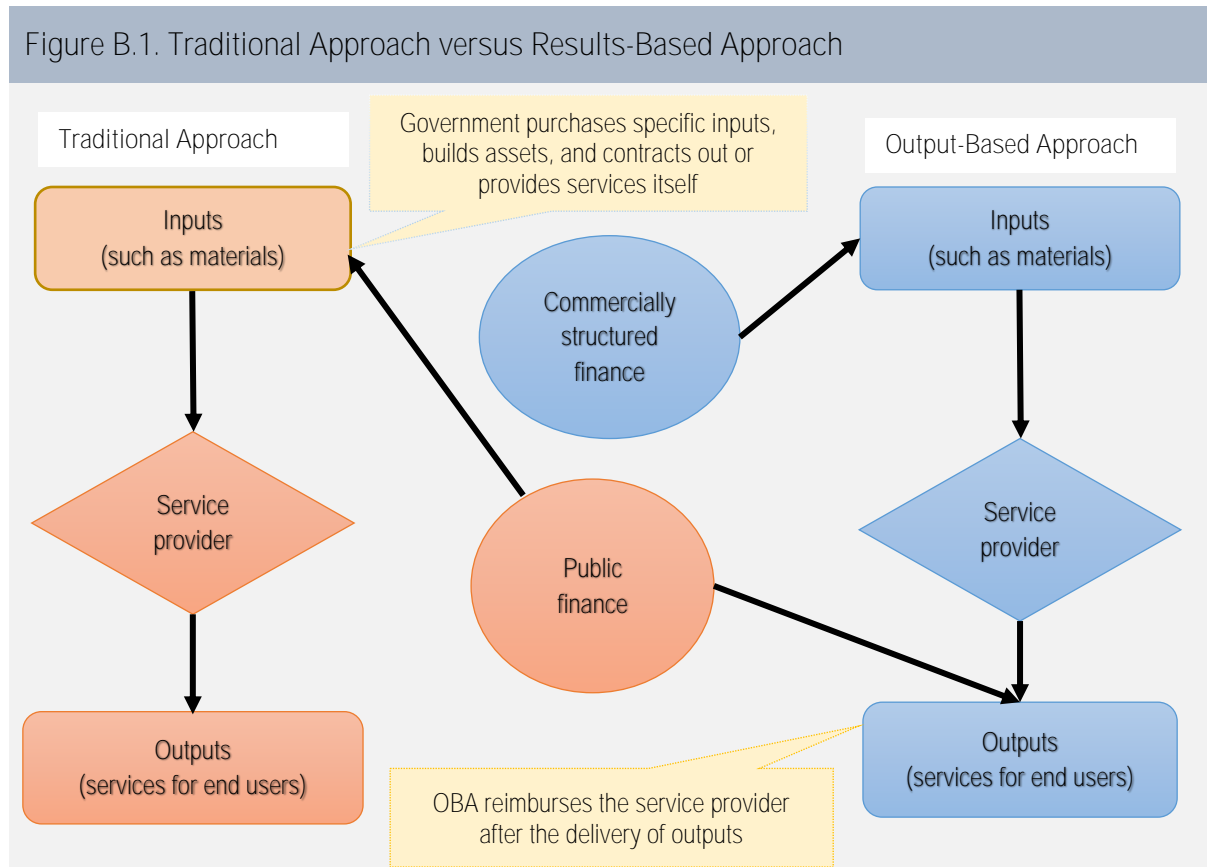
5. One of the key challenges of the RBF approach is how to address the principal-agent problem to ensure that the agent (the party that undertakes the activities to be financed) delivers the intended results sought by the principal (the party that provides the financing). The RBF design would have to include a mechanism that would align the incentives faced by the two parties. In addition, the types of risks faced by the parties would be different in an RBF approach – for example, the agent takes on more of the performance risk (such as financing program implementation with an expectation of disbursements upon completion), and the principal would have less control over the agent's performance.

Output-Based Aid

6. OBA is a “results-based financing mechanism where service provision is contracted to a third party, usually a private sector operator, and subsidies (either one-off, transitional, or continuing) are paid off to the provider after the delivery of specific outputs” (World Bank 2009, p. i). OBA helps finance the gap between the cost of the service delivery and the beneficiaries' ability and willingness to pay user fees for the service. The 2002 Private Sector Development Strategy (World Bank 2002) introduced and recommended the use of OBA as a means of tapping private initiative for delivery of basic services to the poor. In 2003, the Global Partnership on Output-Based Aid (GPOBA) was launched as a Bank-administered donor fund with a view to mainstreaming OBA in the International Development Association (IDA).¹

7. Under an OBA scheme, the contract or other legal arrangement is the mechanism through which the output-based disbursement criteria are established (see figure B.1). Hence, the output has to be well defined and generally under the control of the service provider. To ensure development effectiveness, the causal chain from output (such as number of water connections) to the desired outcomes (for example, reduction in water-borne diseases) and impact (such as improved long-term health) will have to be robust. An underlying rationale for the OBA scheme (that is, the focus on outputs instead of outcomes) is that service providers are not willing to be

accountable for the desired outcomes, which are affected by many factors beyond their control.



8. Output verification is a critical part of the OBA scheme. A review found that the most common way of verifying outputs in OBA projects is to hire independent, specialized consultants with technical expertise to conduct field visits (Mumssen, Johannes, and Kumar 2010). The independent verification agent ensures that only verified outputs are reimbursed by: (a) certifying that the outputs have been physically delivered and the pre-agreed standards of service have been achieved and (b) validating the service provider's reimbursement request (for example, cost reconciliation). The verification protocol and methodology have to be well designed and data sources made available in a timely manner.

9. A *Guidance Note* for procurement staff for OBA projects was issued in 2008 (World Bank 2008). The note identified different scenarios and actions to facilitate compliance with the procurement guidelines in the design of two types of OBA schemes: (a) projects for which there is no existing service provider and (b) projects for which there is already a service provider. Essentially, for projects where there is no existing service provider, a selection will be made through a competitive process.

Downstream procurement carried out by the service provider is not subject to standard procurement post review. For service providers in place but not initially selected through a competitive process, the incumbent will undertake procurement based on Bank guidelines.

10. A key feature of OBA projects is the transfer of risk to service providers. Under an OBA scheme, the service provider, which prefinances the production of the output, bears the performance risk; that is, OBA disburses against verified output, with no financing for implementation (see box B.1). This creates an incentive for the service provider to minimize cost and complete the output in a timely manner. In addition, demand risk can be substantial – there may be less demand than expected for the service. This is a disincentive to the creation of excess capacity, and an incentive to increase access through demand management.

Box B.1. Vietnam Rural Water Supply Development Project

The project objective is to provide access to clean water services through the construction of about 75 new water systems in low-income rural communities in 5 provinces. The project would provide a \$100 subsidy for each household connected, representing about 80 percent of total connection cost. The East Meets West Foundation (EMWF), an international nongovernmental organization operating in Vietnam, is the project implementation entity. The Bank, acting as administrator for GPOBA, entered into a grant agreement with EMWF.

Under the project, EMWF would prefinance the construction of about 75 water systems over the 3 years by borrowing from commercial banks and foundations. The project would disburse 80 percent of the subsidy to EMWF after household connections to water services have been realized and independently verified. The remaining 20 percent would be disbursed after evidence of satisfactory service delivery for at least six months. The water systems would be managed by competitively selected water managers or private operators.

The project includes a financial management action plan with eight actions, four of which were to be completed by project approval. The Bank performed a procurement assessment, which recommended that certain procurement be subject to prior review, and the remainder subject to post reviews. An Environmental Assessment Framework has been prepared in accordance with Bank policy, and an Environmental Management Plan was incorporated in the framework.

Source: GPOBA 2007.

11. A review compared the performance of OBA projects with traditional (input-based) projects with similar objectives (Mumssen, Johannes, and Kumar 2010). Two findings are worth noting. First, OBA provides a stronger platform from which to target infrastructure and social services subsidies than do traditional interventions

in these sectors. Eligibility criteria for beneficiary households are usually clearly defined in OBA, and the output verification process helps provide early evidence that the OBA schemes are reaching the poor. Second, 85 percent of OBA projects reviewed achieved or surpassed desired results within or below budget, compared with 49 percent of traditional projects. Nearly 70 percent of OBA projects were completed below budget, compared with slightly more than half of the traditional projects. The disbursement of funds after service delivery created incentives to deliver outputs in a timely manner. Nonetheless, OBA projects faced several challenges:

- Access to finance can present a hurdle, resulting in difficulty in shifting sufficient performance risk to service providers if the cost of prefinancing the outputs would place an undue burden on the provider.
- Capacity to implement and monitor OBA schemes can be limited, notably in transaction design and implementation, output monitoring and verification, and demand management.
- Some basic institutions and processes that support the development, monitoring, and adjustment of contracts should be in place, including sufficiently transparent legal and regulatory arrangements (for example, tariff setting and adjustment).
- Sustainability of the funding source needs to be addressed, especially in cases where ongoing subsidy schemes are required.

Results-Based Financing for Health

12. RBF² for health is a “cash payment or non-monetary transfer made to national or subnational government, manager, provider, payer, or consumer of health services after predefined results have been attained and verified. Payment is conditional on measurable actions being undertaken” (Musgrove 2011, p. 1). RBF encompasses several types of payment for results (see box B.2). By linking financing to results, RBF shifts emphasis from inputs to outputs and outcomes of health systems, places more of the performance risk on service providers, and gives them incentives to organize and deliver the services more efficiently and effectively. On the demand side, RBF provides payments or transfers to individuals, households, or communities once a precondition in terms of service use has been met. Conditional cash transfers (CCTs) and voucher schemes are examples of demand-side RBF mechanisms.

13. Through the RBF mechanism, donor agencies, ministries of finance, ministries of health, and other government or nongovernmental agencies are able to pay directly for results rather than purchase of inputs. DFID makes a distinction between

RBF and results-based aid (RBA).³ In the DFID formulation, RBF uses government resources in a contractual arrangement between the government and an implementing agent that specifies results to be achieved in return for payment. RBA delivers donor or development partner resources to the government through a contractual arrangement that specifies results to be achieved in return for disbursements. The first part of this section focuses on the experience with RBF, and the second part on RBA.

Box B.2. Glossary of RBF Terms

Performance-based financing is a form of RBF characterized by three conditions: (a) incentives are directed only to providers, not beneficiaries; (b) awards are purely financial, with payment by fee for specified services; and (c) payment depends explicitly on the degree to which services are of approved quality, as specified by protocols for process outcomes.

Performance-based contracting is a form of RBF that departs from simpler types of contracts in setting a fixed price for a desired output by adding a variable component that reduces payment for bad performance and increases payment for good performance, compared with the standard defined in the basic contract.

Cash on delivery is a subset of RBF and is a contract where funders and recipients agree on mutually desired outcomes – rather than inputs or outputs – and a fixed payment for each unit of confirmed progress. The funder uses a hands-off approach. The recipient has complete discretion and responsibility for strategies and programs. Progress toward outcomes is verified by an independent party.

Conditional cash transfer describes demand-side programs where the incentives apply directly to program beneficiaries rather than to agents delivering services. Results are defined by the enrollment of beneficiaries in the program and their compliance with required behaviors, such as consuming specific services.

Sources: Musgrove 2011 and Birdsall and Savedoff 2011.

RESULTS-BASED FINANCING

14. A 2009 review of Health, Nutrition, and Population (HNP) projects approved during fiscal 1995–2008 found that 40 projects supported government health programs involving one or more RBF mechanisms (Brenzel 2009). The range of RBF mechanisms identified in the review included performance agreements between central and subnational government administrative entities; performance-based contracting of public facilities, private providers, and communities; performance-based health worker incentives; vouchers and conditional cash payments; and CCTs. The review identified 24 out of the 40 HNP projects as providing “substantial support” to RBF activities, either because RBF was the focus of the project or the

project had several components related to RBF. The main lessons from the review were the following:

- Political commitment and country ownership are essential. The shift to a results focus requires substantial changes in how ministries and health care providers relate to their work.
- Involvement of all relevant stakeholders in the design of RBF schemes helps to mitigate resistance and facilitate understanding and communication regarding the mechanism.
- Projects designed to increase utilization of health services are more effective when they are complemented by support to improve the quality and quantity of services.
- Several projects were facilitated by complementary reforms, such as decentralization and financial autonomy of health facilities for subnational health authorities.
- Adequate and appropriate M&E frameworks are critical for demonstrating results to stakeholders and for fostering sustainability.
- Independent validation of achievement of indicators linked to performance-based contracts is necessary to mitigate gaming and perverse incentives to overreport.

15. IEG's evaluation of Bank support to health finance noted that a growing share of Bank health operations are supporting reforms of payments to incentivize providers to improve quality and efficiency by increasing the focus on activity or results-based payments (IEG 2014). The report found that Bank support to payment reforms was more effective when integrated with other health-financing functions (risk pooling and revenue collection) and linked to broader institutional reforms. Performance-based payment systems were found to have high overhead costs for verification and were unlikely to improve equity and financial protection.

RESULTS-BASED AID

16. Bank instruments designed to support government health reforms have also been evolving toward RBA (sometimes called results-based lending, or RBL), which links disbursements to performance. In the RBF health projects discussed here, examples of performance criteria linked to disbursements included the number of monitoring systems installed (Argentina Essential Health Functions and Programs Project), number of facilities established (Bolivia Expanding Access to Reduce Health Inequities APL-3), certified enrollment in the program (Panama Health Equity and Performance Improvement Project), number of registrations in a primary health care program (Albania Health System Modernization Project), and percentage of births

attended by skilled personnel (Bangladesh HNP Sector Program). The Argentina Provincial Maternal-Child Health Investment Projects (Phases 1 and 2) linked disbursements to enrollment and output targets (see box B.3). Two projects (Lesotho Health Project and Uganda Reproductive Health Vouchers Project) used the OBA mechanism with funding from GPOBA. All the examples cited are (non-PforR) IPF projects.

Box B.3. Argentina Provincial Maternal-Child Health Investment Projects (Phases 1 and 2)

The Argentina Provincial Maternal-Child Health Investment Projects (PMCHIPs) were approved in May 2004 (Phase 1) and October 2006 (Phase 2). The projects were a follow-on to the Provincial Maternal-Child Health Sector Adjustment Loan, which supported the government's Health Sector Reform Program. The PMCHIP was designed as a three-phase Adaptable Program Loan, with the first phase supporting the nine poorest provinces, and the next expanding to the rest of the country. The main objective is the reduction in the national rate of infant mortality by at least 20 percent (30 percent in the participating provinces) over a 10-year period.

The main component of the PMCHIPs was the capitation payment (about 80 percent of the loan), with disbursements based on submission of audited enrollment lists (60 percent of the capitation payment) and achievement of 10 tracer or output goals (40 percent of the capitation payment). These tracers (for example, proportion of eligible women receiving early prenatal care and the proportion of eligible newborns scoring high in the post-delivery neonatal evaluation) were statistically verifiable from provincial health records. Verification was performed by an independent auditor. The remainder of the loan (20 percent of total) financed goods (such as medical equipment), consultant services, training, and operating costs, which were subject to Bank procurement rules.

An IEG Performance Assessment Report (PAR) that included Phase 1 of the PMCHIP rated the project satisfactory for overall project development outcome, high for design relevance, substantial for efficiency, and moderate for risk to development outcome. The PAR included the following lessons: (i) substantial up-front investments are needed to introduce financial incentives to different levels of government and providers; (ii) financial transfers tied to results require a strong institutional foundation; and (iii) results-linked payments can contribute to better data collection and analysis.

Source: IEG 2011.

Results-Based Lending in Social Protection

17. Social protection was using RBL even before the launch of PforR in 2012. A *Social Protection and Labor Policy Note* (Honorati, Rawlings, and Van Domelen 2011) identified 18 active projects in 2011 that used some form of RBL where disbursements were linked to results. Of the 18 projects listed in the note, 7 were CCT programs – operations supporting CCT programs that disbursed against certain milestones to finance CCT expenditures. Of the non-CCT projects, six had disbursement-linked indicators (DLIs) and five used output-based disbursement (OBD) (World Bank 2007). The use of DLIs was a recent innovation in linking disbursements with agreed results indicators. OBD mechanisms enable investment lending to disburse against delivered outputs or services where procurement methods are satisfactory to the Bank. Social protection used RBL in several areas.

18. **Social safety nets.** RBL in CCT programs typically combines traditional investment lending that finances inputs with RBL mechanisms that finance government transfers when certain milestones are met. In the Brazil Bolsa Familia Project, disbursements and cost-sharing percentages reimburse the Brazilian Treasury for CCTs made under the Bolsa Familia Program and are linked to key milestones of technical improvements. The project design includes an incentive where the cost-sharing percentage increases depending on the actions taken. In addition, the project finances three technical components based on traditional lending arrangements for technical assistance. In Ethiopia's Productive Safety Net Program, the project funds basic operations of the safety net based on the number of people receiving services, and it includes an institutional support component using traditional investment lending mechanisms.

19. **Labor market programs.** Several operations supporting labor-market training programs have begun to finance outputs instead of inputs such as curriculum development and training centers. Examples of outputs include the number of course completions and enrollment of special targeted groups. The Argentina Lifelong Learning and Training Project uses an OBD mechanism to finance payments from the Ministry of Labor to participating organizations based on the number of beneficiaries completing the learning activities. Some projects directly fund the trainee instead of the service provider, as in the case of the Kenya Youth Employment Project.

20. **Decentralized service delivery.** These operations seek to improve service delivery in targeted areas by supporting decentralized service delivery. Ethiopia's Protection of Basic Service Project supports delivery of basic services by subnational governments in five sectors (education, health, agriculture, water supply and

sanitation, and rural roads) by financing block grants for recurrent expenditures. The program uses results-based benchmarks and disbursement triggers to structure the funding. In the Malawi Social Fund, the project finances transfers to local governments based on fully costed community services packages and agreed benchmarks, mainly in terms of local capacity to deliver the service package.

21. **Social protection systems.** There have been RBLs supporting reforms in social protection systems. The Romania Social Assistance System Modernization Project supported reforms based on the government's Social Assistance Reform Strategy. Within the government program, specific programs⁴ would qualify for Bank reimbursement if they met certain criteria for inclusion in the pooled eligible expenditure program (EEP). Disbursements would be triggered by achievement of verified DLIs (see box B.4). The same approach – defining an EEP, disbursing against DLIs – was used in the design of the Albania Social Assistance Reform Project (17 output and 5 action DLIs) and the Moldova Strengthening the Effectiveness of the Social Safety Net Project (13 output and 4 intermediate outcome DLIs). The Albania and Moldova projects included a technical assistance component subject to Bank procurement rules.

Box B.4. Romania Social Assistance Modernization Project

The Romania Social Assistance Modernization Project, approved in March 2012, is a Sector Investment Loan (SIL) that uses a results-based financing approach. The project disbursements are (i) linked to defined eligible line items within the government's program and (ii) triggered by verified achievement of agreed specific results. The project finances a slice (about 10 percent) of the government's program.

The project has 20 DLIs, consisting of 18 outputs (such as adopted action plan for social assistance strategy disseminated) and 2 intermediate outcomes (for example, increase in share of social assistance funds going to the poorest quintile). A specified amount is to be disbursed based on the achievement and verification of the DLIs. The DLIs were chosen from the project results chain. The project drew on lessons from the design of results-based SILs: (i) the importance of a thorough technical diagnostic, including for the expenditure framework; (ii) the importance of clearly defined results chains, with measurable DLIs and clear protocols for verification; and (iii) the provision of ongoing implementation support.

Source: World Bank 2011b.

Sectorwide Approach

22. A sectorwide approach (SWAP) supports a nationally defined program within a sector involving high levels of donor and country coordination toward the

achievement of the program goals. Donor funding could be in the form of parallel or pooled financing, general budget support, or a combination of approaches (Vaillancourt 2009). The SWAP evolved from the Bank's Sector Investment Program instrument, introduced in 1995 (Harrold and Associates 1995). In 1997, a donor meeting focusing on the health sector coined the term SWAP, commissioned a guide (World Health Organization 1995) for its use in the sector, and created an Inter-Agency Group chaired by World Health Organization to foster learning and promotion of health SWAPs. SWAPs were viewed as a vehicle for incorporating the principles of the Bank's Comprehensive Development Framework, introduced in 1999, and for responding to the Millennium Development Goals (MDGs) and various international meetings on aid effectiveness.⁵

23. In 2009, IEG reviewed the efficacy of the approach in the health sector by looking at 18 health SWAPs in 6 countries (Vaillancourt 2009).⁶ These projects had several common characteristics: (a) pooling and joint management of donor funding; (b) mechanisms for coordination between the government and donors, and among donors; and (c) a common M&E framework for measuring program performance used by governments and donors alike, including a mechanism for joint reviews of program performance. The review set out to address three questions:

- Were the anticipated benefits of the approach realized?
- To what extent did the approach facilitate the achievement of national health objectives?
- In what ways did channeling support through a SWAP affect Bank efficiency?

24. The anticipated effects of the approach were: (a) improved sector management and coordination; (b) greater harmonization and alignment of development assistance; and (c) enhanced sector stewardship. The IEG review found that SWAPs generally enabled the development of mechanisms for improved sector management and coordination, with the exception of national M&E systems and capacity, which were a critical weakness. Harmonization and alignment of assistance were achieved through country-led partnership and use of national systems for implementation. However, there were issues with respect to the participation of civil society and transaction costs. With respect to enhanced sector stewardship, while SWAPs contributed to mobilization of resources, there was little evidence of the efficient use of these resources because of gaps in the M&E of efficiency aspects. SWAPs covered by the review had not been effective in establishing mechanisms and incentives to strengthen accountability.

25. The IEG review also found the results focus of the health SWAPs to be modest at best. The strong emphasis on process tasks (such as developing country

systems and coordination mechanisms) distracted attention from health sector performance. The content of the programs was largely input-oriented, with a weak articulation of a results chain that clearly defined and linked inputs, outputs, outcomes, and impact. For example, it was difficult to establish a strong link between improvements in infant and child mortality and the plan supported by the SWAPs. Indicators measuring progress were not coherent with the strategic framework or program of work, and baseline data was poor or missing. Incentives to manage for results and to use results information for decision making were missing. Despite the use of joint program reviews to determine disbursements, the link between sector spending and performance was weak, if it could be detected at all.

26. The IEG review concluded that the ability of the SWAPs to help achieve health sector objectives was mixed. The quality and coherence of many of the programs – a critical factor in achieving health sector objectives – supported by the SWAPs varied considerably across countries. Some programs, while relevant, were complex and overambitious, exceeding the capacities of countries, as in the case of the Malawi health SWAPs. Many of the SWAPs had inadequate assessment of risks and political economy of reforms and provided little guidance on phasing and prioritization of interventions. However, there were improvements in sectoral planning and budgeting capacity.

27. The IEG review included several lessons toward improving design of results-focused SWAPs. First, the program being supported should have a well-articulated results chain, a relevant set of results indicators, and an effective M&E system. Second, management for enhanced development effectiveness requires linking of resources to results. Third, incentives and accountability enforcement – including use of civil society for monitoring – would strengthen the program’s results focus. Finally, assessment of risks and political economy issues would strengthen design and implementation arrangements.

28. In the education sector, the *Education for All* Global Monitoring Report for 2015 (UNESCO 2015) took the position that sectorwide approaches have had mixed success. SWAPs implemented in 25 low-income countries that focused on primary and basic education showed some evidence of success, including efficiency and cost savings brought about by better coordination and flexibility. However, the effectiveness of education SWAPs has been limited by the exclusion of stakeholders other than governments and donors. Nongovernmental organizations have not been involved in SWAPs. In addition, some donors continue to provide parallel aid outside the SWAPs – in 2012, only 7 percent of total aid for education was delivered

in the form of sectorwide support (OECD-DAC 2014). (See box B.5 for the recent introduction of the use of results-based aid by the Global Partnership for Education.)

29. Another paper (McNee 2012) noted that empirical evidence generated by various evaluations of health SWAPs showed that SWAP performance has been mixed. While the underlying theory of change of a SWAP is consistent with principles of aid effectiveness, there was a divergence between the strategic intent of the SWAPs and their implementation. One explanation for the divergence was the focus of SWAPs on technical aspects (national plans, expenditure frameworks, coordination mechanisms, and the like), rather than institutions, incentives, and accountability mechanisms that would enable governments to lead the development process. An Organization for Economic Co-operation and Development (OECD) report (OECD 2012), while not focusing on SWAPs, found that donors focused more on aid effectiveness processes rather than impact, and recommended increased support for country ownership and capacity development.

Box B.5. Global Partnership for Education and Results-Based Financing

The Global Partnership for Education (GPE) provides funds through its Education Sector Program Implementation Grants (ESPIG) to support the implementation of national Education Sector Plans (ESP). GPE also provides technical assistance for the development of new or updated ESPs, but the bulk of its financial support is through the ESPIG.

The GPE introduced a new funding model for the 2015–18 period, whereby the maximum country allocation (MCA) for the ESPIG is divided into two parts – a fixed portion accounting for 70 percent, and a variable portion accounting for 30 percent. Eligible countries could apply for the fixed portion only, or both the fixed and variable portions.

To access the fixed portion, countries must meet three requirements: (i) a credible, endorsed ESP; (ii) evidence of availability of financing for the implementation of the ESP from the government and development partners to complement GPE funds; and (iii) availability of critical data for analyses, monitoring, and evaluation. Countries meeting these requirements could use 70 percent of the MCA to finance inputs and other eligible expenditures toward implementation of the ESPIG.

To access the variable portion, countries must meet additional incentive-based requirements by providing evidence of actions and corresponding indicators that confirm transformative strategies to improve equity, efficiency, and learning outcomes in basic education. Once GPE has verified that the actions have been completed and target indicators reached, GPE would disburse 30 percent of the MCA. This is the results-based financing feature of the MCA.

Source: Global Partnership for Education website.

30. To strengthen the links between disbursement and results, many SWAP operations have begun to use DLIs in recent years.⁷ A review⁸ for this appendix of 10 SWAP operations with a total of 118 DLIs found that 56 percent of the DLIs were action or process indicators, 24 percent were output indicators, and 20 percent were intermediate outcome indicators. The large share of action or process indicators points toward linking disbursements to efforts in improving systems and institutions that support the programs.

Disbursement-Linked Indicators in (Non-PforR) IPF Projects

31. In recent years, the use of DLIs has been growing in (non-PforR) IPF projects. There is no comprehensive list of such projects, but a 2013 paper (O'Brien and Kanbur 2013) compiled a list of 29 projects (PforRs and IPFs) using DLIs that had been approved since 2006. The 20 (non-PforR) IPF projects⁹ on the list had a total of 237 DLIs. Most of the DLIs for (non-PforR) IPF projects were actions or outputs, with fewer than 30 percent in the intermediate outcome category (see table B.1), compared with 40 percent for PforRs (IEG 2016). About 95 percent of the (non-PforR) IPF project amounts were to be disbursed using DLIs, with all except one project having one or more components supporting institutional-strengthening measures.

32. A common feature of (non-PforR) IPF projects with DLIs is the definition of eligible expenditures to be financed – these expenditures are either not subject to Bank procurement rules (salaries, for example) or are below the threshold for Bank procurement prior review. In addition, these projects include technical, fiduciary, and environmental and social assessments. Verification is typically done by government agencies, although in one case an independent auditor was used as an additional check. The non-DLI project components were subject to Bank procurement rules (that is, the traditional input-based approach was used in financing these components).

Table B.1. DLIs by Type in IPF

	Number of projects	Number of DLIs	Share of total DLIs by type		
			Action (%)	Output (%)	Intermediate outcome (%)
SWAPs	10	118	56	24	20
Non-SWAPs	10	119	13	53	34
Total	20	237	35	38	27

Sources: Project PADs. IEG allocation to types of DLIs.

Experience of Multilateral and Aid Organizations

33. **U.K. Department for International Development.** DFID has been increasing its use of payment by results, which is characterized by: (a) disbursements tied to clearly specified results; (b) recipient discretion in how the results are to be achieved; and (c) robust verification of results as triggers for disbursement. DFID differentiates payment by results according to organizations receiving payment: (a) payments from funders to partner governments are classified as RBA; (b) payments from funders or governments to service providers are classified as RBF; and (c) returns to investors in development impact bonds based on results of a program being funded by the instrument. In the 12-month period ending September 2013, 71 percent of contracts for services issued by DFID had a performance-based element. About a third of financial support to governments is linked to performance. DFID is increasing use of outcomes as the basis for payment – as of June 2014, DFID had 21 outcomes-based programs (such as completion rates in education, safe deliveries for women, access to water points), with 17 more in the planning stage (DIFD 2014a).

34. A DFID paper emphasized that payment by results will not always be the best instrument for delivering aid (DFID 2014a). Based on early experience with the approach, two considerations in determining whether to use payment by results were identified. First, targeting specific actions to improve overall performance may be more appropriate than paying for results, especially where the approach is based on recipient capacity to determine how the results are to be achieved. Second, the potential benefits of using the approach may not outweigh the costs relative to other mechanisms and the risk that the recipient will not be able to take on the performance risk. Nonetheless, the paper provided early indications of benefits of the approach: (a) improvement in results data; (b) strengthening of empowerment and accountability; and (c) greater focus on performance and efficiency. To systematically determine what works, DFID is strengthening its evaluations of payment-by-results programs and projects – including providing a framework for the evaluations (DIFD 2014b).

35. **Inter-American Development Bank.** In 2003, the IDB approved a six-year pilot program of performance-driven loans (PDL), which would improve development effectiveness by emphasizing verified development outcomes instead of focusing on project inputs (IDB 2003). The PDLs disburse against achievement of outcome targets verified by independent consultants and verification that the expenditures subject to the disbursement were used to achieve the outcomes. PDLs require well-functioning performance measurement and monitoring systems, and they allow use of fiduciary country systems acceptable to the IDB. Borrowers are required to present documentation that procurement was done based on acceptable IDB procedures. The PDL program suffered from low demand because it had

greater disbursement requirements than other investment lending instruments. The last PDL was approved in 2009, and the pilot program was allowed to lapse pending further review and evaluation.

36. There were reviews of the experience from the 17 PDLs approved during the life of the pilot program (IDB 2014a). In Guatemala, there was a mismatch between the PDL and program scope – the borrower did not have the resources to achieve the outcomes. In Honduras, there was a mismatch between the PDL and country capacity – an investment loan was subsequently approved to strengthen capacity. More generally, there were issues with the verification of outcomes and expenditures: (a) difficulty in linking outcomes and expenditures; (b) lack of clarity in the roles of financial auditors and results auditors; and (c) difficulty in finding capable firms to verify outcomes. Partial achievement of agreed outcomes caused delays in disbursements – partial disbursements were not allowed. Most borrowers did not meet IDB requirements for fiduciary systems, and instead used IDB procurement systems. Finally, PDLs were designed around outcomes; outputs were not included in the results against which the PDL would disburse (IDB 2014b).

37. **Asian Development Bank.** The ADB approved the piloting of results-based lending for programs in 2013, with the pilot to last six years (ADB 2013). The RBL would support government-owned sector programs, link disbursement with program results, and use program systems after they have been assessed by the ADB as conforming to accepted good principles. To date, five RBLs have been approved, four in the education sector and one in the health sector. There were some early lessons from the experience (ADB 2015). First, the quality of the DLIs is critical, since these are important tools for system improvement. Second, fiduciary systems and safeguards assessments are important inputs toward strengthening the institutional systems required for implementing RBLs, but the ADB should not use a cookie-cutter approach to these assessments. Third, information-sharing about experience and best practices in RBL would help provide feedback to strengthen the effectiveness of the instrument – the ADB has established RBL champions. Finally, the ADB will work with partners during program implementation, focusing on issues relating to how to achieve program results and improve systems, rather than on project inputs.

Findings

38. There has been an evolution in the use of results as a basis for disbursement in Bank IPF projects. Various types of results have been used – outputs, performance measures, milestones, completed actions, and intermediate outcomes.

39. RBF disburses against outputs and intermediate outcomes, based on the Bank's assessment that the borrower has the capacity to implement an agreed project or program consistent with Bank standards. In theory, the borrower takes on the performance risk. In practice, the Bank mitigates that risk through various mechanisms of involvement in the implementation of the project or program, including implementation support by Bank staff and accompanying technical assistance and institution-strengthening loan components.

40. A common characteristic of (non-PforR) IPF projects with DLIs is the definition of expenditures eligible for reimbursement, which would not be subject to Bank procurement prior review. One of lessons from the IDB experience with its results-based instrument is the need for borrowers to meet fiduciary standards.

41. Fiduciary and environmental and social assessments in (non-PforR) IPF projects are not performed systematically, unlike the practice for PforRs, where there are clear guidelines. In addition, strengthening fiduciary and environmental and social institutions is not an area of focus of (non-PforR) IPF projects.

42. There are lessons from the use of results-based instruments that were known prior to the introduction of PforR, including the importance of: (i) thorough technical diagnostics, including for the expenditure framework; (ii) clearly defined results chains with measurable DLIs and clear protocols for verification; and (iii) the provision of ongoing implementation support.

43. Compared with PforRs, the IPF projects with DLIs have greater flexibility in project design, in particular the ability to combine different approaches (input and results-based) in one instrument, depending on client needs. There has not yet been any formal guidance for staff using results-based approaches or DLIs for IPFs, but IEG understands that such guidance is now in preparation.

44. There has not been a systematic review of the effectiveness and efficiency of various results-based approaches, including whether Bank guidelines have kept up with evolving practices. Performing this review would generate important lessons (including on the design of DLIs, fiduciary arrangements, implementation support, and conditions under which a results-based approach would be effective).

45. The preceding analysis emphasizes that the PforRs represent an evolution rather than a sharp departure from previous Bank instruments, with many overlapping features. While the SWAPs have many of the features of a PforR instrument, the latter provides greater flexibility in strengthening programs, promotes interdepartmental cooperation, and minimizes the cost of dealing with compliance requirements. However, these advantages may also have to do with the

design of SWAP programs, such as the imposition of additional disbursement conditions beyond the achievement of DLIs. As an instrument, the PforR has certain advantages over the SWAP: stronger verification systems for disbursement, more rigorous assessments of country systems, and greater emphasis on improving program and country systems and capacity through the Program Action Plans (PAPs).

¹ The *IDA 15 Mid-term Review* (World Bank 2009) identified 18 OBA projects financed by IDA and 17 projects financed by GPOBA during IDA 14.

² Pay for performance and performance-based payment and performance are considered synonyms for RBF.

³ The Asian Development Bank uses the term results-based lending for its new instrument that links disbursements with results.

⁴ State Child Allowance, Indemnity for Child Raising, Family Allowance Program, and Guaranteed Minimum Income Program.

⁵ For example, the 2005 Paris Declaration, to which the Bank is a signatory, emphasized the following principles that were consistent with SWAPs: country ownership, alignment of donor support with the country situation, harmonization of their support with country systems, managing for results, and mutual (that is, country and donors) accountability.

⁶ The six countries were Bangladesh, Ghana, Kyrgyz Republic, Nepal, Malawi, and Tanzania. In five countries, IEG had field studies as part of the preparation of Project Performance Assessment Reports or case studies for the preparation of an IEG health sector evaluation (IEG 2009). In one country, the review used the findings of an evaluation of Tanzania's health SWAP commissioned by the government and financed by development partners.

⁷ None of the SWAP operations reviewed in the IEG paper on Health, Nutrition, and Population SWAPs used DLIs.

⁸ The review used the list of SWAP operations with DLIs in O'Brien and Kanbur (2013).

⁹ The sectoral composition of the projects was: Social Protection – 6; Health, Nutrition, and Population – 4; Education – 3; Infrastructure – 1; Public Sector Management – 1; and Multisector – 5. Half of the projects were SWAPs.

Appendix C. PforR Portfolio

Table C.1. Summary of Approved PforR Operations as of June 30, 2015 (US\$, millions)

Country	Program name	Board approval	Global Practice	IDA/IBRD	PforR Bank financing	Other donor financing	Total program cost
Bangladesh	Revenue Mobilization Program-for-Results: VAT Improvement Program	5/9/2014	Governance	IDA	60	0	73
Brazil	Strengthening Service Delivery for Growth, Poverty Reduction and Environmental Sustainability in the State of Ceará	11/21/2013	Trade & Competitiveness	IBRD	315	0	416
Croatia	Health System Quality and Efficiency Improvement Program	5/8/2014	Health, Nutrition & Population	IBRD	103	0	248
Egypt	Inclusive Housing Finance Program	5/5/2015	Finance & Markets	IBRD	500	0	1,982
Ethiopia	Health Millennium Development Goals Program	2/28/2013	Health, Nutrition & Population	IDA	100	556	676
Ethiopia	Local Government Development II Program	5/2/2014	Social, Urban, Rural & Resilience	IDA	380	0	557
India	Third Maharashtra Rural Water Supply and Sanitation Program	3/12/2014	Water	IDA	165	0	235
India	Enhancing Teacher Effectiveness in Bihar	5/19/2015	Education	IDA	250	0	357
Kenya	National Integrated Safety Net Program	7/23/2013	Social Protection & Labor	IDA	250	261	952
Mexico	Oaxaca Water Supply and Sanitation Sector Modernization Program	6/6/2014	Social, Urban, Rural & Resilience	IBRD	45	0	93.5
Moldova	Health Transformation Program	5/22/2014	Health, Nutrition & Population	IDA	28.7	0	114
Morocco	National Initiative for Human Development (INDH) Phase II	6/28/2012	Social, Urban, Rural & Resilience	IBRD	300	0	1,115
Morocco	Improving Primary Health in Rural Areas	4/24/2015	Health, Nutrition & Population	IBRD	100	111	226.2
Mozambique	Public Financial Management for Results Program	6/24/2014	Governance	IDA	50	0	130.6
Nepal	Results-Based Bridges Improvement Program	6/28/2012	Transport & ICT	IDA	60	0	148
Nigeria	Program to Support Saving One Million Lives	4/23/2015	Health, Nutrition & Population	IDA	500	0	1,052
Pakistan	Punjab Governance Reforms for Service Delivery	11/14/2013	Governance	IDA	50	0	77
Rwanda	Transformation of Agriculture Sector Program	10/31/2014	Agriculture	IDA	100	800	1,200
Rwanda	Public Sector Governance Program	10/31/2014	Governance	IDA	100	30	172
Tanzania	Urban Local Government Strengthening	10/23/2012	Social, Urban, Rural & Resilience	IDA	255	0	255

APPENDIX C
PFORR PORTFOLIO

Tanzania	Big Results in Education	7/10/2014	Education	IDA	122	130	416
Tanzania	Strengthening Primary Health Care for Results Program	5/28/2015	Health, Nutrition & Population	IDA	200	350	2,620
Tunisia	Urban Development and Local Governance Program	7/24/2014	Social, Urban, Rural & Resilience	IBRD	300	0	363
Uganda	Support to Municipal Infrastructure Development	3/28/2013	Social, Urban, Rural & Resilience	IDA	150	0	160
Uruguay	Road Rehabilitation and Maintenance	11/13/2012	Transport & ICT	IBRD	66	160.5	510
Vietnam	Results-Based Rural Water Supply and Sanitation	11/1/2012	Water	IDA	200	0	260
Vietnam	Northern Mountains Urban Program	6/5/2014	Social, Urban, Rural & Resilience	IDA	250	0	300
Total					4,999.70	2,398.50	14,708.30

Note: Data for Bank financing, other donor financing, and total program cost can be found in PADs (data pages on financing source). PADs can be found at <http://www.worldbank.org/en/programs/program-for-results-financing#2>

Table C.2. Approved PforR Operations in Fiscal 2016 as of March 31, 2016 (US\$, millions)

Country	Program name	Board approval	Global Practice	IDA/IBRD	PforR Bank financing	Other donor financing	Total program cost
Burkina Faso	Public Sector Modernization Program	7/9/2015	Governance	IDA	40	6	59.3
Ethiopia	Enhancing Shared Prosperity through Equitable Services	9/15/2015	Social Protection & Labor	IDA	600	257.2	8,010.5
Kenya	Statistics Program-for-Results	9/10/2015	Poverty & Equity	IDA	50	3	138.4
Kenya	Devolution Support Project	3/15/2016	Social, Urban, Rural & Resilience	IDA	200	0	287.3
Vietnam	Results-based Scaling Up Rural Sanitation and Water Supply Program	11/12/2015	Water	IDA	200	25.5	226
Pakistan	Punjab Jobs & Competitiveness P4R	3/31/2016	Trade & Competitiveness	Blend	100	0	280
China	Financing for Air Pollution Control	3/22/2016	Energy & Extractives	IBRD	500	0	1,000
Costa Rica	Strengthening Universal Health Insurance	3/15/2016	Health, Nutrition & Population	IBRD	420	0	1,575
Egypt	Sustainable Rural Sanitation Services Program-for-Results	7/28/2015	Water	IBRD	550	530	1,250
India	Swachh Bharat Mission Support Operation	12/15/2015	Water	IBRD	1,500	0	22,000
Morocco	Urban Transport Project	12/9/2015	Transport & ICT	IBRD	200	0	350
West Bank and Gaza	Local Governance and Services Improvement Program	11/2/2015	Social, Urban, Rural & Resilience		5	15	20
Total					4,365.0	836.7	35,196.0

Note: Data for Bank financing, other donor financing, and total program cost can be found in PADs (data pages on financing source). PADs can be found at <http://www.worldbank.org/en/programs/program-for-results-financing#2>

Appendix D. Disbursement-Linked Indicators

1. The PforR instrument was introduced through the 2011 Board paper (World Bank 2011a). This instrument disburses the Bank's loan or grant funds upon achievement of program results through the disbursement-linked indicators (DLIs). This appendix uses the following definition for DLIs – taken from the *Guidance Notes to Staff* (World Bank 2012c, p. 16):¹ “Disbursement-linked Indicators (DLIs) are a limited set of key indicators that are considered critical toward achievement of the program development objectives (PDOs) and would be the basis for disbursement of World Bank funding. They may also include key performance actions deemed necessary to strengthen the performance of the government's financial management and procurement processes, environmental and social aspects, oversight and controls (including integrity systems), monitoring and evaluation of the program.”
2. The Board paper (World Bank 2011a) explains (as do the *Guidance Notes*):
 - DLIs can be outcomes; intermediate outcomes, outputs, or process indicators; or financing indicators.
 - They can also be key actions aiming to address specific risks or constraints to achieving the results.
 - Their selection would be driven by the desired outputs and outcomes of the programs, while considering the practical aspects of measuring, monitoring, and verifying the achievement of the results.
 - DLIs will be tangible and verifiable, and a given DLI can have one or more value to be achieved over the lifetime of a PforR operation.
3. Each DLI is required to have a credible verification protocol that will define the DLI and set out how it will be measured, whether it is discrete (all or nothing) or scalable, how the verification will be done, and who will be responsible for verifying its achievement. Verification can be carried out in a variety of ways and by various parties, including government agencies and third-party entities, with the primary objective of ensuring that a credible mechanism is in place for monitoring, measuring, and verifying the achievement of the DLIs.
4. In June 2012 Bank management issued the *Guidance Notes* referenced above, which supplemented the Board paper (World Bank 2011a) discussion of the following key points – considerations that clearly may require the task teams to compromise in their DLI designs in consideration of the following:

- While DLIs will vary in nature, they should be driven by desired outcomes or outputs.
- The selection of DLIs should focus on the indicators that provide evidence of continued progress toward the PDO, and the choice of a DLI should be clear with respect to its signaling purpose. As the notes emphasize: “Does the DLI serve the role of signaling and monitoring a critical milestone along the results chain without which the PDO could not be achieved? Or does the DLI serve the role of signaling incentives for rewarding performance (outputs, outcomes) to encourage the practice of managing for results?” (World Bank 2012c, p. 34).
- The DLIs should also be clearly defined and measurable, with clear protocols for monitoring. They should take into account the country’s context and borrower capacity and the feasibility of the results within the PforR period.
- The design should take into consideration the borrowers’ need for budget predictability and flow of funds.
- Maintaining a simple design for the operation is also critical.
- Finally, some DLIs may have a final achievement date, and others will be achievable during the program period, and some DLIs may be scalable, and others (such as a specific action) are either achieved or not.

The Relationship of the DLIs to the Results Frameworks

DLIS ARE NOT THE RESULTS OF A PFORR

5. During program preparation, and especially during implementation, the DLIs may get the most attention, since their progress will determine the pace of disbursements. However, a comprehensive assessment of the results of a PforR must cover several dimensions, with the PforR results framework at the core. The 2011 Board paper (World Bank 2011a) on PforRs makes clear that—in keeping with the instrument name—these are programs aimed at achieving results, but the paper otherwise does not provide much discussion of what should be understood by results, beyond the brief discussion of the wide range of possible DLIs. For instance, there is virtually no discussion of any possible sustainability requirements. The *Guidance Notes* stipulate that results are “the output, outcome, or impact of a development intervention. In general, the Bank seeks to encourage results that support sustainable improvements in country outcomes—that is, evident changes in peoples’ lives, and/or the behaviors of targeted organizations, households, or firms” (World Bank 2012c, p. 16). There is thus an expectation that the DLIs will be important to the achievement of results, but with several considerations that may serve to limit the strength of the linkages between the results matrix and the DLIs.

DLIS AND THE RESULTS FRAMEWORKS

6. The PforR instrument finances the implementation of a program, and the support for program implementation should, in turn, determine the choice and design of the DLIs. Unlike output-based aid (OBA), where financing is directly linked to outputs or outcomes (that is, program/project implementation is financed elsewhere), the PforR instrument has to find a way of designing disbursements through DLIs to ensure that there is sufficient flow of funds during program implementation – this is probably one of the factors that explains the number of action/process DLIs (also the significant use of advances). An additional issue is that the absence of a clear theory of change in most project appraisal documents (PADs) makes it difficult to discern with precision the likely linkages between the individual DLIs and the eventual expected results, according to the PAD PDOs and supporting PDO indicators (with the various intermediate indicators in the results frameworks).

7. There are, however, some PforRs with quite close and clear integration between the results framework and the DLIs (and also with the Program Action Plan [PAP]).² One example is the Croatia Health Sector Reform PforR, where the DLIs are fully integrated into the results framework matrix, and with key PAP actions directly linked to the achievement of some DLIs. In that program, six DLIs are linked to the reorganization and management improvement components of the program, three are linked to the quality and preventive care components, and one to the financial stability of the hospital system.

8. DLI sets may be more or less representative of the results frameworks, and the linkages between the two may be more or less clear. In two cases, the DLI set consists of all the indicators in the results framework for their PforRs, and these DLI sets are thus fully representative of the results frameworks. In all other cases the DLI sets include some indicators not in the results framework. For some PforRs, such as the Ethiopia Health MDG Support Operation and the Bangladesh Value Added Tax Improvement Program, the DLI sets are well structured to represent both the results chains and the set of indicators in the results framework. However, there are other cases where a heavy focus of the DLIs on intermediate steps is less representative of the expected results in the results matrix, and may also be less directly linked to program objectives. Box D.1 shows one example.

Box D.1. Example of a DLI Set Different from the Results Framework

In the Tanzania Big Results in Education Now Program, there is a clear difference between the results framework (PAD annex 2) and the set of DLIs (PAD annex 3):

- The PDO is to improve education quality in Tanzanian primary and secondary schools, and the results framework supports this through appropriate PDO indicators that include reading levels, subtraction levels, teachers found in classrooms, and teachers' knowledge levels.
- But the six DLIs are overwhelmingly (five DLIs, 80 percent by amounts) institutional – completion of foundational activities, resource flows, annual reporting, deployment of teachers across districts, and schools receiving incentive grants. Only one DLI addresses education results – for reading improvements.
- It is thus hypothetically possible for all DLIs to be met in full, and on time, with 100 percent disbursements, but with only one of the four PDO indicators (reading levels) being met. (However, management believes in this case that the DLIs will support the achievement of the PDO indicators.)

Source: World Bank 2014d.

9. Some DLIs may, in effect, give preference to less challenging areas. The focus in the DLIs on achieving results may sometimes encourage concentration on areas or target groups where results toward specified criteria may be easier (or less difficult) to achieve. Thus, based on IEG's discussions during its visit to the ongoing Vietnam Rural Water Supply and Sanitation (RWSS) PforR, the DLIs for reaching commune-wide sanitation targets appear to have led to the selection of the more capable communes by the participating provinces in the first years. These communes were already close to the agreed aggregate targets, while communes that were left for the later years of the program had farther to go to reach the targets. This could point to a possible conflict between aiming for "low-hanging fruit" or for trying to reach the more disadvantaged communes or population groups.

Aspects of the DLIs

DISBURSEMENTS CAN BE FRONT-LOADED

10. The Board paper (World Bank 2011a) offers two timing options that would encourage early disbursements:

- *Disbursements against prior results:* In some circumstances, certain results may need to be realized before the legal agreement is signed for the desired program results to be achieved (for example, establishing monitoring mechanisms or setting baselines). In such cases the Bank will be able to disburse against DLIs achieved between the date of program concept review

- and the date of the legal agreement, up to an aggregate amount of 25 percent of Bank financing. The Concept Notes amplify the need for the task teams to ensure that such results are within the scope of the program and that the systems used to achieve such results are assessed by the Bank in adherence with Operational Policy/Bank Procedure (OP/BP) 9.00.
- *Rolling advances:* Advances can also be considered to achieve not only the initial set of DLIs, but also subsequent DLIs during the implementation period, in cases where advances may be helpful or necessary for the borrower to finance the activities needed to achieve the results of one or more of the DLIs. Such advances should normally not exceed 25 percent of total PforR financing. The amount is deducted from the total to be disbursed under a subsequently met DLI, and further advances can then be made once an advance has been recovered, or partially recovered, as long as the overall limit is not exceeded. The Concept Notes add that the need for and justification of advances should be described in the PAD, and the specific amount and allocation should be specified in the Financing Agreement.
 - *Maximum advances:* The combined amount of financing under the two bullets above may not exceed 50 percent of PforR financing.

11. Disbursements have been very fast in several instances as a result of the above two options, plus the general structure of the DLIs, including the potential to scale up. As one example, the two ongoing programs in Rwanda, both made effective in December 2014, had, by November 2015, disbursed well over 60 percent of their Bank commitment amounts. But generally the pace of early disbursements has varied considerably. Thus, for the 20 programs that became effective before January 1, 2015, as of mid-November 2015, cumulative disbursements, including advances, varied between more than 60 percent (four programs) and no or very low disbursements (three percent or less – also four programs).

DO DLIS REPRESENT “**S**TRETCH” TARGETS?

12. The *Two-Year Review* (World Bank 2015e) did not discuss the degree of difficulty or stretch in the DLIs, but noted (in paragraph 37) that the formulation of the DLIs (and the associated disbursement mechanisms) had been one of the most challenging aspects of the preparation of PforR operations. This evaluation has not been able to make a general determination of the degree of stretch in the DLIs, but field missions noted several cases where a number of the DLIs seemed relatively easy. Box D.2 shows one example of an apparently quite unambitious set of DLIs.

Box D.2. Vietnam—DLIs in the Results-Based National Urban Development Program in the Northern Mountains Region

The program has eight DLI indicators, divided among four DLIs:

- **DLI 1.1: Enhanced annual city plans approved and disclosed to the public.** These plans to be produced annually. Presumably, once a plan has been produced, it will be easy for the cities to modify it as appropriate in the following years.
- **DLI 1.2: Professionally staffed management units in place within each participating city People’s Committee.** Presumably, once a unit is in place, it will be straightforward to maintain it.
- **DLI 2: Local urban infrastructure investments delivered as per each participating city’s approved enhanced annual city plan.** One of two DLI indicators concerning outputs on the ground.
- **DLI 3.1: Asset management plan adopted and local urban infrastructure subprojects in full service after completion.** This indicator also concerns outputs on the ground. For the first two years, the indicator concerns asset condition assessment and asset management plans. From the third year, condition that at least 80 percent of completed local infrastructure subprojects are free of physical damage and operational.
- **DLI 3.2: Increased annual own-sources revenue in participating cities.** The requirement (12 percent increase in nominal terms) may not appear difficult considering an inflation rate of 6-7 percent and a GDP growth rate of 7 percent in 2014, but at least one participating city is concerned because of limitations on their revenue-generation possibilities by central government policies.
- **DLI 4.1: Implementation strategy for National Urban Development Program adopted with annual milestones.** A set of gradual milestones, with approval of a Policy Note in year one and management capacity in place in year four.
- **DLI 4.2: Professionally staffed unit in place in the Ministry of Construction, preparation of annual capacity development plans, and capacity-building support provided to cities in accordance with such plans.** So the degree of progression under this DLI will depend on the degree of progression of the annual capacity building. No apparent progression in requirements after year one, but annual disbursements based on continuation of the requirement.
- **DLI 4.3: Completed program report.** An annual report.

Of the eight indicators, six are thus institutional (and therefore under the control of the central or city authorities), and two concern the investments on the ground, with a modest degree of apparent progression over the years.

Source: World Bank 2014e.

HOW STRONGLY DO DLIS INCENTIVIZE PERFORMANCE?

13. The role of DLIs to incentivize performance is mentioned regularly in PforR PADs, in particular in the risk assessments. There can be two aspects to incentives. A well-structured set of DLIs responding to borrower priorities will clearly help to

encourage performance through a set of logical, achievable steps, and this has been observed in some cases by IEG field missions. But it is a separate question whether the prospects of an amount of financing will encourage prompt and timely actions—and the field missions saw examples where such prospects did not appear to do so. And if such incentivizing should be present, then large early disbursements may possibly have an adverse effect on motivation later in some programs. A related question is whether small DLIs (in dollar terms) will have much of an incentive effect. As one example, for the relatively recent Vietnam National Urban Development Program, one DLI accounts for 1 percent of total financing, one for 3 percent, and two others for 4 percent each. At the other end of the scale, one DLI is expected to account for 62 percent of overall disbursements. It is too early to assess whether the “small” DLIs will have much of an incentivizing effect—but there is reason to be somewhat skeptical.

DLI Risks and Risk Management

14. The integrated risk assessments in the PADs show that the DLI risks are not seen as particularly high, and generally as a bit lower than for the overall program risks. In 1 case the DLI risk was rated high (against 4 programs for the overall program risks), in 8 cases substantial (16 for overall program risks), and in 18 cases moderate (7 for overall program risks). The DLIs are thus normally not seen as risk drivers.

15. There are no dominant reasons for a DLI risk rating of substantial or high, but such reasons include that a program or institutions are new, possible concerns that the DLI targets may be too ambitious, linkages to technical risks, or possibilities that delays in disbursements may affect participants’ motivations.

16. At the same time, mitigating measures to deal with DLI risks tend to focus on the same aspects across the PforR operational spectrum: the care taken to define the DLIs, role of the verification mechanisms, the incentives for participating entities from targeted DLI disbursements, use of scalable DLIs, and that the DLI targets have been set carefully to help ensure achievement, and thus disbursements. From these discussions it could be inferred that the DLIs generally do not use stretch targets—that attention to timely disbursements may have been more important than possible ambitious performance incentives. See one illustration in box D.3.

Box D.3. An Example of DLI Risk Mitigation

The Morocco National Initiative for Human Development PforR, in PAD annex 7, lists as DLI risk-mitigation measures factors that ought to be part of a program under any circumstances (*italics used here for emphasis*):

“DLIs have been firmed up during appraisal. *Key risk management measures* included: (a) using *straightforward and scalable DLIs* as much as possible; (b) *exert realism in setting DLI*; (c) *computing DLI targets in a conservative fashion*; (d) *building prudent scenarios in terms of the program budget ceiling*; (e) *ensuring budget predictability*; and (f) *seeking synergies and alignment with other donors ... results indicators/DLI* in order to alleviate reporting burden on client. Risk management during implementation will include (a) continuous monitoring of DLI progress; (b) strengthening capacities in weak areas potentially affecting DLI achievement” (World Bank 2012b, p. 86).

Source: World Bank 2012b.

17. DLIs are sometimes used to mitigate other risks. The role of DLIs to incentivize performance is mentioned for a number of programs. It remains to be seen whether these incentives will work as intended in all instances. DLIs are sometimes also used as tools for mitigating other or overall risks, but less frequently and less importantly than overall program structure and preparations, or the PAPs. Examples of the use of DLIs for risk mitigations include:

- **Bangladesh VAT:** Mitigating technical risk of complex procurement by having the DLI specify the process to be followed.
- **Ethiopia Health:** Fiduciary risk addressed by DLI to improve transparency and increase use of competitive bidding.
- **Ethiopia Urban Local Government:** Technical risk of lack of motivation among regional governments addressed through DLI amounts to incentivize for the delivery of results. Environmental and social risk of quality of implementation addressed by DLI providing resources to support performance and for backstopping.
- **Mexico Oaxaca Water Supply and Sanitation:** Technical risk addressed by DLI designed to generate financial incentives.
- **Nepal Bridges:** Technical risk addressed through DLIs that will only disburse for investments that follow priorities.

DLIs as Discussed in the IEG Templates

18. Evaluation team members have reviewed all 27 PforR operations approved by June 30, 2015, in accordance with a standard template. This has primarily been a desk review, except for nine PforRs where the templates were revised following

field missions. Table D.1 shows the distribution of the ratings concerning results frameworks and DLIs. Important findings:

- The ratings have the same distributions for the quality of results frameworks and quality of DLIs (with verification protocols). This is not surprising, given the strong overlaps between the two.
- Overall, reviewers found these two sets to be largely satisfactory – with 78 percent of the PforRs rated satisfactory or moderately satisfactory for these core dimensions.
- The results frameworks and DLIs are supported by the PAPs, which were rated moderately unsatisfactory in only 2 of 26 cases (8 percent).
- The DLI risk ratings were also found generally satisfactory – with only 4 of 25 (16 percent) PforRs rated moderately unsatisfactory.

Table D.1. DLI Ratings

DLI ratings by IEG reviewers	HS	S	MS	MU	U	Total
2.4 Quality of the results framework and M&E arrangements, including capacity		9	12	5	1	27
2.5 Quality of the DLIs and verification protocols		9	12	5	1	27
2.6 Adequacy of Program Action Plan and conditionalities	1	10	13	2		26
5.5 Appropriateness of DLI risk	1	14	6	4		25
6.4 Quality of implementation of Program Action Plans		1	9	3	1	14

Note: The totals differ since reviewers did not fill out the ratings in all cases (or qualified as NA)—in particular for rating 6.4, due to lack of information about implementation of the PAPs (such as that this implementation had not yet started).

19. Results frameworks and/or DLIs rated either moderately unsatisfactory or unsatisfactory were typically found to have some of the following weaknesses:

- PDO not fully representative for program objectives.
- PDO-level indicators do not cover expected real results well (results on the ground).
- Lack of baseline data and monitoring indicators.
- Weaknesses in verification protocols.
- Weak attention to important gender or to social or environmental dimensions of results.
- Lack of economic indicators of results, especially noted as a weakness for operations with strong economic evaluation models.
- DLIs not covering (or covering late) important parts of results frameworks, so that disbursements may largely take place even with little or no progress for these aspects.

DLI Classifications

20. DLIs can be classified along several dimensions, as described in this section.

21. The average size of DLIs varies. The average PforR operation has 8.1 DLIs, and ranges from a high of 12 (4 programs) to a low of 3 (1 program). The average size of planned aggregate disbursements for each DLI also varies considerably (influenced, in part, by the size of the programs). The highest average size (\$66.7 million) is for the \$200 million Vietnam RWSS Program, with only three DLIs, while the lowest (\$3.1 million) is for a small program (\$30.8 million total) in Moldova. There is no particular relationship between commitment amounts and number of DLIs.³

22. Table D.2 shows that the DLIs are divided almost equally between capacity building and institutional development (CBID) and others (that include results on the ground and fiduciary DLIs), but with some modest differences by amounts versus by number, since, on average, the CBID DLIs have somewhat smaller financing attached. Thus, for the sample, CBID DLIs account for 48 percent by amount and 52 percent by number. The distribution varies considerably among operations, from a low of zero CBID in one program (Pakistan Punjab Government) to 100 percent at the other extreme (Ethiopia Local Government), and with 17 of the 27 operations below the average of 48 percent, with the 5 largest percentages for programs in Africa.

Table D.2. Distribution of Capacity-Building and Institutional Development DLIs by Number and Amount

Nature of DLI	Number of DLIs	Percentage of total number	Financing amount (US\$, millions)	Percentage of total financing
Capacity building / institutional development	116	53	2,430.0	48
Other	103	47	2,651.9	52
Total	219	100	5,081.9	100

23. In some cases, high percentages for institutional-type indicators are to be expected. Thus, in the Rwanda PforR for Public Sector Governance, the PDO is itself institutional – to enhance the country’s public financial management and statistics systems to improve transparency and accountability in the use of public funds, revenue mobilization, and the quality and accessibility of development data for decision making. In this case, the one non-CBID DLI is measuring the extent to which sectors are using a specific accounting and financial reporting system. In other cases – such as the example from Tanzania in box D.1 – a relatively high

percentage of CBID DLIs may mean that the DLI sets in question are more focused on intermediate institutional steps, and thus are less representative of the overall program objectives. It may also mean that the achievement of the DLIs in such cases may or may not mean that program objectives are being met.

24. Table D.3 shows the distribution among the four categories of intermediate outcomes, outputs, actions, and process. There is thus not even a separate classification for full outcomes. This absence is understandable, given the limited period for the PforR operations, but indicates that the name of this instrument should not be understood to mean program for outcomes.

Table D.3. Distribution of DLIs by Number and Commitment Amounts

Type of DLI	Number of DLIs	Percentage by number	Commitment amounts (US\$, millions)	Percentage by commitment amount	Average amount per DLI (US\$, millions)
Intermediate outcomes	90	41	2,305	45	25.6
Actions	60	27	1,129	22	18.8
Outputs	39	18	1,068	21	27.4
Processes	30	14	580	11	19.3
Total	219	100	5,082	100	23.2

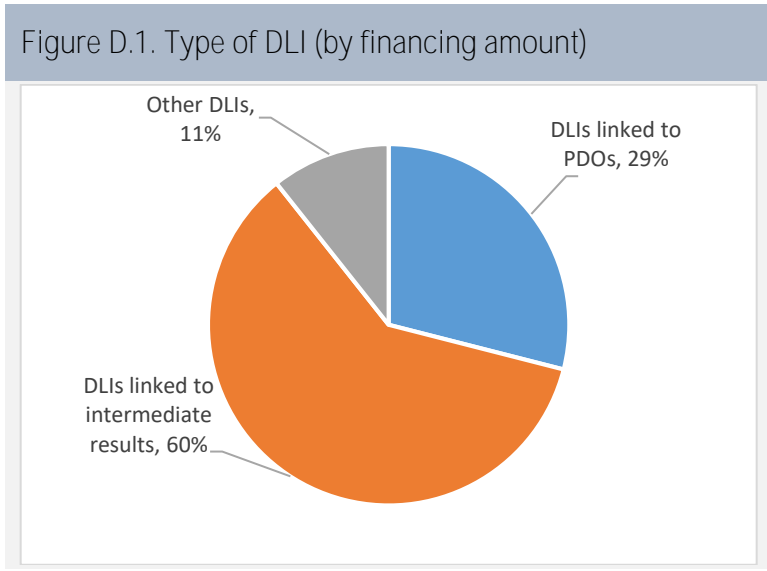
25. Most DLIs extend over several disbursement periods, typically with somewhat changing conditionalities over the period – for example, establishing minimum standards in the first year, completing upgrades of some facilities in the second year, completing upgrades of more facilities in the third year, and so on. The periods are normally one year each, but there are also cases with periods of six months, such as the Brazil PforR.

26. In some cases, including the India Enhancing Teacher Effectiveness in Bihar PforR, the nature of some DLIs can change significantly during implementation. For this program, an example is DLI4 – Accountability and Monitoring System. For periods 1–4, the indicators are for the development of such a system, while in the last period (period 5), the DLI shifts to average teacher performance score.

27. By financing amount, 60 percent of DLIs are linked to intermediate results in the results framework, 29 percent to PDOs, and 11 percent to other reasons, such as fiduciary. This is shown in figure D.1. By numbers, the linkages to intermediate results are even stronger – 64 percent, with 20 percent linked to PDOs and 16 percent for other reasons. This reinforces the impression that, overall, the DLIs are to support processes and institutional improvements toward the PDOs more than to

directly support the PDOs themselves.

28. There are only modest differences in the DLIs between International Bank for Reconstruction and Development (IBRD) and International Development Association (IDA) programs. There are eight programs in IBRD countries, for an average of \$221 million, and 19 programs in IDA countries, for a somewhat smaller average of \$172 million. The largest program in each category is for \$500 million – for Egypt and Nigeria, respectively.



29. A final consideration is the degree of scalable DLIs. Table D.4 shows a total of 153 such indicators – 70 percent of all DLIs in the IEG sample. However, in at least one case – Rwanda Agriculture – there were a few institutional indicators that were erroneously (in IEG’s view) treated as scalable.⁴ There also seems to be an uneven practice when it comes to the “bottom line” for scalability – in the same Rwanda case, any improvement in the performance of the indicators will cause at least some disbursements to take place, but in other cases, there will be a floor for eligible performance. (And in this Rwanda case, achievement of 75 percent of stated targets is sufficient for full disbursement.)

Table D.4. Distribution of Scalable DLIs

Type of DLI	Number of DLIs	Scalable DLIs	Percentage	Average amount per scalable DLI (US\$, millions)	Average amount per DLI (US\$, millions)
DLIs linked to PDOs	44	37	84	36.0	33.5
DLIs linked to intermediate results	141	93	66	25.3	21.8
Other DLIs	34	23	68	15.2	15.9
Grand total	219	153	70	26.4	23.2

Sources: OPCS DLI Database as of March 31, 2015. The five additional operations, approved after March 31, were coded by the IEG evaluation team.

Note: The Tanzania Primary Health Operation DLI#4 is linked to both PDOs and intermediate results, based on the components of the LGA (Local Government Agency) Balance Score Card, as shown in the appendix. However, for **statistical simplification, this DLI was treated as “linked to PDOs” when calculated** for this table. Scalability: several DLIs were coded differently for different years in the DLI verification protocol tables. For statistical simplification, changes are made as follows: (i) Egypt Housing Project DLI#4: Scalability = yes (verification protocol table lists no for year 1 but yes for years 2–5); (ii) India Teacher Effectiveness Project: DLI#1: Scalability = yes (verification protocol table lists no for year 0, but yes for years 1–5), DLI#3: Scalability = no (verification protocol table lists no for years 0–3, part (ii) in years 4 and 5, yes for part (i) in years 4 and 5); and DLI#5: Scalability = no (verification protocol table lists no for years 2 and 3, yes for year 5).

DLIs Outside PforRs

30. As discussed elsewhere, DLIs are also being used increasingly outside PforRs—having performance-linked disbursements without the use of country systems. Thus, many sectorwide approach (SWAP) operations have begun using DLIs in recent years. A review of 82 DLIs for 7 SWAP operations showed that 30 percent of the DLIs were action or process indicators, 32 percent were output indicators, and 29 percent were intermediate outcome indicators. That review noted that the large share of action or process indicators points toward linking disbursements in efforts to improve systems and institutions that support programs. IEG observations indicated that not using country systems can have the advantage of putting less pressure on government entities that may already have heavy workloads in relation to their human and financial resources, and also to help to insulate the operations better—for example, from problems of government budgeting. But it was also indicated that the verifications may be addressed less rigorously when outside PforRs. There has not yet been any formal guidance for staff using results-based indicators or DLIs for investment project financing (IPF), but IEG understands that such guidance is now under preparation.

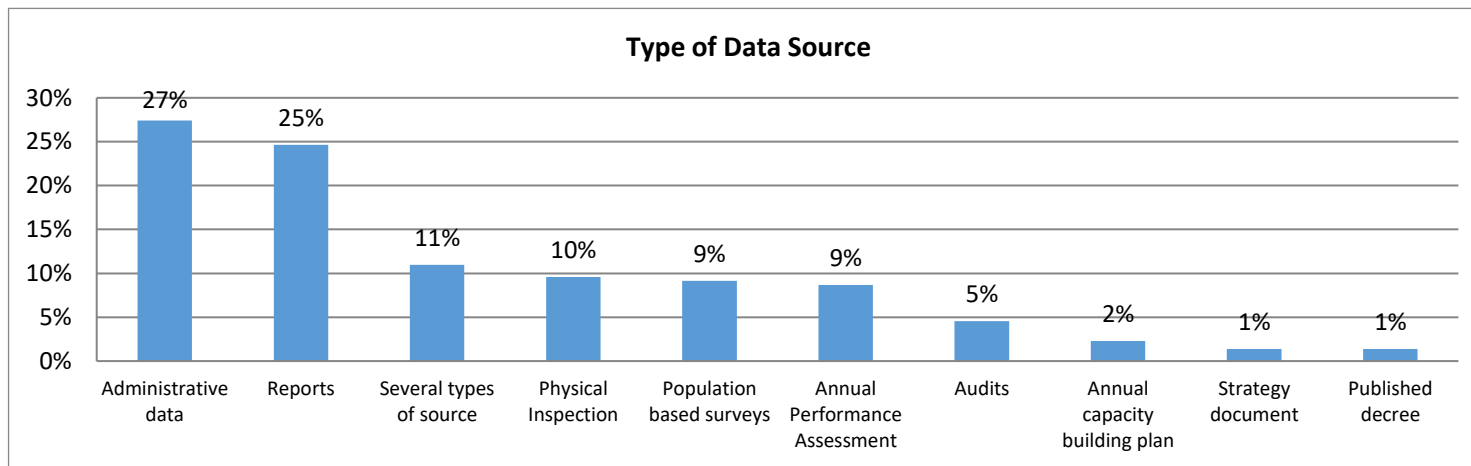
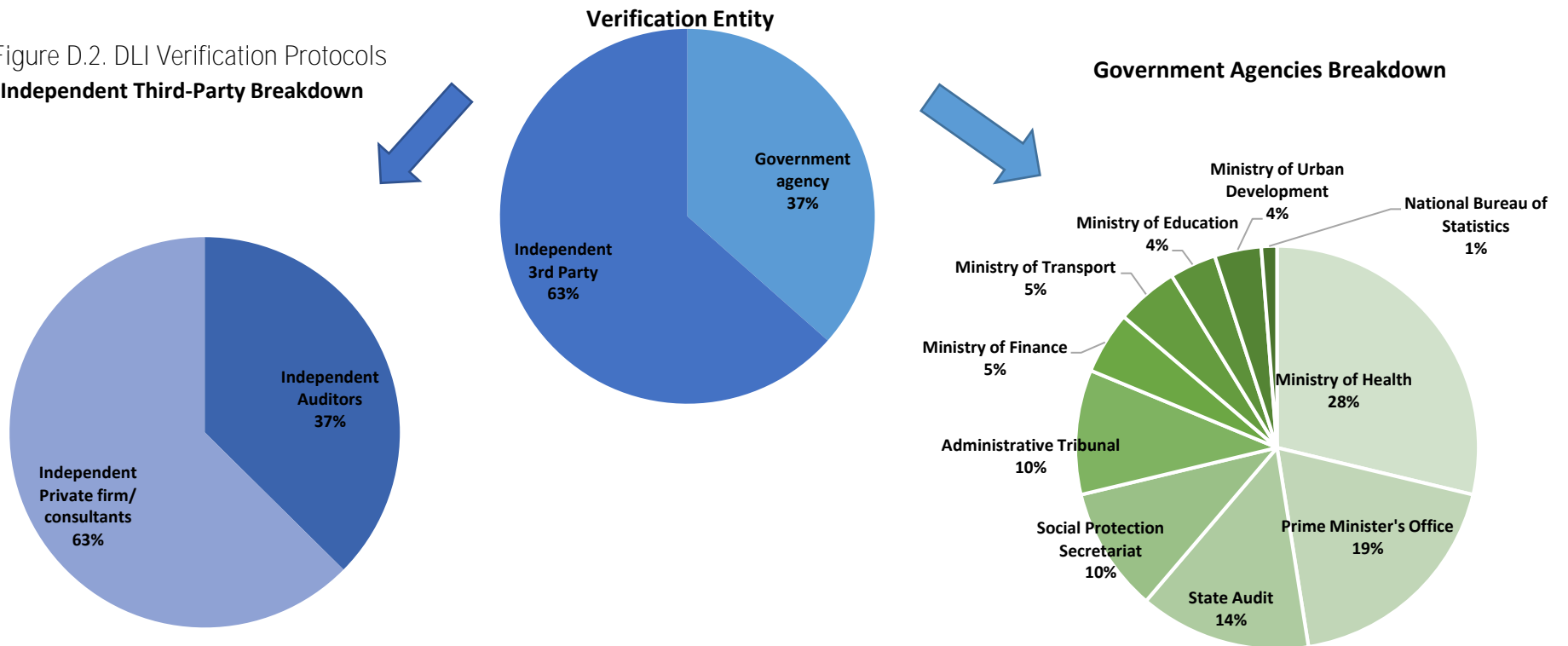
Verification Protocols

31. All PforRs are required to have a credible verification protocol for the DLIs that will define them and set out how they will be measured, whether they are

APPENDIX D
DISBURSEMENT-LINKED INDICATORS

discrete (all or nothing) or scalable, how the verification will be done, and who will be responsible for verifying their achievement. All programs in the IEG sample have detailed verification protocols, and the programs examined through field visits have shown that the verifications (where applicable from program progress) are taken seriously by all parties, although the sample of such programs is necessarily small, so it is premature to try and pass judgment on how the process works in practice and over time (see figure D.2).

Figure D.2. DLI Verification Protocols
Independent Third-Party Breakdown



32. Verification can be carried out in a variety of ways and by various parties:
- Figure D.2 shows that 63 percent of all DLIs by number are to be verified by independent parties, and 37 percent by government agencies.
 - For the independent verifications, 63 percent of DLIs are to be verified by private third parties and 37 percent by independent auditors (including auditor-generals).
 - For the government agencies, there are a variety of entities. Thus, for the two Rwanda PforRs, verification is done in one instance by the Auditor-General's Office and in the other by the Prime Minister's Office. The common thread is that while formal independence may vary, the key aspect – which should always be verified during program implementation – is that these entities will have sufficient competence and behavioral independence in relation to the executing agencies.
33. In a few cases, the verification protocols also list the Bank as the verification entity. Thus, in the case of the Tanzania Urban Local Government Strengthening Program, the Prime Minister's Office, Regional Administration and Local Government, is the "normal" verification entity for information coming in from other sources. However, for three DLIs, this office will itself be the data source, and the PAD verification protocol then lists the Bank as the verification entity. Given the Bank's general oversight role (quite extensive in some cases), it is not to be recommended that it also to be responsible for the verification of specific DLIs.
34. Verification processes are often – and rightly – demanding (complex and lengthy) and time-consuming, which is warranted on program grounds. Thus, as one example, the verifications for the Vietnam RWSS Program have so far taken nine months each year – between the end of the results period and the disbursement of funds to the implementing provinces – although the mid-term report expresses hopes that this can be reduced to six months. Such long processes can create issues of liquidity for participating entities. More generally, it is conceivable (but premature to assess) that such lengthy processes may have a limited exemplary effect – that they may not be maintained by the governments or expanded into other areas without Bank engagement.

DLIs During Implementation

35. So far, very few DLIs have been modified, indicating care in their initial formulation. The Board paper (World Bank 2011a) stated that "if DLIs were not well defined during the preparation of a Program-for-Results operation, or cannot be met because of unexpected events, the definition of these DLIs may be modified during

the operation's implementation, following normal Bank project restructuring processes" (p. 28).

36. One case of proposed mid-term modifications: The aide-memoire from the recent mid-term review of the Vietnam Rural Water Supply and Sanitation Program included proposals for modifying the DLIs. These included de-linking water supply and sanitation sub-DLIs (this would provide more flexibility but could also lead to a return of the traditional pattern of increased attention to water supply rather than sanitation), extending program implementation by one year, reducing the target for DLI1.1 (water supply connections), and introducing a new DLI on the sustainability of sanitation in schools and clinics.

Lessons

- The PforR objectives are in the results frameworks, not in the DLIs. A comprehensive assessment of PforR results must cover several dimensions, with the results framework as the foundation. The DLIs, which naturally will get much attention during implementation, should not be considered the most important aspects of program results.
- The DLIs must address a number of considerations. In their DLI designs, task teams may need to weigh various considerations, including the balance between relative ease of timely disbursements and the degree of ambition (stretch) in the DLI targets. IEG has seen several cases of DLIs perhaps being unduly easy, either individually or in the aggregate.
- DLI sets should be clearly linked to, and preferably reasonably representative of, the results frameworks, so that achievement of DLIs would provide strong assurance that the PDO/PDO indicators will also be achieved. Task teams should be cautious of DLI sets that concentrate heavily on institutional matters if this means that there will be only modest attention to the planned results on the ground.
- It should be recognized explicitly that many intermediate indicators and DLIs are likely to be for supporting processes and institutional improvements toward the PDOs more than in direct support of the PDOs themselves. This also means that few outcome indicators are likely to be for final outcomes.
- Preferably, the Bank should not be the responsible verification entity for specific DLIs, given the Bank's general oversight role.
- The Bank could consider, where possible, steps to simplify and shorten the verification processes where this could be done without reducing the quality of the processes. This could help reduce participants' liquidity issues and could also work to strengthen the general attractiveness of these processes in the countries.

¹ Two interim guidance notes were issued on June 18, 2012, one on technical assessment, the other on DLIs and disbursement arrangements. In this paper, the two notes are referred to collectively as the Guidance Notes.

² In this background note, all references to PforR programs refer to the programs approved through June 30, 2015, unless otherwise stated.

³ In a few cases, the commitment amounts presented include not only the planned Bank commitments, but also some other amounts, in one case from a Bank-managed trust fund, and in another case an amount from the borrower itself. There thus does not seem to be fully consistent practice among PforRs about how to record the DLI amounts.

⁴ For that program, the PAD annex table 3.3 treats all DLIs as scalable under the formulation "Payments will be made in proportion to the achievements (an agreed minimum value of at least 75 percent of the agreed target value to obtain the 100 percent disbursement target," even for institutional DLIs such as DLI 6 (Updated Gender Sensitive MIS Framework and Action Plan) and DLI 7 (approval of various policies and begin implementation of action plans), for which scalability is meaningless.

Appendix E. Fiduciary Assessments

1. The Bank has been committed to increased use of country systems for some time (World Bank 2003). Country systems include, for instance, designation of the project as on-budget, aligning with the fiscal calendar, and relying on the range of national budget preparation and execution procedures, as well as intergovernmental transfers, audit, and national competitive bidding procedures for procurement. Country systems for safeguards, results frameworks, and monitoring may also be used.
2. PforRs are part of this long-term attempt to design operations that are a better fit for country contexts. The PforR approach to country systems is different from the Bank's earlier country systems pilots. First, it uses the systems that are responsible for the program being supported, which may be at a higher standard than those for other programs. Second, it does not insist that these program systems be at the same standard as the Bank's policies and procedures for investment lending operations.

Integrated Fiduciary Systems Assessments

3. Each PforR produced an Integrated Fiduciary Systems Assessment (IFSA), typically summarized in an appendix of the program document, and covering the following issues:
 - Adequacy of program procurement system
 - Adequacy of program financial management system
 - Adequacy of program-specific system to handle issues relating to fraud and corruption, based on the Bank's Anti-Corruption Guidelines (ACG)
 - Adequacy of stakeholder involvement in program implementation and oversight, including verification of disbursement-linked indicators (DLIs) and establishment of a complaint and grievance mechanism
 - Level of transparency in program decision-making processes and performance reporting
 - Appropriateness of exclusions.
4. Overall, these assessments have been comprehensive, and cover most of the key aspects that one would expect to find. Each area of discussion typically begins with an overview of the laws and processes applied in each area of the government's program, the shortcomings in practice, and how the shortcomings will be addressed and monitored. Most PforRs have also drawn from extensive previous analytical work on fiduciary systems, including Public Expenditure and Financial Accountability (PEFA) assessments, Country Procurement Assessment Reports

(CPARs), and a range of other fiduciary assessments done by development partners in connection with budget-support operations. The IFSAAs are generally thorough in identifying relevant transparency measures, such as timely provision of information to stakeholders, disclosure of tender notices and award decisions, and parliamentary oversight. But some contextual or political economy issues may not be addressed, as discussed in this appendix.

Costings

5. A shortcoming is the lack of details on the costing methodologies used for the government programs being supported. Most PforRs seem to base their costing on the costing of the underlying government program, but with few or no details on the costing methodology used. An exception and example of good practice is the Ethiopia Health Millennium Development Goals PforR, where costing was done using Marginal Budgeting for Bottlenecks (see box E.1).

Box E.1. Good Practice in Costing: Ethiopia Health Millennium Development Goals PforR

Costing was done using Marginal Budgeting for Bottlenecks, a tool created by development partners and the national ministries of health from several countries. The tool helps to remove health system bottlenecks by helping to prepare strategic plans and expenditure programs to increase the quality of high-impact health, nutrition, HIV/AIDS, tuberculosis, and malaria interventions. The approach focuses on three service-delivery modes: family-oriented, community-based services; population-oriented schedulable services; and individual-oriented clinical services. It helps to assess current performance of health services and to identify bottlenecks in both supply and demand, and enables a tailored approach to each country's specific situation. This helps to facilitate selection of the types, quantities, and costs of salaries, drugs, training, and other inputs that are needed to overcome bottlenecks to achieve optimal results.

Sources: World Bank 2013b; UNICEF, World Bank, and African Development Bank 2011.

Procurement

6. Most program procurement systems are assessed as consistent with good public procurement principles, although implementation is often uneven due to lack of political commitment, rigorous follow-up, trained staff, and effective demand-side processes. Challenges are sometimes pointed out without corresponding proposed actions, such as possible entry barriers for contractors created by their mandatory registration.

7. The fiduciary systems used for PforRs are broadly the same as those for other areas of the government programs. One exception is the PforR for Brazil: Strengthen

Service Delivery for Growth, Poverty Reduction and Environmental Sustainability in the State of Ceará, where technical assistance procurement (10 percent of Bank-funded project costs) uses Bank rules because local rules were not thought to give enough attention to quality. However, IEG's field mission found that the technical assistance component had become unduly complex and had turned into a source of delays for the implementation of the PforR. In hindsight, it is reasonable to assume that it might have been more efficient for the Bank to have provided some initial capacity building and training for the government's own procurement staff in participating agencies to implement its own "technical quality + price" methodology, which, although rarely used, had been legally available from the start.

8. Another exception is the high-value procurement exclusion, which has reportedly excluded financing in certain areas supported by the PforR in some countries. PforR financing cannot normally be used for procurement packages of high-value contracts above certain thresholds, based on the type of procurement and the level of fiduciary risk. If such procurements are required in programs supported by a PforR, they may be financed from another source. Following the *Two-Year Review* (World Bank 2015e), a provision was added that such contracts may be financed if they are important to the integrity of the program and their cost is less than 25 percent of the cost of the overall program. Such an exception needs to be approved by the Bank's managing director and chief operations officer.

9. The Bangladesh Revenue Mobilization Program-for-Results: Value Added Tax (VAT) Improvement Program included the financing for a large information technology system. The hardware and software components were initially bid as separate procurements, but the Cabinet Committee on Government Purchase decided not to accept the successful bidders from the initial procurement as recommended by the Technical Committee, to cancel the initial tender, and to retender, merging the two components. The resulting successful tender was for \$29 million, above the \$20 million threshold allowed under the program. As a result, and following lengthy discussions, the full component (hardware and software) will now be financed by the government, while the PforR will support the implementation of the system in other ways. In this case, the IPF instrument might have been a better choice, at least for this component, given the challenges in Bangladesh with large, complex procurements.

10. In a sense, any PforR that requires additional measures to mitigate fiduciary risk is creating potential differences between the PforR system and the relevant overall program and country systems. In some cases, these measures may be applied not only to the PforR, but to the larger program as well, although this often is not clear from the documentation. Thus, under Uruguay's PforR, the Program Action

Plan (PAP) calls for identifying bidding bottlenecks and reducing bid duration by improving agency efficiency of award decision making. In addition, the PforR will establish a database in the implementing agency to register, track, and follow up on complaints linked to program-related activities, including procurement complaints. Generally, if high-value procurements take place outside the PforR, the PforR fiduciary risk is transformed into a funding risk for the portion of the program not funded by the Bank, which could, in turn, become a risk to the development objective supported under the program. In addition, the high-value procurement exclusion diminishes the ability of the Bank to strengthen some of the more difficult components of the country fiduciary systems.

Anti-Corruption

11. Under PforR financing, the borrower is responsible for taking actions in this area. The Bank also can investigate allegations, and sanction parties if appropriate. Concerns about the Bank's ACG have reportedly been responsible for reducing the size of some PforR loans. In Brazil, to avoid inconsistency at the local-government level, some activities were taken out of the scope of the PforR operation. In Morocco, government counterparts interviewed by the IEG mission repeatedly raised the Bank's ACG as an issue. They questioned why the Bank agreed to the use of all other aspects of national systems, but insisted on the application of its ACG without even assessing the quality of the country's own anti-corruption system. The Integrity Vice Presidency to date has received only five complaints alleging corruption in 3 of the 27 PforR operations assessed by IEG. The complaints were judged to not warrant investigation, and no investigations were carried out.

12. Following the *Two-Year Review* (World Bank 2015e), these rights and responsibilities did not change. However, there is an added provision that the Bank and borrower will formulate a program-specific protocol on how to take actions. The protocol may be a legal agreement, a memorandum of understanding, or part of the negotiation minutes. In addition, new language is added emphasizing that the country has the sovereign right to take actions in this area, and that Bank actions are solely for the purpose of determining compliance with Bank policies. If any action taken by the borrower conflicts with the laws and regulations of the country, the Bank and borrower will consult to agree on alternative actions that will avoid such a conflict.

Political Economy Aspects

13. Political economy issues are behind many of the constraints raised in fiduciary assessments, but they are rarely discussed directly. Any critical analysis tends to address only low-level issues, such as in one case: "The membership of

district council internal audit committee cannot be easily controlled since they are elected politicians. This can be addressed by co-opting competent members based on an agreed sitting fee.” The IEG evaluation found several examples through simple literature searches where outside parties, such as researchers, had raised critical views about the actual state of procurement and other fiduciary practices that did not seem to be reflected in the IFsAs, where the focus is instead typically on institutional reforms to address corruption, including program-specific measures. There are seldom any convincing arguments why such measures will work in the case of the new programs. Several interviews with team members confirm their awareness of such information and issues; the lack of published analysis is reportedly because of the sensitivity of this area for borrower governments.

14. An example of good practice is the Mozambique PFM PforR. The project appraisal document (PAD) includes a good summary of the political economy challenges, the identification of winners and losers, and design features that address these challenges: strong, senior-level ownership across key ministries, a shift in incentives that creates greater benefits to cooperation, flexibility to account for unexpected impacts, the use of change agents at the deconcentrated levels to drive change, and sharing of risk between the Bank and the client. The design also addresses the complex political economy of medicine supply chain reform, including interdepartmental coordination and competing policy aims.

15. The Systematic Operations Risk-Rating Tool, introduced for PforR operations in January 2016, includes political and governance risk as one of the nine risk categories. Another category measures the related aspect of institutional capacity risk. IEG supports the use of this new tool to address the political economy challenges raised in this report.

Risk and Mitigation Measures

16. Nearly all PforRs mention fiduciary capacity constraints, often with mitigating measures enforced by DLIs, PAPs/legal covenants, and/or during supervision missions. In most cases, these actions are not conditions precedent to disbursements or PforR effectiveness, though some may be linked to a portion of the disbursements through the DLIs as a performance incentive. The timing of PAPs/legal covenants for the PforRs assessed seems reasonable.

17. One example is Rwanda, where a key noted constraint is the high turnover of procurement officers. The Bank’s earlier Public Sector Capacity Building Project had worked to address this issue, which is pervasive across professional positions in the public sector. But despite extensive support provided by that operation, and related capacity-building operations funded by other donors, the problem has persisted.

The PAD correctly points out that better incentives for attracting, promoting, and retaining staff are needed to reduce high (by regional standards) vacancy and turnover rates. This issue was apparently awkward to address through the PforR, but the PFM basket-fund donors are addressing it through other support.

Findings and Conclusions

18. PforRs are part of a long-term attempt by the Bank to move toward operations that are a better fit for country contexts. The instrument is designed to support the implementation of government programs, using the government's own fiduciary, environmental, and social safeguard systems. The intention is that PforRs, working in tandem with other Bank instruments, will help to deliver the Bank priorities around results, scale up in poorer countries, and strengthen country systems.

19. Overall, the fiduciary assessments in PforRs, along with the mitigating measures carried out, generally provide reasonable assurance that program funds will be used as intended, including funds from sources other than the PforR. A shortcoming is the lack of details on the costing methodologies used for the government programs being supported. Most PforRs seem to base their costing on the costing of the underlying government program, with few or no details on the costing methodology used.

20. The fiduciary systems used for PforRs are broadly the same as for other areas of the government programs, with two main exceptions. The first is the high-value procurement exclusion, which normally prevents the use of PforR financing for contracts above certain thresholds, based on the type of procurement and the level of fiduciary risk. This exclusion has prevented the use of PforR financing for areas important to program integrity. This exclusion has not been a major obstacle in the PforRs reviewed, because these areas can be financed from other sources. In any case, the Bank introduced a policy change in 2015 allowing for exceptions to the procurement threshold in some situations.

21. The second exception is the perceived scope of the Bank's ACG, which were viewed by some borrowers as limiting the scope of PforRs. The Bank made some changes in 2015 to the PforR ACG to address these concerns, but it is too early to determine the result.

22. Political economy issues are behind many of the constraints raised in fiduciary assessments, but are rarely discussed directly. However, interviews with team members confirm their awareness of such issues; the lack of published analysis is reportedly because of the sensitivity of this area for borrower governments.

23. Finally, IEG finds some evidence that capacities strengthened by PforRs will be sustained – for example, through support from a new results-based project in Mozambique’s education sector, and through reforms supporting the medium-term expenditure framework in Kenya.

Appendix F. Environmental and Social Systems

Background and Context

1. The Bank's environmental and social safeguards for investment project financing (IPF) have been undergoing a multiyear review and reform process. In anticipation of that process, IEG prepared an evaluation report (IEG 2010), which found that the Bank's safeguards policies had helped to avoid or mitigate large-scale social and environmental risks associated with the Bank's IPFs. However, the implementation of these policies had meant enforcing compliance with mandatory policies and procedures, without engendering strong client ownership. The report pointed to the need for a systems approach—linking policy regulations to project design, supervision, monitoring, evaluation, and disclosure—and partnership with clients, third parties, and local communities to enhance ownership of results, integrating elements of the Bank's safeguards with some of the practices under the International Finance Corporation/Multilateral Investment Guarantee Agency Performance Standards. On this basis, IEG recommended the revision of the policy frameworks to harmonize thematic coverage and guidance across the Bank Group and enhance the relevance of those frameworks to client needs.

2. A 2015 report (IEG 2015c) found that the application of the elements of the Bank's Operational Policy on Development Policy Financing (OP 8.60) related to environmental and social risks had been inconsistent. In addition, the Bank's assessments of the clients' capacity to manage these risks had often been perfunctory. Finally, there was no formal system in place in the Bank to monitor and evaluate environmental and social risks and their mitigation in development policy financing (DPF), and the Bank's monitoring documents seldom contained information on actual environmental or social effects or the efficacy of mitigation measures. In light of these findings, the report suggested that Bank Procedure (BP) 8.60 be revised to specify procedures to ensure that the provisions of OP 8.60 with respect to environmental and social risks associated with DPF are adequately applied. These provisions relate to the screening and identification of environmental and social risks; guidance to task teams; incentives and training for Bank staff; and the strengthening of monitoring, reporting, and evaluation of environmental and social effects in DPF documents. The Bank has accepted these suggestions and recommended their implementation (World Bank 2015c).

Environmental and Social Provisions for Program-for-Results Financing

3. As stated in the fiscal 2012 Board document (World Bank 2011a), PforRs will seek to make sure that the environmental and social effects of the program are adequately addressed. During the preparation of a PforR operation, the Bank will assess, against the requirements of Operational Policy/Bank Procedure (OP/BP) 9.00, the degree to which a program's systems manage and mitigate the environmental and social impacts of the overall program. The assessment will also identify and exclude high-risk activities – that is, those that pose a risk of potentially significant adverse impact on the environment or affected people (activities classified as Category A under the IPF safeguards).

4. In line with the above, OP 9.00 – Program-for-Results Financing, spells out the requirements for the Bank teams' assessments of each program's environmental and social systems (World Bank 2015b, p. 69ff). Specifically, an assessment is expected to consider to what degree the program's systems:

- Promote environmental and social sustainability by avoiding, minimizing, or mitigating adverse impacts and promoting informed decision making.
- Avoid, minimize, or mitigate adverse impacts on natural habitats and physical cultural resources.
- Protect public and worker safety against potential risks associated with construction activities, exposure to hazardous materials, and natural hazards.
- Manage land acquisition and loss of access to natural resources in a way that avoids or minimizes displacement, and assist the affected people in improving, or at the minimum restoring, their livelihoods and living standards.
- Give due consideration to the cultural appropriateness of, and equitable access to, program benefits, giving special attention to the rights and interests of the indigenous peoples and to the needs or concerns of vulnerable groups.
- Avoid exacerbating social conflict, especially in fragile states, post-conflict areas, or areas subject to territorial disputes.

5. Unlike the practice for IPFs, the Bank itself is expected to undertake the environmental and social systems assessment (ESSA) for a PforR, and the ESSA is required for every operation. The ESSA is expected to identify actions, as needed, to enhance the systems during program preparation and implementation (the latter to be included in the Program Action Plan [PAP]). The Bank is also expected to consult with program stakeholders and disclose the results and recommendations of its ESSA before the appraisal of the PforR.

6. In addition, the environmental and social risks and proposed mitigation measures are expected to be inputs to the integrated risk assessment of the PforR. During implementation, the Bank is expected to pay particular attention to reviewing the borrower's compliance with its contractual undertakings in the environmental and social management areas, including those related to the PAP; monitor the evolution of risks; and make adjustments as appropriate.

Findings from the *Two-Year Review*

7. Management's *Two-Year Review* of the PforR instrument found that the implementation of environmental and social dimensions was "working as intended" (World Bank 2015e). The review found that PforR activities had generally presented low-to-moderate environmental or social risk. In the two cases where risks were rated as substantial, adequate risk management and mitigation measures had been put in place. For all projects, environmental and social risk management measures had been reflected in 10 percent of the disbursement-linked indicators (DLIs) and 16 percent of all PAP actions, and the screening process had effectively excluded high-risk activities. However, the review also found that the exclusion of potentially significant environmental and social impacts was one of the most important factors accounting for differences between the scope of PforR programs in relation to the government programs supported by the PforRs (World Bank 2015e, p. 15). Finally, the review identified two areas where improvements were warranted. First, the ESSAs were uneven in their depth of assessment, focus, and clarity. They could do more to prioritize potential issues according to their relevance, devise appropriate actions to address priority issues, and emphasize measures to strengthen the borrower's overall environmental and social risk management systems. Second, more integration at the team level could help ensure that the ESSA is part of the overall focus on capacity building and an integral part of the broader dialogue. As will be discussed below, IEG broadly agrees with these findings, and identifies additional issues that deserve greater attention for the future design and implementation of PforRs.

Identification, Assessment, and Mitigation of Environmental and Social Risks

8. **Quality of environmental and social assessments.** The desk review of the PforRs found that the Bank teams had prepared an ESSA for every operation, and that their quality had been satisfactory or moderately satisfactory in every case. The most frequent shortcomings related to the extent to which the capacity-building measures recommended in the ESSA were followed up in the PAPs and technical assistance provisions, the coverage of social issues, the adequacy of outreach and consultation with poor and vulnerable beneficiary groups, and the absence of monitoring

indicators and reporting on safeguards implementation in the results frameworks of the operation. Each of these areas will be discussed in greater detail below.

9. **Appropriateness of environmental and social risk rating.** Given the exclusion of high-risk components, as mandated by OP 9.00, the actual environmental and social risks faced by the sample PforRs ranged from low to substantial and have been satisfactorily rated in the project documents in most cases. In two cases, however, the ratings tended to be overcautious. For the Kenya National Safety Net Program, which has no direct environmental effects and only a very limited range of potentially adverse social effects, the environmental and social risk should have been appropriately rated as low rather than moderate. Similarly, in light of the fact that the Vietnam Northern Mountains Urban Development Project has excluded investments in environmentally sensitive areas, avoided investments in ethnic minority areas, and is addressing identified social risks through appropriate PAP actions, a risk rating of moderate would seem more appropriate than one of substantial.

10. **Appropriateness of measures for improving program systems for managing environmental and social effects.** The provisions for improving the programs' environmental and social systems were satisfactory in most of the review sample, and moderately satisfactory in two cases. Here again, the most frequent gaps related to limited or vaguely specified inclusions of ESSA-recommended capacity strengthening measures in the PAP and, in some cases, inadequate support for the PAP in the technical assistance provisions of the programs. For example, the ESSA of the Uruguay Road Rehabilitation and Maintenance Program identified a need for strengthened environmental management and supervision, and recommended seven actions to address this gap, but only two of these were included in the PAP. Similarly, the PAP of the Mexico Oaxaca Rural Water and Sanitation Program did not mention any of the actions the ESSA identified as required to address the substantial social risks of the program. It is possible in these and other cases that such actions have been covered through other means, such as an environmental and social management manual. However, the absence of their explicit mention offers little assurance that they will be adequately addressed and implemented.

11. **Quality of implementation of environmental and social systems.** Very little information on the quality of implementation of environmental and social systems can be gleaned from the projects' Implementation Status and Results Reports (ISRs).¹ In two of the sampled projects, the implementation of environmental and social systems is also directly tied to the achievement of DLIs, and their progress can be followed in the ISRs. Thus, for the Brazil Ceará Strengthening of Service Delivery Project, the ISR regularly reports on the progress of DLI#9 – linked to 9.9 percent of

total project financing – which is an index of environmental enforcement quality that combines indicators of (i) the enforcement of water pollution control regulations, (ii) the use of appropriate water-quality monitoring technology, and (iii) citizen participation in environmental monitoring.

12. The ISRs also appear to have very little information on the quality of environmental and social systems implementation. Thus, while all the ISRs that were reviewed include a section on the management of environmental and social risks, this section simply includes a description of the risk – synthesized from the risk assessment in the project appraisal document (PAD) – and an update on the environmental and social risk rating. A review of these environmental and social risk ratings in the latest ISRs indicates that they were the same as those in the PAD, except in five cases, where they have been raised. One of these cases was the Mozambique Public Financial Management for Results Program. This was initially rated as a low risk, but the rating was raised to moderate because of continued delays in the implementation of regulations for the disposal of expired medicines. Implementation delays were also the reason for the raising of the risk ratings in other cases.

13. In the course of its country visits, IEG found that the environmental and social systems were being diligently supervised by the Bank teams and implemented much as expected. No serious issues were foreseen, because the programs had excluded investments in environmentally sensitive areas and generally minimized land acquisition. In the few instances (Vietnam, Brazil) where the programs involved indigenous peoples, adequate provisions for their culturally sensitive participation had been addressed through the PAP and supported by technical assistance. Even so, the mission learned about important challenges faced in the adequate implementation of the Bank's policy requirements:

- One recurrent challenge relates to the **adequacy of the program systems**, which most ESSAs have described as adequate from a legal and policy perspective, but as stretched and uneven in terms of implementation capacity. IEG found that this situation continued for the programs visited. While technical assistance has been provided in every case to prepare environmental and social guidelines and train program staff to implement them, the missions found no indications that the resource and staffing shortages diagnosed in the ESSA had been addressed. Thus, the enforcement of environmental requirements and monitoring of environmental performance continued to be sporadic and uneven.
- Another recurrent challenge relates to the **adequacy of compensation for land acquisition**. Overall, IEG missions found that the programs that

required land acquisition had supported the preparation of land acquisition and resettlement guidelines and trained staff to implement them in conformity with Bank requirements. Verifying the adequacy of land compensation, however, remains a difficult challenge, as illustrated by the case of the Vietnam Northern Mountains Urban Development Program (box F.1).

Box F.1. Compensation for Land Acquisition—Vietnam Northern Mountains Urban Development Program

- In this program, the Program Action Plan supports “independent market-based assessments” of land prices.
- IEG was informed that such assessments had arrived at prices that were 2–5 percent higher than the established government guidelines, but that it was difficult to establish market prices because of the common practice of underreporting them in the official records in order to minimize land transfer taxes.
- A possible indication of the distortion created by this underreporting and under-assessment is that out of the 40+ households that needed to be relocated for the first investment project in a major city, all had accepted the “land compensation” option in an outlying resettlement area, and not a single household chose the “cash compensation” option that was based on the “independent market-based assessment.”

Source: World Bank 2014e.

14. **Appropriateness of stakeholder engagement and program transparency during implementation.** The ISRs in the PforR portfolio contain virtually no information on stakeholder engagement in the implementation of environmental and social management aspects in any of the ISRs in the PforR portfolio, even though in most cases, specific provisions and requirements for such engagements were made, or should have been expected, at an early stage of implementation. The apparent absence of any progress reports and monitoring information is a significant gap in the results reporting systems of these programs.

15. At the same time, IEG found that for the sample of visited projects, the stakeholder engagement associated with the programs’ social systems was being fully implemented and supervised by the Bank teams. Thus, for example, the Brazil Ceará Strengthening Service Delivery Program had established a publicly accessible website with all of the program documents, as well as the results framework, targets, and status.² In addition, the level and frequency of stakeholder engagement seemed to be appropriate for the type and implementation status of the programs. Thus, the skills development program involved consultations with private sector

organizations, and the family assistance program had involved consultations with local community leaders for the development of an outreach strategy for indigenous peoples and other vulnerable minorities.

Strengthening of National and Program Environmental and Social Systems

16. Adequacy of national systems for managing environmental and social effects of government programs. The ESSAs of the sample operations mostly discuss the countries' national systems for managing environmental and social effects only to the extent that they are applicable to the programs at hand. In each case, the countries' policies and legislation are described as comprehensive and in line with international practice, and the identified gaps relate to their uneven implementation, associated with local-level and program-specific capacity, resource, and incentive constraints. IEG's field visits found that the planned technical assistance and PAP actions had supported a number of improvements in the environmental and social systems of the supported programs. There were, however, few indications that such improvements were being adopted for broader use in the national systems.

17. Adequacy of the program systems for managing environmental and social effects. IEG's desk review found that the majority of the programs' systems appeared satisfactory. In the moderately satisfactory cases, systemic weaknesses were related to the adequacy of capacity and resources; deficiencies in the social management systems, such as with respect to transparency, accountability, and grievance redress mechanisms; and the disposal of medical and pharmaceutical waste. In these cases, specific measures to address the shortcomings identified were included in the technical assistance and PAPs.

18. IEG's field visits found that the planned technical assistance and PAP actions—primarily involving the preparation of specific guidelines and training for program staff—were being implemented, but had been unable to address the preexisting capacity and budgetary constraints that were limiting the effectiveness of the program systems. For example, the PAD of the Vietnam Rural Water Supply and Sanitation Project appropriately mentions the Ministry of Natural Resources and Environment—and, by implication, the provincial Departments of Natural Resources and Environment—as responsible for monitoring and evaluating the environmental aspects of program implementation. In the course of IEG's visit, the provincial departments confirmed that they were responsible for monitoring and enforcing environmental requirements, but had been unable to carry out these functions as systematically as expected due to continuing staff and budgetary shortages. Thus, while the PforR was appropriately using the program's existing system, there was no indication that it had helped address its capacity and resource limitations.

19. **Appropriateness of identification and exclusion of high-risk activities.** All of the sample PforRs have appropriately excluded high-risk environmental and social activities in line with the requirements of OP 9.00. In one case, however, the interpretation of this requirement appears to have been overly cautious, to the extent of having an impact on the achievement of the PDO: For the India Maharashtra Rural Water Supply and Sanitation Program, the exclusion of “schemes involving highly polluted water sources” appears to be at variance with the operation’s PDO “to improve access to quality and sustainable services in peri-urban villages and in water-stressed and water-quality affected areas” (World Bank 2014a, pp. vii, 6). In a few other cases, the exclusion of high-risk environmental and social activities has significantly reduced the scope of the PforRs in relation to the supported government programs, and the attendant challenges and opportunities associated with strengthening the programs’ systems. Thus, programs that were intended to strengthen local infrastructure development have excluded water treatment plants, new landfills, or road construction outside existing rights-of-way, as well as any investments with potential impacts on natural habitats or cultural areas. This raises the concern of how such investments will be handled if they are not subject to the oversight associated with Bank involvement.

Monitoring and Reporting of Environmental and Social Effects

20. The extent of monitoring of environmental and social effects of a project is difficult to establish, since reporting has been sparse. While most (78 percent) of the reviewed PforRs have been rated as facing moderate or substantial environmental and social risks, and two-thirds (63 percent) of the ESSAs point to the need for monitoring and reporting, only 10 (37 percent) PADs discuss specific provisions for the monitoring of environmental and social effects, and 5 (18 percent) of the ISRs provide any information on their implementation.

21. In three of the five ISRs that provide information on the implementation of environmental and social systems, the reporting is directly tied to their inclusion in a DLI. In other cases, information on the environmental and social effects is only available in the ISRs of two additional projects. The ISRs also appear to have very little monitoring information on the environmental and social effects of the programs. Thus, while all reviewed ISRs include a section on the management of environmental and social risks, this section simply includes a description of the risk – synthesized from the risk assessment in the PAD – and an update on the environmental and social risk rating, with no additional data.

22. In the course of its country visits, IEG found that the environmental and social systems were being implemented, but with virtually no reporting of environmental

and social effects, even at the program-management level. Thus, while the missions were adequately up to speed on the status and extent of land acquisition, involuntary resettlement, community consultations, extent of compliance with environmental permitting regulations, and so on, little of this information was – with some exceptions – reflected in the programs’ progress reports, and even less in the ISRs.

23. Overall, the evidence from the review of program documents and field visits indicates that environmental and social effects have not normally been seen as an integral part of PforR reporting. This finding is surprising since environmental and social risks have been rated as moderate or substantial in all but six (77 percent) of the sample portfolio and, as already noted, these risk ratings were raised in the course of supervision in five (18 percent) cases. Thus, while IEG found no evidence that environmental and social risks are not being managed, the very limited reporting about them that is available in the Bank’s systems is a major concern, given the essential role of monitoring for the environmental and social management framework and the important place of environmental and social risks within the PforRs’ integrated risk management framework.

Summary Findings

- Overall, IEG’s desk review indicates that the majority of the program systems were deemed to be satisfactory from a legal and policy perspective, but with frequent gaps in implementation.
- The available information suggests that the ESSAs have been reasonably effective in identifying and assessing environmental and social risks and in the identification of remedial actions.
- The most frequent shortcomings relate to the coverage of such actions in the PAPs and in technical assistance.
- Where systemic weaknesses were identified, the PforR’s provisions for improving the programs’ environmental and social systems appear to be adequate.
- The initial evidence about their mitigation has been encouraging, but there is a shortage of environmental and social monitoring indicators in the results frameworks and the ISRs.
- However, in the visited programs, the implementation of environmental and social systems was found to be adequately supervised and monitored by the Bank teams.
- Since the programs had been purposely defined to exclude subprojects in environmentally sensitive areas, and also to minimize land acquisition and impacts on indigenous peoples, IEG found no indications of significant adverse impacts so far.

- In a few instances involving indigenous peoples, adequate provisions for their culturally sensitive participation had been addressed through the PAPs and supported by technical assistance.
- Many of the ESSAs and PADs point to staff and budgetary constraints that have limited the effective implementation of the country program systems.
- In relation to the strengthening of national and program systems: While the remedial technical assistance and PAP actions are being implemented, there is no indication that the best practice models supported by these actions are being adopted for wider use beyond the scope of the PforRs.
- There is also no indication that the PforRs have had any effect on addressing the staff and resource constraints affecting the program and country systems.

Lessons

24. Given the above findings on the assessment and implementation of environmental and social systems in PforR operations, two lessons can be derived:

- **PforR operations are unlikely to have an impact on national environmental and social systems.** While IEG's country visits found that programs' environmental and social systems were being implemented largely as expected, it found no indication that the PforRs had yet been effective in addressing systemic weaknesses affecting the national systems, such as budget and staff constraints, or that the supported guidance and capacity building were being applied beyond the scope of the programs. This is not surprising, given the program focus of these operations and the pervasive budget and staff constraints on both the Bank and borrower sides. While it is possible that the improvements supported by the PforRs could have some demonstration effects in the future, IEG could not find indications for such broader impacts at this early stage of the programs' implementation.
- **Environmental and social performance is more likely to be monitored and reported if it is tied to DLIs.** As already noted, only five of the sample ISRs included information on the performance of the programs' environmental and social systems, and in three of these cases, the reporting was directly tied to their inclusion in a DLI. Thus, environmental and social performance information was only reported in two (8 percent) of the 24 projects that did not associate environmental and social performance with the DLIs, which is notable, given the important role of environmental and social performance to the Bank's overall performance, and the important role of ISRs for the Bank's overall reporting of operational results.

¹ One exception is the Uganda Support to Municipal Infrastructure Program, for which the ISRs have regularly reported on the average scores for the participating municipalities' performance with respect to the implementation of environmental measures and resettlement action plans.

² <http://www2.ipece.ce.gov.br/pforr/index.htm>.

Note: To find PADs for all the PforR projects, visit:

<http://www.worldbank.org/en/programs/program-for-results-financing#2>.

Bibliography

- ADB (Asian Development Bank). 2013. *Piloting Results-Based Lending for Programs*. Manila: ADB.
- — —. 2015. "Q&A on Results Based Lending." With Kiyoshi Nakamitsu. ADB, Manila.
<http://www.adb.org/site/public-sector-financing/results-based-lending-programs/q-a-results-based-lending>
- Birdsall, Nancy, and William Savedoff. 2011. *Cash on Delivery: A New Approach to Foreign Aid*. 2nd ed. Washington DC: Center for Global Development.
- Brenzel, Logan. 2009. "Taking Stock: World Bank Experience with Results-Based Financing (RBF) for Health." Technical Brief. Report No. 54034. World Bank, Washington DC.
- DFID (U.K. Department for International Development). 2014a. *DFID's Evaluation Framework for Payment by Results*. London: DFID.
- — —. 2014b. *Sharpening Incentives to Performance: DFID's Strategy for Payment by Results*. London: DFID.
- GPOBA (Global Partnership on Output-Based Aid). 2007. "GPOBA Commitment Paper: Vietnam Rural Water Supply Development Project." GPOBA, Washington, DC.
- Harrold, Peter, and Associates. 1995. "The Broad Sector Approach to Investment Lending: Sector Investment Programs." World Bank Discussion Paper No. 302. Washington, DC.
- Honorati, Maddalena, Laura Rawlings, and Julie Van Domelen. 2011. "Results-Based Lending Approaches in Social Protection and Labor." Social Protection & Labor Policy Note, October, No. 1. World Bank, Washington DC.
- IDB (Inter-American Development Bank). 2003. *Annual Report*. Washington, DC: IDB.
- — —. 2014a. *Lending Instruments Background Paper*. Office of Evaluation and Oversight. Washington, DC: IDB.
- — —. 2014b. *Performance-Driven Loans: Lessons Learned*. Office of Evaluation and Oversight. Washington, DC: IDB.
- IEG (Independent Evaluation Group). 2009. *Improving Effectiveness and Outcomes for the Poor in Health, Nutrition, and Population: An Evaluation of World Bank Support Since 1997*. Washington, DC: World Bank.
- — —. 2010. *Safeguards and Sustainability Policies in a Changing World: An Independent Evaluation of World Bank Group Experience*. Washington, DC: World Bank.
- — —. 2011. *Argentina Provincial Maternal and Child Health Adjustment Loan, Argentina Provincial Maternal and Child Health Investment APL1, and Brazil Family Health Extension Adaptable Lending Program*. Project Performance Assessment Report. Washington, DC: World Bank.
- — —. 2014. *World Bank Group Support to Health Financing*. Washington, DC: World Bank.
- — —. 2015a. *Approach Paper: Program-for-Results (PforR), IEG Process Evaluation*. Washington, DC: World Bank.
- — —. 2015b. *Learning and Results in World Bank Operations: Toward a New Learning Strategy*. Washington, DC: World Bank.
- — —. 2015c. *Managing Environmental and Social Risks in Development Policy Financing*. Washington, DC: World Bank.

BIBLIOGRAPHY

- — —. 2015d. *The Quality of Results Frameworks in Development Policy Operations*. Washington, DC: World Bank.
- — —. 2016. "Disbursement-Linked Indicators – Background Paper." World Bank, Washington, DC.
- McNee, Andrew. 2012. *Rethinking Health Sector Wide Approaches through the Lens of Aid Effectiveness*. Development Policy Centre Discussion Paper No. 14. Canberra: The Australian National University.
- Mumssen, Yogita, Lars Johannes, and Geeta Kumar. 2010. *Output-Based Aid: Lessons Learned and Best Practices*. Directions in Development: Finance. Report No. 53644. Washington DC: World Bank.
- Musgrove, Philip. 2011. *Financial and Other Rewards for Good Performance or Results: A Guided Tour of Concepts and Terms and a Short Glossary*. Washington DC: World Bank.
- O'Brien, Thomas, and Ravi Kanbur. 2013. "The Operational Dimensions of Results-Based Financing." *Public Administration and Development* 34(5): 345–58.
- OECD (Organization for Economic Co-operation and Development). 2012. *Aid Effectiveness in the Health Sector: Progress and Lessons*. Paris: OECD.
- OECD-DAC (OECD Development Assistance Committee). 2014. *International Development Statistics: Creditor Reporting System*. Paris: OECD.
- UNESCO (United Nations Educational, Scientific and Cultural Organization). 2015. *Education for All 2000–2015: Achievements and Challenges*. Paris: UNESCO.
- UNICEF, World Bank, and African Development Bank. 2011. *Marginal Budgeting for Bottlenecks: An Analytical Tool for Evidence-Based Health Policy, Planning, Costing, and Budgeting – Concise User Guide*. <http://www.devinfolive.info/mbb/mbbsupport>. Washington DC.
- Vaillancourt, Denise. 2009. "Do Health Sector-Wide Approaches Achieve Results? Emerging Evidence and Lessons from Six Countries: Bangladesh, Ghana, Kyrgyz Republic, Malawi, Nepal, and Tanzania." Independent Evaluation Group Working Paper 2009/4. IEG, Washington, DC.
- World Bank. 2002. *Private Sector Development Strategy – Directions for the World Bank Group*. Washington, DC: World Bank.
- — —. 2003. *Strategic Framework for IDA's Assistance to Africa (SFIA): The Emerging Partnership Model*. Washington, DC: World Bank.
- — —. 2006. "Guidelines on Preventing and Combating Fraud and Corruption in Projects Financed by IBRD Loans and IDA Credits and Grants." World Bank, Washington, DC. http://siteresources.worldbank.org/INTRUSSIANFEDERATION/Resources/ibRD_IDA_AnticorruptionGuidelines.pdf
- — —. 2007. "Implementing Output Based Disbursement Mechanisms for Investment Operations, Operational/Technical Guidance Note to Staff." Operations Policy and Country Services. World Bank, Washington, DC.
- — —. 2008. "Procurement Issues in Structuring OBA Operations Financed by the World Bank – Guidance Note for Staff." World Bank, Washington, DC.
- — —. 2009. *IDA15 Mid-term Review: A Review of the Use of Output-Based Aid Approaches*. Washington, DC: World Bank.

- — —. 2011a. *A New Instrument to Advance Development Effectiveness: Program-for-Results Financing*. Operations Policy and Country Services. Board Report. Report No. 66193. Washington, DC: World Bank.
 - — —. 2011b. *Romania: Social Assistance Modernization Project*. Project Appraisal Document. Report No. 58280. Washington, DC: World Bank.
 - — —. 2012a. "Guidelines on Preventing and Combating Fraud and Corruption in Program-for-Results Financing." World Bank, Washington, DC.
<http://siteresources.worldbank.org/INTLAWJUSTICE/Resources/ACS.pdf>
 - — —. 2012b. *Morocco – Second National Initiative for Human Development Program Project*. Project Appraisal Document. Report No. 67496. Washington, DC: World Bank.
 - — —. 2012c. *Program-for-Results Financing: Interim Guidance Notes to Staff on Assessments*. Operations Policy and Country Services. Washington, DC: World Bank.
 - — —. 2013a. "Bank Procedure: Policy and Procedure Framework." World Bank, Washington, DC.
 - — —. 2013b. *Ethiopia – Health Millennium Development Goals Program-for-Results Project*. Project Appraisal Document. Report No. 75101. Washington, DC: World Bank.
 - — —. 2013c. *Uganda – Support to Municipal Infrastructure Development Program Project*. Project Appraisal Document. Report No. 73174. Washington, DC: World Bank.
 - — —. 2014a. *India – Maharashtra Rural Water Supply and Sanitation Program Project*. Project Appraisal Document. Report No. 76172. Washington, DC: World Bank.
 - — —. 2014b. "India – Maharashtra Rural Water Supply and Sanitation Program Project, Fiduciary and Systems Assessment." World Bank Working Paper. Report No. 87000. World Bank, Washington, DC.
 - — —. 2014c. *Mexico – Oaxaca Water and Sanitation Sector Modernization Operation Project*. Project Appraisal Document. Report No. 87104. Washington, DC: World Bank.
 - — —. 2014d. *Tanzania – Big Results Now in Education Program Project*. Project Appraisal Document. Report No. 84545. Washington, DC: World Bank.
 - — —. 2014e. *Vietnam – Results-Based Northern Mountains Urban Program Project*. Project Appraisal Document. Report No. 87773. Washington, DC: World Bank.
 - — —. 2015a. "Bank Directive – Program-for-Results Financing." Washington, DC.
 - — —. 2015b. "Bank Policy – Program-for-Results Financing." Washington, DC.
 - — —. 2015c. *2015 Development Policy Financing Retrospective: Results and Sustainability*. Washington, DC: World Bank.
 - — —. 2015d. *India – Enhancing Teacher Effectiveness in Bihar Operation Project*. Project Appraisal Document. Report No. 92972. Washington, DC: World Bank.
 - — —. 2015e. *Program-for-Results: Two-Year Review*. Operations Policy and Country Services. Washington, DC: World Bank.
 - — —. 2016. "Bank Procedure, Program-for-Results Financing." Washington, DC: World Bank.
- World Health Organization. 1995. *A Guide to Sector-Wide Approaches for Health Development: Concepts, Issues and Working Arrangements – Jointly Involving Governments and Development Agencies in Healthcare Provision through SWAps*. Geneva: WHO.