

The Anatomy of Program-for-Results: An Approach to Results-Based Aid

Alan Gelb and Nabil Hashmi

Abstract

The World Bank's new Program for Results (PforR) instrument is only the third instrument approved by its Board and the first to directly link disbursements to results. Designed to support programs of service delivery, the program is still in its early stages. This paper provides an overview of the approach and some of the debates on the design of the instrument, including the approaches to safeguards and to "results", which encompass the strengthening of systems of service delivery as well as the actual delivery of services. It develops a classification of Disbursement-Linked Indicators (DLIs) that can be used to situate the results-based instruments in the context of investment loans (IL) and development policy loans (DPLs) and applies this to the first four PforR operations. Some are shown to approach other results-based formulations (for example COD Aid) while others have a larger overlap with DPLs. The paper notes a number of features of the operations, including the still-modest use of system performance indicators, as opposed to action-based indicators, to link disbursements to systems reform, and the implications for the PforR approach. While it is far too early to judge the success of PforR and its various design features, the paper considers some of the implications, including the meaning of "success" for a results-based operation.

JEL Codes: F35, F36, O19

The Anatomy of Program-for-Results: An Approach to Results-Based Aid

Alan Gelb
Center for Global Development

Nabil Hashmi
Center for Global Development

We thank Bill Savedoff, Ravi Kanbur, and an anonymous referee for very helpful comments and express appreciation for the insights gained from discussions with Fadia Saadah, Aymeric Meyer, and participants at a number of seminars and workshops on PforR. All errors, including of interpretation, are those of the authors.

CGD is grateful for contributions from the Australian Department of Foreign Affairs and Trade, the Bill & Melinda Gates Foundation, and the UK Department for International Development in support of this work.

Alan Gelb and Nabil Hashmi. 2014. "The Anatomy of Program-for-Results: An Approach to Results-Based Aid." CGD Working Paper 374. Washington, DC: Center for Global Development.

<http://www.cgdev.org/publication/anatomy-program-results-approach-results-based-aid-working-paper-374>

Center for Global Development
2055 L Street, NW
Fifth Floor
Washington, DC 20036

202.416.4000
(f) 202.416.4050

www.cgdev.org

The Center for Global Development is an independent, nonprofit policy research organization dedicated to reducing global poverty and inequality and to making globalization work for the poor. Use and dissemination of this Working Paper is encouraged; however, reproduced copies may not be used for commercial purposes. Further usage is permitted under the terms of the Creative Commons License.

The views expressed in CGD Working Papers are those of the authors and should not be attributed to the board of directors or funders of the Center for Global Development.

Contents

| | |
|--|----|
| I Introduction | 1 |
| II Some Considerations around Program-for-Results | 3 |
| III A Classification of Disbursement-Linked Indicators | 7 |
| IV The First Four Projects | 10 |
| V Conclusion | 17 |
| References | 20 |
| Annex | 21 |

I Introduction

The World Bank's new Program-for-Results (PforR) instrument is one of a number of approaches to link the flow of assistance to development achievements.¹ It is only the third financing instrument approved by the Bank's Board and promises to become a significant financing mechanism alongside traditional project loans and development policy loans. PforR is intended to improve effectiveness in a particular area – programs of service delivery – and to bridge a perceived instrument gap between investment loans (IL) best suited to large, lumpy, capital projects and quick-disbursing policy loans (DPLs) designed to encourage policy reforms. The central design feature of the results-based approach in general, and of PforR in particular, centers on “results”. What qualifies as a result? How are results captured and measured in the indicators used to trigger disbursements? How exactly do they determine disbursements and allocate risk between the lender and the borrower? This paper considers these questions, and how they define the anatomy of PforR operations.

Section II provides a brief overview of the PforR approach and some of debates on its particular design features. It notes that the concept of “results” in PforR extends beyond traditional service delivery—measured by outputs such as school enrollments or outcomes such as learning—to include indicators of the performance of the systems that provide the services. The fact that the instrument is called “Program for Results” rather than “Payment for Results” is no accident; the aim is to improve the systems for service delivery as much as it is to deliver services. This makes PforR different to some other approaches such as Cash on Delivery (COD) aid that consider a major benefit of results-based financing to be a transparent “hands-off” approach where project implementation is left entirely in the hands of the client. Nevertheless PforR does raise some of the same questions as other results-based approaches, including how to measure “success” in cases where the results are not achieved but the operation functions as it is designed to do and fails to disburse.

Another feature of PforR is its approach to fiduciary, social and environmental safeguards, where it builds on country systems rather than imposing the specific safeguard policies of the Bank. This has been an area of debate but is central to the objective of leveraging the PforR approach, including its results frameworks, to cover expenditure programs that are

¹ Other approaches include the EU's MDG Contracts, Cash on Delivery (COD) Aid pilots in Ethiopia and Rwanda, REDD+ programs, the Performance-Based Grants of the Global Fund, the Global Partnership on Output-based Aid, as well as IDB partnerships on health programs in Central America. For an extensive review of results based aid and financing, see Pearson, Johnson and Ellison 2012.

only partly financed by the Bank. Some safeguard-related issues are noted for the way in which they shape the operations, but this area is not discussed in depth.

Section III develops a classification for Disbursement Linked Indicators (DLIs) and Section IV applies this to the first four PforR operations in Nepal, Uruguay, Morocco and Tanzania. As explained in more detail below, DLIs are not peculiar to results-based instruments. Every type of financing operation has some sort of DLI, whether it is submitting documentary evidence of approved expenditures for reimbursement, undertaking particular policy actions or achieving measured outcomes. We develop a classification of DLIs that helps to understand how the PforR operations are designed and how different they are (or are not) from other instruments. As seen in the first four operations a wide range of “results” have been considered as eligible for PforR and the operations differ a great deal. Some have an overlap with development policy loans, reflecting the challenge of constructing “objective” measures of system performance that are not actual outputs. They also show the difficulty of developing contractual arrangements based on results for demand-driven service delivery systems where it is difficult to specify outputs or outcomes in advance.

Section V concludes. It is still far too early to try to assess the PforR instrument and the paper does not try to do this, but rather sets out some implications of the analysis. The program is still at a “learning” stage but appears to have made a strong start and appears to be seeing rising demand. Traditional investment loans and policy loans are still available, so that the approach will have to evolve to compete with these instruments if it is to become a widespread modality for World Bank support to programs of service delivery. In theory, just as the smarter, more adaptable mammals overcame the dinosaurs, the rationality of a results-based approach should enable it to prevail in such a Darwinian contest at least for the kinds of programs for which the approach is intended. In practice, it remains to be seen whether the push towards results and the willingness to take some risks to achieve them – both on the clients’ side and on the part of the Bank and its shareholders—is strong enough for such a natural selection process to occur. If not – just as the extinction of the dinosaurs seems to have been necessary to enable the mammals to emerge – a conscious decision to limit the use of other instruments might be needed for such a results-based approach to become the norm. Any comments on this question would be purely speculative at this time.

II Some Considerations around Program-for-Results

Until January 2012 when PforR was formally approved after a lengthy process of consultation, the Bank had only two types of financing instrument, investment loans (IL) and development policy loans (DPL).² ILs have accounted for about 80 percent of the Bank's lending and typically have a longer time horizon than DPLs. They typically disburse against records of expenditures incurred for agreed project inputs. DPLs generally disburse more quickly and support specific policy reforms, and they can be designed with one or multiple tranches. In addition to financing inputs such as construction materials or power plants and being conditioned on policy changes, both types of operations may include capacity-building components to strengthen implementation.³

PforR was designed to fill a perceived gap in the instrument space between investment and policy loans, to support programs of spending that are neither large, lumpy, investments nor centrally focused on policy change. To some extent the new approach formalized and built on the growing use of sector-wide approaches and other programmatic operations. In some cases the Bank funded only a “slice” of these programs, some of which incorporated performance-based elements. While sectors like education and health come immediately to mind, such service delivery programs can be important in many other sectors, for example, repairing roads or providing water connections. In the first two years PforR commitments are limited to 5 percent of annual IDA/IBRD commitments. As of early 2014 some 24 operations had either been approved (10) or were under preparation (14) in nineteen countries, both IDA and IBRD. They cover a wide range of sectors, infrastructure (10), education, health and social protection (8), as well as public sector management, with some operations having a mix of sector emphasis. All involve the delivery of some form of service through an identified program of activities. The Bank is to conduct a two-year review of initial experience in 2014, following which the limit on commitments will be reviewed.

PforR has been welcomed by major stakeholders, including the US⁴ and other donor countries, as well as by officials from many borrowing countries - whether in finance

² The Bank has at various times appeared to have many more types of instrument, including SWPs, LILs, SILs, APLs, etc., but all fall under either the investment loan or the policy loan umbrella. For more details see http://siteresources.worldbank.org/PROJECTS/Resources/ALR06_20_01.pdf and <http://www.rojasdatabank.info/saprin/367720rev0pdf.pdf>

³ This is not to say that investment operations are devoid of policy content. For example, a power operation may require tariffs to be raised to ensure sustainability.

⁴ http://www.treasury.gov/resource-center/international/development-banks/Documents/1_30_2012_P4R_US_Position%20Statement_Final.pdf

ministries or sectors - who see it as supporting their own efforts to introduce results frameworks into their programs and to strengthen their effectiveness. Early indications are that the approach has indeed increased attention to the importance of results frameworks⁵ and encouraged partnerships through the focus on shared goals. As with any new approach, the learning curve has been steep, both for officials in client countries and for the Bank staff responsible for preparing and supervising the projects.

As the approach was being developed debates fell into two main categories. The first set of concerns came from **safeguard constituencies**, individuals and organizations concerned with social or environmental protection, as well as fiduciary risk.⁶ Bank investment loans must accord with a standard set of safeguard policies that aim to protect people and the environment and also to reduce the risk that funds are diverted to unauthorized uses. In contrast, PforR operations assess country safeguard systems and apply them to the program of focus, and where necessary seek to strengthen them, rather than segregate Bank funds into a special pool administered through Bank processes. This is because a key goal is to leverage the project to build wider institutional capacity and improve the effectiveness of the entire program—PforR operations typically cover only part of the costs of the programs they support. If the results framework for the Bank’s funding pushes the borrower to operate more effectively throughout the whole of the project –not just the portion financed by the Bank—PforR has the virtue of leveraging relatively small pools of financing for large development gains and creates a distinctive partnership role for the Bank.

Is it necessary to be concerned about safeguards if disbursement is conditional on achieving well-defined results? In the area of fiduciary safeguards the results-based approach has a potential advantage. If the costs of achieving the results can be accurately estimated and built into a program and the project does not provide more funding than needed to achieve them, results-based conditionality logically substitutes for fiduciary safeguards. If the money is lost, stolen or wasted, the result cannot be achieved so that the loan cannot disburse.⁷ The argument is even stronger if the program leverages the World Bank funds so that achieving the results provides some assurance on the use of other funds as well. This argument has not

⁵ A results framework is a brief schematic of overall goals, inputs, outputs, outcomes, and monitoring and evaluation strategies. Bank stakeholders and borrowers may use these frameworks to understand what their major goals are, what intermediate steps are necessary, and how progress towards those outcomes will be realized and measured. For a good overview of the Bank’s use of results frameworks, see http://siteresources.worldbank.org/EXTEVACAPDEV/Resources/designing_results_framework.pdf.

⁶ Many of the concerns are summarized in Alexander 2011, 2012; see also Crippa et al 2011.

⁷ For a good overview of how results-based programs may be less vulnerable to corruption, see Kenny and Savedoff, “Can Results-Based Payments Reduce Corruption?” (2014).

been fully accepted by the Bank. Individual procurements in excess of \$50 million are excluded from PforR projects. In addition, the Bank's anti-fraud and corruption department (INT) still retains authority to investigate the operations. This can be a concern for the clients since the operations also leverage the potential involvement of INT beyond the use of the Bank's own funds and across the entire program.

Results-based disbursement does nothing in itself to alleviate possible concerns in the areas of environmental and social safeguards. For example, a project to extend or repair a road network and funded only on the basis of mileage rehabilitated could open up environmentally sensitive areas to logging, threaten rare species or displace vulnerable communities. Results-based projects will therefore be more attractive – at least to donors – if they address less sensitive sectors so that the focus can be on results rather than on risks in other areas. In response to these concerns, all Category A operations are excluded from PforR, these being ones that risk significant and irreversible negative impacts on the environment or affected peoples. Critics, including some environmental and social NGOs, were not fully convinced by this exclusion, or by the argument that it is not helpful to apply Bank processes to a possibly small part of a program and ignore the rest) and argued that the Bank's abdication from its own safeguard standards would increase risk and dilute the accountability built into the system for investment loans.

The approach to safeguards in PforR therefore has two implications. On the one hand, it strengthens the partnership approach, avoiding the segmentation of donor funds into a separate pool and enabling the results framework to extend to expenditures funded by the country or other donors. On the other hand, excluding particularly sensitive segments of a sector (or large procurements) to mitigate risk could carve out a results-based program that resembles a Swiss cheese, full of holes and excluding critical areas.

The second set of questions around PforR is raised by “**results constituencies**”, organizations and individuals pressing for simpler, more effective and more credible aid instruments. A results-based approach should be simpler than the convoluted agreements and intrusive, yet frequently ineffectual, monitoring that characterizes so much aid delivered through traditional mechanisms. As outlined in the COD approach – the purest model of results-based financing (Birdsall and Savedoff 2009)–donors and recipients first decide what observable outputs or outcomes they want. They agree on a very few transparent indicators that can be monitored in a timely and unambiguous way. They define baselines, and agree on how gains should be reflected in disbursements. They ensure that there is a competent,

timely and independent monitoring system in place. The donor disburses according to the progress the recipient has made toward the specified goals. The donor does not prescribe how to achieve the goals; this is the responsibility of the recipient who has space to experiment with different approaches. Accountability is clear. The more rapid is the progress, the faster flows the money.

In practice, life is not so simple. In addition to concern over safeguards, donors may be under pressure to increase the complexity of the operations in response to a variety of concerns and mandates. In the case of PforR, the focus on the systems as well as the services they deliver can reflect a concern that results should be sustainable after the conclusion of the project.⁸ It could also reflect a continuing mindset that the Bank has a responsibility to engage into the nuts and bolts of implementation, especially when dealing with client countries with modest levels of capacity. Simply focusing on outputs or outcomes may seem a step backwards relative to more traditional financing instruments that seek, in one way or another, to reform policies and strengthen capacity. Donors – especially donors with a traditionally large “policy weight” like the Bank—may thus have wider horizons than the actual results (in the sense of outputs or outcomes) targeted by a project.^{9,10} Recipients may also have concerns, including over how disbursement based only on the achievement of particular outcomes or outputs allocates performance risk and the cost of measurement errors. They might therefore favor some disbursements set by measures of systems performance, or efforts towards improving it.

This of course raises the question of whether it is possible to objectively measure the performance of systems in ways that do not depend on measuring actual development outputs or outcomes produced by these systems. It also raises the risk of increased

⁸ “There is also a risk that the schemes (RBF in particular) become a “sticking plaster” which may reduce the perceived need for the more fundamental, and often politically sensitive, reforms which might be required to sustain progress. A distinction needs to be made between the valid long term role of RBA/RBF and the shorter term “catch up” role they might play which suggests the need for a long term vision of how schemes will contribute to system development over time.” Pearson, Johnson and Ellison 2012, p3.

⁹ World Bank (2012). Early PforR experience - Exchanging views and emerging lessons. Center for Global Development, Washington, DC. It is notable that Institutional and financial sustainability are included as standard evaluation criteria in IEG project evaluation methodology.

¹⁰ To take an example, a one-off road repair might have little permanence without reforming a corrupt and inefficient contracting system. What are the donors to do when the newly improved road crumbles – fund the same repairs again and again? The system approach seeks to eliminate this aid trap although, just like the road itself there is no guarantee that any systemic improvements will endure after the end of the operation. A similar sustainability issue arises when considering repeated applications of COD aid. In principle, each successive baseline benchmark should be raised to reflect progress already made, but this leaves the donors in a difficult situation if for some reason a country slips badly in terms of development results. They cannot apply COD aid again to assist recovery. If the newly-repaired road network deteriorates, the country must be left to its own devices until it has returned to its previous peak.

complexity; the more complex is the approach the lower is its transparency and the greater is the risk of diluting the simple results-based model. Will the resulting instrument be credible in the face of pressure to disburse? If performance is “close” to triggering disbursement might the country or the Bank encourage the monitors to be lenient, especially if disbursement is conditional on reaching a particular threshold? Can monitors really be independent if hired by the country—or by the Bank?¹¹

All development projects are implemented under conditions of uncertainty. Some performance risks are specific and predictable, but especially if performance has a measure of “judgment” the risk is harder to quantify. Those managing the country relationship might argue that, with the benefit of hindsight, the targets were set in an over-optimistic way. Under what conditions is it reasonable to revise them? This raises the possibility that if the parties are both risk averse and both have a bias towards disbursing funds rather than withholding them, the only credible projects will be those with low levels of ambition.

The link between results and disbursements could also be weakened through indirect mechanisms. PforR operations may be part of a country portfolio – as in an IDA program—where the overall level of resources is not itself based on results and it is possible to restructure the portfolio by shifting funds out of poorly performing projects without losing them. For a traditional portfolio this flexibility is helpful; it encourages pro-active management of problem projects and also facilitates response to unexpected needs. However, for a results-based operation it opens up the unsettling possibility that a country that fails to achieve the results could still get the money in the end through other projects.¹²

III A Classification of Disbursement-Linked Indicators

All financing instruments have conditions for disbursement. In practice a single operation may combine conditions of different types, but it is useful to start off with a stylized picture that distinguishes three types of operations.

Investment operations generally finance (or reimburse) the purchase of an agreed bundle of project inputs. Their DLIs can therefore be classified as Input-Based (I). Of course they may

¹¹ This issue is not specific to results-based aid: similar concerns have been levied over rating agencies which are generally paid by the entities they rate (“issuer pays”) rather than paid by those who invest based on the ratings (“subscriber pays”). Research conducted by the American Enterprise Institute in 2008 suggested that subscription-based ratings were more accurate: see <http://www.cfr.org/financial-crises/credit-rating-controversy/p22328>.

¹² One way to approach this problem is to drive IDA allocations more strongly by the results obtained through the IDA portfolio. Gelb 2013 argues for such an approach, which would also encourage the use of results-based instruments more widely in these countries.

also require changes in policies or regulatory frameworks as conditions of disbursement and may include elements of capacity building. It is technically possible to add on conditions of a results-based nature as a condition of disbursement as has been done in some investment operations that preceded PforR, but this creates a situation of “double jeopardy” for the client who is forced to shoulder additional performance risk on the use of funds whose use is heavily determined and closely monitored by the Bank.¹³

Disbursements under DPLs are conditional on specified policy actions. Some may be preconditions for a single-tranche operation and others specified in advance within a multi-tranche framework. All should be under the control of the government. These actions are, in principle, straightforward to monitor. Their baselines are sometimes explicit (reduce the average tariff from 40% to 20%) and sometimes implicit (introduce a new policy where none existed before). They are denoted as Action DLIs (A).

Disbursements under PforR are conditional on “results”. These are usually measured at regular intervals within a project horizon of several years. They can be outputs, for example, the number of bridges repaired or roads rehabilitated. They can also be outcomes further down the results chain, for example the reduction in measured transit time between two points on the road network. While the conceptual distinction is clear, it is not always easy to classify a particular result as an output or outcome because this depends on the length of the causal chain downstream from the operation.¹⁴ DLIs considered to be Outputs are denoted as O and those that are considered to be outcomes as OO.¹⁵

Indicators for the performance of delivery systems raise similar issues to measuring the quality of governance. As debated in an extensive literature (for example, Kaufmann and Kraay 2007), governance indicators can be of many different types: *de jure* or *de facto*, “objective” or “subjective”, “actionable” or otherwise. In practice, “objective” indicators

¹³ Elements of a results-based approach can be found in some investment operations. For example, in the Minas Gerais Partnership for Development program, IBRD funds are disbursed against indicators in six sectors related to improved public sector management. The government finances the bulk of the program with IBRD funding about 18 percent.

¹⁴ For example, consider the following results chain for an education project: classrooms constructed, teachers hired, textbooks provided, pupils enrolled in school, pupils attending classes, completing primary schooling, achieving learning results as measured by test scores, achieving gainful employment, increases in household income, multi-dimensional advance out of poverty, happiness. While the first is surely an output and the last surely an outcome, the intermediate results could be considered as either. In addition, a given input, output, or outcome within a project may not be the same category in a different context. The output of a construction or manufacturing project could be the input to an education or health program.

¹⁵ Outcomes are generally preferred as better representing development goals, but outputs can have advantages as a contractual basis for loan disbursement. Many can be monitored through technical assessments or surveys to provide a rigorous indicator. It may also be easier to cost outputs than outcomes, to help provide an empirical scale to link disbursements to achievements if that is desired.

often focus on *de jure* policy actions or system inputs – for example, whether a particular report was presented to the Auditor General – rather than *de facto* measures of actual government performance or system outputs. The latter are often based on expert surveys, which raises the question of whether they measure actual performance or perceptions of performance that could be influenced by other things, such as the release of new information.¹⁶ It is possible to construct more “objective” indicators for *de facto* performance and easier when a system produces a clearly defined service. Savedoff (2011) offers a number of performance indicators for health systems; for customs administration Rabelland (2013) suggests a combination of higher revenue and faster clearance, both of which can be measured objectively.

System DLIs can be classified based on this literature. They may be individual policy actions (A) as above. They could also be groups of actions, bundled together in complex plans of institutional reform. These “System Actions” could of course be broken apart into their individual components, but sometimes at the cost of introducing an extremely long list, so that it is convenient to bundle them together. They are denoted SA and because of the bundling will be less transparent than simple individual actions. System Output indicators that measure its performance are denoted SO. Table 1 lays out these six types of DLIs with an example for each.

Even if the differences between these DLIs is conceptually clear there can be some degree of ambiguity in determining whether a particular criterion should be seen as an action or an intermediate output, particularly in capacity-building programs where successive actions (SA) can also reflect improvements in system capacity that results from previous actions. The classification below does not break out composite system DLIs in this way but the ambiguity needs to be kept in mind.

¹⁶ For example, the breaking of sensational news on a case of grand corruption may increase awareness of corruption that has been endemic all the time.

Table 1: DLI Classification

| DLI Type | Abbreviation | Example |
|---------------|--------------|--|
| Input | I | Presentation of invoice for approved purchase of construction materials |
| Output | O | Number of new bridges built according to agreed plan |
| Outcome | OO | Percent of secondary-school girls reaching agreed standard of achievement on standardized test |
| Action | A | Preparation of Environmental and Social Guide |
| System Action | SA | Implementation of an agreed program to strengthen the Bridge Management System |
| System Output | SO | Increased percentage of bridge works completed on pre-agreed schedule |

The balance between DLIs shapes the anatomy of a PforR operation, in particular its relative emphasis on outputs or outcomes, policy changes and institutional development. A heavy emphasis on O or OO indicators would position the project more closely as an alternative to an investment loan. Having mostly A type DLIs (or SA if these are indeed mainly actions) makes the operation more comparable to a DPL policy loan. A heavy emphasis on systems as shown by SA and SO suggests a comparison with more traditional capacity-building operations, but hopefully with a stronger emphasis on measuring and rewarding the results of capacity-building efforts rather than just the efforts made to strengthen capacity.

IV The First Four Projects

Full information was publicly available for the first four projects at the time of writing. Table 2 presents summary information on the projects and more detailed information can be found in Annex 1. Their commitments total \$681 million but cover only part of the costs of the defined programs which in three cases are supported by other donors and the country's own funds.

Table 2: Summary of the First Four Operations

| Project Name | Country | Project Objective | Loan (Program) \$million |
|---|----------|---|-----------------------------|
| Bridges Improvement and Maintenance Program (P125495) | Nepal | Maintenance and construction of bridges in Nepal's Strategic Road Network | 60 (148) |
| Road Rehabilitation and Maintenance Program (P125803) | Uruguay | To ensure that at least 35 percent of the National Road Network is in good or very good condition and to improve road sector management | 66 (510) |
| National Initiative for Human Development (INDH) Phase II (P116201) | Morocco | To improve participatory local governance mechanisms, social services, basic infrastructure, and economic opportunities for targeted groups | 300 (2,100) |
| Urban Local Government Strengthening Program (P118152) | Tanzania | To improve institutional performance for urban service delivery in local government authorities | 255 (255) |

The Nepal Bridges Improvement and Maintenance operation supports the Bridge Program for Nepal's strategic road network that covers the main arteries for trade and economic activities. The Bridge Program spans 2013 to 2017 and will cost \$147.6 million; of this, IDA will finance \$60 million. The Uruguay Road Rehabilitation and Maintenance project aims to ensure that at least 35 percent of the national road network is in good or very good condition and also to improve road sector management. The program's total size is \$510 million; of this the Bank contributes \$66 million. The Morocco Human Development operation focuses on increasing access to and improving the quality of local governance mechanisms, basic infrastructure, social services, and economic opportunities. It supports Phase II of a large national program (INDH) with a total cost of \$2.1 billion. The Bank will contribute up to USD 300 million. The Tanzania Urban Local Government Strengthening operation aims to improve institutional performance for investments and service delivery in specific urban local government authorities through a new Urban Performance Grant system. This new system is at present fully funded by the Bank to the amount of \$255 million.

All of the projects support Category C operations, those seen as posing the lowest social and environmental risks. In the case of Nepal, this is achieved by ring-fencing the program to

exclude bridges in sensitive areas of national parks. While this reduces the risks of the operation, it also illustrates the downside of the exclusions: their repair could be supported by other donors and projects, without whatever level of oversight and assurance that might be provided by inclusion in the PforR project.¹⁷

Table 3 sets out the DLIs for the operations, distinguishing them by type (I, O, OO, A, SA or SO). The indicators are also classified by whether disbursement is scaled in proportion to performance (S) or is conditional on achieving a threshold level for the DLI (T), and by whether a baseline for the DLI is set out in the project document. In some cases baselines may be implicit (Im) as in the case of a simple DLI of type A. The disbursements associated with the DLIs are cumulative, over the life of the project. Table 4 summarizes projected disbursements by type of DLI.

Table 3: DLIs for the First Four Operations

| Country | DLI Description | Type (I,O,OO A, SO, SA) | Scale or Threshold (S,T) | Baseline (Y, N, Im) | Value, \$ million |
|---------|--|-------------------------|--------------------------|---------------------|-------------------|
| Nepal | Completion of major maintenance of bridges on Strategic Road Network (SRN) (cumulative meters) | O | S | Y | 30 |
| | Completion of minor maintenance of bridges on SRN (number of bridges) | O | S | Y | 3 |
| | New bridges built or improved on SRN (cumulative meters) | O | S | Y | 18 |
| | Strengthened performance management in bridge sector (percentage of works completed on planned schedule) | SO | S | Y | 3 |
| | Improved Bridge Asset Management (Bridge Management System operational, undertake surveys) | SA | S | Im | 3 |
| | Increased effectiveness of the institutions responsible for bridge sector management | SA | T | Im | 3 |
| Uruguay | Cumulative km rehabilitated on the National Road Network, at a minimum rating of 85 | O | S | Y | 26 |
| | Number of km of the National Road Network maintained through performance-based contracts | A | S | Y | 30 |
| | Approval of the multimodal plan for Montevideo seaport access | A | T | Im | 2.5 |
| | Approval of the catalogue for technical solutions of pavement rehabilitation | A | T | Im | 2.5 |
| | DNV Environmental Manual has been updated and at least 75 percent of technical staff have been trained up to a standard acceptable to the Bank | A | T | Im | 2.5 |
| | The guidelines for expropriation and social management processes have been approved and an international workshop on best practices for road works social management has been carried out. | A | T | Im | 2.5 |

¹⁷ Another way to address exclusions could be hybrid projects with some components PforR and others associated investment operations.

| | | | | | |
|----------|---|--------------|---------------|-------------|-----------------|
| Morocco | Percent of girls who reside in educational dormitories graduating to the next grade | OO | S | Y | 40 |
| | Percent of population provided with access to improved water supply in targeted rural communes by the Program | O | S | Y | 40 |
| | Percent of income-generating activities implemented by cooperatives, associations or companies which are viable two years after having benefited from financing under the Program | OO | S | Y | 40 |
| | Percent of infrastructure projects financed under the Program judged by the auditors as conforming to technical specifications, after final commissioning | O | S | Y | 20 |
| | Percent of women and youth in certain local governance bodies | A | S | I | 30 |
| | Percent of projects under the rural and urban subprograms of the Program contracted by the local government, associations, or cooperatives. | A | S | Y | 35 |
| | Percent of provinces and prefectorates in the Program Area which have put in place a plan of action to address audit recommendations | A | S | Y | 35 |
| | Percent of priority audit recommendations included in the action plan which are implemented | A | S | Im | 20 |
| | i) Preparation of environmental and social guide related to the Program and ii) percent of key actors trained in the use of such a guide | A | i) T ii) S | Im | i) 20 ii) 20 |
| Tanzania | Urban Local Government Authorities (ULGAs) have strengthened institutional performance and achieve Program minimum conditions in the annual assessment | SA note 1 | S | N | 45 |
| | ULGAs have strengthened institutional performance as scored in the annual performance assessment | SA note 1 | S | N | 106 |
| | Local infrastructure targets as set out in the annual action plans are met by ULGAs utilizing the Program funds | O | S | N note 2 | 50 |
| | Number of ULGAs with all core staff in place | A | S | N | 14 |
| | Completion of annual capacity-building activities for Program ULGAs | A | S | Im | 30 |
| | The Prime Minister's Office has adopted an enhanced Local Government Development Grant assessment system. | A | T | Im | 10 |

Note 1 DLIs 1 and 2 are scored by a composite index comprising 49 actions.

Note 2 Scoring for this indicator is against the annual action plan, not an absolute outcome, hence no baseline.

Table 4: Disbursement by Type of DLI\$ Millions (% of Loan)

| | Output/Outcome | Actions | System Actions | System Outputs | Total |
|----------|----------------|------------|----------------|----------------|-------|
| Nepal | 51 (85) | 0 (0) | 6 (10) | 3 (5) | 60 |
| Uruguay | 26 (39) | 40 (61) | 0 (0) | 0 (0) | 66 |
| Morocco | 140 (47) | 160 (53) | 0 (0) | 0 (0) | 300 |
| Tanzania | 50 (20) | 54 (21) | 151 (59) | 0 (0) | 255 |
| Total | 267 (39.2) | 254 (37.3) | 157 (23.1) | 3 (0.4) | 681 |

The DLIs. Overall, about 40 percent of total disbursements for the four operations are conditional on outputs or outcomes, mainly the former. Slightly more than one third is conditional on specific stand-alone actions. In most cases these are clear and they also have clear, sometimes implicit, baselines. About one quarter of disbursements depend on systems indicators. These are almost all system actions or individual actions, with only one systems indicator of a clear output (SO) type – the percentage of Nepalese bridge works completed on schedule as planned. Since it is not apparent that the schedules have been specified in advance, it is not clear that this constitutes an entirely robust basis for a contract. Most systems DLIs are of type SA and involve a large number of actions as well as intermediate steps. In Tanzania, the score-card measuring system performance is composed of 49 criteria each representing about 2% of the total maximum score of 100.¹⁸

The operations are very different in their emphasis on outputs or outcomes versus systems. From Table 4, disbursements for Nepal are mainly against outputs, with modest systems and actions components all in areas closely related to the delivery of the outputs. Uruguay and Morocco have lower shares of disbursements linked to outputs and disburse the remainder based on action DLIs. These operations blend a version of the simple results-based (O or OO) model with policy loan content. The Tanzania operation is the outlier. System improvements are the focus, measured as SA. These do include some criteria that could be considered as intermediate system performance measures¹⁹, but the heavy weight on actions raises the question of the weight of actual results in disbursements. Twenty percent of disbursements are indeed against outputs, measured by the completion rate for demand-

¹⁸ Previous governance projects in Tanzania have also featured extremely complex evaluation frameworks that were rated by the Bank's Independent Evaluation Group to have been of limited operational value World Bank Independent Evaluation Group. 2013. [Project Performance Assessment Report: Tanzania – Public Service Reform Project](#). Washington, D.C.

¹⁹ For example, consider the sequence: computerized property register in place; increase in the number of properties in the register; increase in taxable properties valued; increase in property tax collected. While the first is reasonably an action and the last reasonably a system output, the others are both actions and intermediate results from the program.

driven decentralized infrastructure plans at the local government level, but as these plans will evolve over time it is not possible to specify a baseline and an ex-ante contractual performance requirement in advance. This points to the difficulty of designing results-based operations for flexible and decentralized programs of service delivery that are expected to evolve over time.

Disbursements and Performance Risk. Performance risk is somewhat moderated by having disbursement depend on a number of DLIs rather than a single indicator. In addition, most disbursements are scaled in line with achievement, providing relatively stable incentives compared with threshold-type on-off programs that might or might not disburse even if there is some progress. A few DLIs are threshold-based; this provides a high-powered incentive at the margin but also runs the risk of discouraging effort when attaining 100% of the target seems doubtful. It may also increase pressure to modify the DLI in response to concerns about disbursement. In the case of Tanzania, achieving the minimum access conditions on DLI 1 for strengthened institutional performance is a condition for receiving any further support on the basis of the other DLIs. This could be a factor behind the concern about performance risk flagged in the documentation. If the minimum conditions local governments must meet to trigger disbursement are perceived to be too tough, some might drop out of the program or simply not choose to make any changes that would incur costs. Many dropouts could, in turn, create pressure to relax standards.

Performance Monitoring. There is certainly no shortage of performance monitoring. From Table 2, the four combined operations muster a total of 27 DLIs. When combined with multi-year disbursement monitoring, this results in an expected total of some 84 individual DLI decision points. On average, a country receives only \$8.1 million for each review of a DLI. Some of the complexity is due to combining different types of objectives within one operation but it also reflects the reality that even an apparently simple output goal, such as “rehabilitate the road network” cannot be linked to funding through a single indicator such as mileage rehabilitated if different sections of the network require different levels of repair.

The average payment per condition is even less if we decompose the systems DLIs into their multiple underlying actions. For Tanzania there are 18 municipalities; a single action taken by an average-sized municipality to strengthen institutional performance would be rewarded by a disbursement of only \$200,000.

Leverage. The operations are mostly highly leveraged, pointing to the potential ability of the new instrument to have a wide impact. Those for Nepal, Uruguay and Morocco provide only 15% of the total costs of their programs. The rest is covered by other donors, as well as funds from the countries themselves. Especially if the costing of the program outputs is sound, high leverage should strengthen the fiduciary comfort created by the results-based approach; to achieve the outputs, countries need not only to apply the Bank's funds effectively but those from other sources.

Of course, high leverage could also reduce the direct financial incentives provided by the operations to achieve the targets unless the other sources of funding also condition their support on performance. In the Uruguay case, the PforR operation leverages its approach beyond the Bank's own funds through the action DLI that disburses according to the number of kilometers of road maintained through performance-based contracts, irrespective of the funding source. This provides an additional incentive to leverage the results-based approach.

Tanzania is again the outlier. There is no clear potential leverage from other funding sources and the system-based nature of the DLIs provides no robust costing basis for the disbursements. In this case performance-based lending offers little comfort from the fiduciary perspective.

Baselines and Monitorability. Except for the case of Tanzania mentioned above, baselines, progress units and evaluation criteria are specified in loan documents for output/outcome DLIs. It is also generally straightforward to monitor the action DLIs. Was a specified report produced or a manual written? Did a capacity-building seminar take place as planned? If yes, disburse. Measurement is less clear for improvements in system performance. Even if the individual criteria are transparent and monitorable, their number increases the overall complexity of the operation.

Independent monitoring. The operations create some space between those responsible for implementing the project and those responsible for monitoring or auditing the DLIs. In many cases both the implementers and the monitors are government agencies, though not the same. In these cases, there are sometimes provisions for "independent" auditors, although in some cases the verification of results is to be by Bank staff.

However, as noted previously, this may not fully solve the problem, which is not peculiar to results-based aid. When implementers are required to hire a private firm to carry out an

evaluation, there is always the risk that they will choose a firm with low capacity to execute an in-depth audit. A firm paid by the project may not be impartial. Bank staff may also not be impartial particularly if they face pressure to disburse. They may be hesitant to overturn results from a (supposedly independent) agency of their development partner, especially when those results would trigger disbursement.

Despite the Bank's emphasis on using country systems in PforR operations, national statistics agencies have not been given a large role in monitoring performance. These agencies may be more impartial than implementing agencies, and tasking them with measuring key indicators could help build their capacity, improving the overall quality of a borrower's statistics.

Transparency and Citizen Engagement. All four operations refer to the importance of publicizing DLI results, but it is not always clear how and how soon the results will be put into the public domain. In particular, the publication of results achieved according to the monitoring of the DLIs is not a precondition for disbursement. Conditioning disbursements on the publication of the results of DLI reviews could increase transparency and create a valuable role for civil society organizations to monitor progress, and further strengthen domestic accountability and the focus on results.

V Conclusion

The PforR approach introduces an important new option for countries seeking financing from the Bank. It is still in its initial "learning-by-doing" stage. The operations included in this paper are all in the general financing niche for which PforR was developed: –programs of service delivery implemented on an ongoing basis over several years. Appropriately for a learning phase, the operations span a wide range of sectors and countries and they also show a mix of emphasis between services delivered and the performance of the delivery systems.

One contribution of the paper is to propose a classification of operations in terms of their DLIs. All financing instruments have some type of indicator related to the decision to disburse. We distinguish three basic types of operation, investment loans that disburse against inputs, policy loans that disburse against (policy) actions and results-based loans including PforR. These can vary a great deal, including in their relative emphasis on services delivered (outputs or outcomes) and the strengthening of the delivery systems providing the services. Some come reasonably close to the simple model of a results-based project approach like COD aid, with a heavy emphasis on major service outputs, if not outcomes, in

terms of DLIs and disbursements. Others are closer to policy and institution-building operations with a stronger disbursement weight on the delivery system than on any particular output or outcome produced by the system. Overall, almost half of the disbursements of the first four operations depend on outputs or outcomes, with the rest based on policy actions or programs of institutional strengthening.

The examples confirm the promise and flexibility of the approach but also raise some questions. On the positive side, the high financial leverage of most operations suggests the potential of PforR to strengthen partnerships around results frameworks beyond the activities directly financed by the projects and to further encourage reform-minded constituencies to improve the effectiveness of programs in their own countries. As the share of aid declines relative to the total resources available to developing countries this should become an increasingly important role for more analytically-oriented lenders such as the IFIs that combine financial assistance with sector expertise and advice.

Most of the cases specify performance baselines and provide for some institutional separation between implementers and monitors, if not actual independence. The cases also suggest that the performance risk inherent in results-based operations can be moderated by using several scalable DLIs that provide a graduated basis for disbursement rather than threshold-type on-off conditions.

On the questions, embedding programs of institutional strengthening into a results-based approach is clearly not simple. There is still debate on how to design objective measures of system performance that do not rely on the actual outputs or outcomes produced by those systems. Without more imaginative use of measures of system performance PforR operations risk being sucked into the orbit of more traditional capacity building loans or development policy lending by the exclusive recourse to complex action-based DLIs. This will reduce their transparency, as well as dilute the accountability that is a virtue of the results based approach.

As suggested by other studies, it should be possible to develop performance metrics that relate to the quality of service, such as leakages and deviation of resources, provider absenteeism, responsiveness and timeliness of delivery and unit costs. Also, new technology is increasingly being used by firms in the service sectors to measure and strengthen their

operational performance.²⁰ Use of such systems could help to improve the quality of data and provide feedback in real-time. PforR should push the dialogue in these directions.

PforR might also move further in the area of transparency, by requiring the posting and wide dissemination of DLI achievements as a condition of disbursement. This would offer an expanded role for civil society in terms of monitoring the veracity of assessments.

Another challenge suggested by the cases is the difficulty of applying a results-based approach to decentralized development projects that aim to build local capacity to both articulate demands and to provide desired services. Such operations cannot be expected to set out a predetermined sequence of outputs or outcomes, so that new ways will need to be found to link their performance to disbursements.

Finally, PforR raises the question of how to judge success, not of the program itself, which should be encompassed by the DLIs, but of the actual operations. If the results are not achieved and the project (appropriately) fails to disburse, should it be considered as a success or a failure? Without clear guidance on this question, which is bound to surface soon, there is a risk of biasing incentives against a demanding and truly results-based approach.

These are all still open issues. As stressed previously, it is still too early to assess either the individual operations or the PforR approach in general and this paper has not sought to do so. The next few years will provide essential information on how the new instrument works out for the future.

²⁰ United Parcel Service (UPS) for example has made a massive investment in its performance tracking and monitoring systems. These enable real-time feedback on distribution and deliveries and optimization of routes. See for example <http://www.wired.com/2013/06/ups-astronomical-math/>

References

- Alexander N. 2011. ["The World Bank's Proposed Program for Results \(P4R\): Implications for Environmental, Social, and Gender Safeguards and Corrupt Practices"](#). Heinrich Boell Foundation.
- Alexander, N. 2012. ["What is 'Country Ownership'? Do Social and Environmental Safeguards Destroy It: The Case of the World Bank's Program-for-Results"](#). Heinrich Boell Foundation.
- Birdsall N. and W. Savedoff. 2010. Cash on Delivery: A New Approach to Foreign Aid. Center for Global Development, Washington DC.
- Kaufmann D. and A. Kraay. 2007. ["Governance Indicators: Where Are We, Where Should We Be Going?"](#). Policy Research Report 4370, World Bank.
- Crippa, L. , P. Kebec and G. Gordon. 2011. ["Comments and Recommendations on the World Bank's Draft Op/Bp 9.00 Program-for-Results Financing"](#). Indian Law Resource Center. .
- Pierson, M., M. Johnson and R. Ellison. 2012. ["A Review of Results-Based Aid \(RBA\) and Major Results-Based Financing Schemes. \(RBF\) "](#). UK Department for International Development, Human Development Resource Centre.
- Raballand, G. and A. Rajaram. 2013. ["Behavioral Economics and Public Sector Reform: An Accidental Experiment and Lessons from Cameroon"](#). Policy Research Working Paper 6595, World Bank.
- Savedoff W. 2011. ["Governance in the Health Sector: A Strategy for Meaningful Definitions of Performance"](#). Policy Research Working Paper 5655, World Bank.
- World Bank (2012). [Early PforR experience - Exchanging views and emerging lessons](#). Center for Global Development, Washington, DC.

Annex

The First Four Operations

The Nepal Bridges Improvement and Maintenance operation supports the Bridge Program and Nepal's Strategic Road Network (SRN), the main roads for trade and economic activities. The Bridge Program spans 2013 to 2017 and will cost USD 147.6 million. IDA will finance approximately \$60 million, or \$12 million per year for 5 years, while the government will finance the remainder.

There are no other external partners, but other development partners are supporting standalone SRN bridge projects. The Asian Development Bank, the Japanese International Cooperation Agency, and the governments of China and India are funding bridges associated with individual road projects. Implementation for these individual projects will be governed by their own procedures.

There are six DLIs in total—three for outputs and three for systems. Output DLIs (DLIs 1 through 3: the length and number of bridges built, improved, or maintained) account for 85 percent of the total loan amount. DLIs 4 through 6 relate to systems and represent 15 percent of the loan. Nepal will receive these disbursements for introducing new management systems, publicly disclosing information related to grievances, and completing works on schedule. Every DLI is scalable except for DLI 6, which will only be disbursed if consultant reports and the National Planning Commission decide that the DLI has been achieved.

Baseline data for the output-based DLIs are 0 by definition since the DLIs concern repairs or improvements made after the project's launch. In most cases, the Bridge Management System (BMS), which is managed by the Bridge Project team at the Department of Roads (DOR), as well as consultant reports, will be used for reporting. The National Planning Commission and a Bank team will provide verification, and the Commission will also public post information about the program and its achievements. The Office of the Auditor General will perform financial and performance audits, while the National Vigilance Centre will conduct a technical audit. Other entities, including external firms, will conduct internal reviews.

The Bank has given the project an overall risk rating of "Substantial". The primary DLI-related risk is that the DLI targets may not be achievable given resource constraints. The Bank has concerns about the technical capacity of both bridge designers and local

construction companies. Fiduciary risks relate to improper procurement, diversion of funds, and corruption. Existing audit procedures may also be weak. The Bank has noted that the Environmental and Social Management Framework developed for the SRN may lack bridge-specific considerations. Furthermore, the unit with the DOR that manages social and environmental risks may lack the mandate, resources and staff capacity to support the SRN Bridge Program.

One unique safeguard is the exclusion of bridges that fall within the boundaries of Nepalese national parks. These include 49 existing SRN bridges, 5 SRN bridges under construction, and 21 planned new SRN bridges. Their construction and repair may be financed by some of Nepal's other development partners – without the benefit of any scrutiny from the PforR operation. The DOR will upload results from the BMS to its website for the public to see and post other relevant information on its website or in accessible media outlets.

The Uruguay Road Rehabilitation and Maintenance project aims to ensure that at least 35 percent of the Uruguay National Road Network is in good or very good condition and to improve road sector management. The program's total size is USD 510 million. Of this total, \$283.4 million will be financed by the government of Uruguay. The Bank is contributing \$66 million, and IADB has committed \$67.7 million. Finally, FONPLATA, the Fund of the Development of the Countries of the Rio del Plata Basin has loaned \$92.8 million. IADB and FONPLATA's contributions go towards a set of specific contracts related to road rehabilitation, road safety, and bridges.

Output-based DLIs 1 and 2, which concern the cumulative number of kilometers rehabilitated and the number of kilometers maintained through performance-based contracts, account for 84 percent of the loan. The remaining 16 percent goes to the four sub-indicators of DLI 3, which are all for institutional strengthening. DLI 2 disburses based on the number of kilometers of the National Road Network maintained through performance-based contracts. This feature has the potential to leverage the focus on results beyond the relatively small share of financing provided by the Bank. DLIs 1 and 2 are scalable, while the four sub-indicators that make up DLI 3 are binary.

The Road Department is responsible for collecting data on most of the DLIs, but the Planning and Logistics Department will monitor DLI 3-1, and the Topographic Surveying Department will monitor DLI 3-4. Baseline data exist for DLIs 1 and 2, while the baselines for the other DLIs are trivial because these are actions. For verification of DLIs 1 and 2, the

Ministry of Transport and Public Works (MTOP) will contract out an independent auditor through international bidding. The Bank will verify all of the sub-indicators in DLI 3.

The overall risk of the program has been assessed as “Moderate”. There were no major technical, fiduciary, or environmental risks identified. One risk could be that DLIs are overly ambitious; the Bank notes, however, that the DLIs are realistic and targets were computed in a conservative fashion.

Reports issued through the country budgetary system show information on spending and budget execution, but do not provide enough detail to track spending at the activity level. As part of the program, the Bank and the government agreed to rely on these reports for the first year, before determining how best to revise the system to include sub-activity data in the future. Public information about the program will be available on the MTOP website, and the Ministry will conduct an annual satisfaction survey of road users.

The Morocco Human Development operation focuses on increasing access to and improving the quality of local governance mechanisms, basic infrastructure, social services, and economic opportunities. It supports Phase II of the INDH program, following INDH I, which lasted from 2005 through 2010. While INDH II has five sub-programs, the Bank will support only three of them: the rural sub-program, the urban sub-program, and the crosscutting sub-program. The total cost of INDH II is USD 2.1 billion. Most of this will come from the central government and local governments. The Bank will disburse up to USD 300 million towards the program. Other external partners include the European Union, which has committed €20 million and also plans to use a results-based approach. The Bank has already disbursed USD 54.7 million or about 18 percent of the total possible.

There are nine Disbursement Linked Indicators. Output-based DLIs, like the percentage of girls who reside in the educational dormitories and complete the school year, represent 56 percent of the loan. The remaining 44 percent is conditional on systems-based DLIs. Of these, some are actions such as preparing an Environmental and Social Guide, while others are system-strengthening goals such as implementing audit recommendations.

Disbursements are scalable; faster progress will result in faster disbursement. Of the USD 300 million loan, 60 million per year was scheduled to be disbursed in 2012 and in 2013 with USD 90 million allocated for each of 2014 and 2015.

The National Observatory for Human Development (ONDH) is responsible for impact evaluation. It carried out a baseline survey during INDH I with Bank support. Baseline data

for all of the output-based DLIs are available. Various agencies are responsible for collecting data on different DLIs. For most, INDH is responsible for data collection but for some auditors from the Interior Ministry and the Ministry of Finance will collect data. For a single DLI concerning access to improved water supply in rural areas, the National Office for Water & Electricity will take charge of data collection. There will be regularly conducted audits by the Interior Ministry and Ministry of Finance auditors. The Bank will also hire independent firms to spot check on DLIs that cumulatively account for 74 percent of the total loan amount. In addition, the support provided by the EU includes an external and a financial performance audit.

The fiduciary risk of the program was rated as “Substantial”. There are provisions to track spending, but many government accounts currently do not conform to International Public Sector Accounting Standards. The project document notes a lack of transparency in the procurement process, especially with regards to reporting and resolving grievances. The Bank has also expressed concern about the capacity of government staff, community associations, and technical supervisors to execute the procurement process properly and to deliver high-quality works or services to rural areas. There are, however, several policies in place to improve transparency about program performance and procurement. Public authorities must disclose their decisions on approved and rejected activities, release information about procurement bids, and publish annual implementation and financial reports.

The Tanzania Urban Local Government Strengthening operation aims to improve institutional performance for urban service delivery in specific Urban Local Government Authorities (ULGAs). The project became effective in February 2013 and will close at the end of 2018. The operation supports Urban Performance Grant system, a new window that is expected to make up about a quarter of the Local Government Development Grant (LGDG) system that awards funding from the central government to Local Government Authorities. This funding comes in the form of Urban Performance Grants (UPGs), which are focused on meeting the investment needs of Tanzania’s secondary cities. The LGDG system is jointly funded by the Government of Tanzania, seven bilateral partners, and the Bank. The government wishes to use the Bank as the initial funding source for the operation, with the possibility of adding additional ULGAs to the project as it develops. It is unclear whether any additional ULGAs could be part of the PforR operation or whether

they would receive funding from other donors. As the project stands now, it is funded entirely by the Bank. The total amount is USD 255 million.

All six DLIs are of the systems type. Institutional and delivery performance accounts for 80 percent of the credit while 20 percent goes towards strengthened management capacity. Most of the DLIs focus on system performance, not on specific policy actions. Only the ULGAs which fully meet minimum access conditions set out in DLI 1 regarding annual institutional performance assessments will be eligible to receive UPG funds and disbursements through the other DLIs. ULGAs that fail to meet the minimum conditions will receive capacity-building support from the Regional Administration and Local Government unit of the Prime Minister's Office (PMORALG).

DLIs 2 through 5 are scalable, while DLI 1 is a pass/fail measure. DLI 6 will only be disbursed once PMORALG sends the Bank task team evidence that it has improved the annual performance assessment and has publicized the changes made to the assessment. Although there does not seem to be a clear link between expected disbursement levels and the estimated costs of achieving the DLIs, UPG funding is assigned to local governments on a per-capita basis. The UPGs are scheduled to increase in value over time to reflect the increasing capacity of the ULGAs.

Only one of the six DLIs (DLI 4) has a baseline; the other five DLIs will be assessed only after the project's launch. A private firm, hired by PMORALG will conduct annual performance assessments for DLIs 1 through 3, and the Bank will review the assessment results, allocation amounts, and timeliness of the UPG disbursements. For DLIs 4 and 5, PMORALG will submit relevant documentation to the Bank, which will then review the documents prior to disbursement. The Controller and Auditor General will conduct a regular program audit. A supplementary audit may be done on the flow and timing of PforR funds to beneficiary institutions. Disbursements will be based on an annual performance assessment and other measures of institutional capacity. Within 30 days of the new fiscal year, PMORALG will be responsible for submitting an implementation report about the previous year's plan. The Bank will then verify the report and determine the disbursement level.

The overall performance risk against the DLIs is "Substantial". If the performance assessments are not conducted on time or are not of high enough quality, the disbursement schedule may be thrown off. Another risk is that disbursements could be delayed by factors

outside of the control of PMORALG and the ULGAs. For example, if the Treasury is slow to process disbursements, ULGAs may not receive funds promptly.

Although the Bank will review the results for each DLI, one might be concerned that PMORALG will hire a firm that could be pressured into producing a more generous assessment of the ULGAs' progress. Concerning DLI 4, which focuses on staffing the ULGAs, one worry is that PMORALG will pressure the ULGAs to produce identical lists of core staff so that the Bank disburses fully after verifying the staff lists from PMORALG and the ULGAs.

The fiduciary risk of the program is rated as "High". Automation of financial management is low, financial statements audited in the past have not been assessed favorably, and staff capacity remains low at the ULGA level. Past procurement plans are often lacking in information and there have been serious delays in evaluating bids and awarding contracts. Nevertheless, the operation will use country system safeguards, including Tanzania's Public Procurement Act, which specifies minimum tendering periods, procurement limits, and bidding processes.

The Environmental and Social System Assessment found large differences between Tanzanian systems and the Bank's requirements for resettlement and compensation. An Environmental and Social Management Manual (ESMM) was developed to address these gaps. The ESMM contains criteria that will be used to exclude certain new landfill and road projects from the project, to analyze induced impacts, and to increase public participation. The Tanzanian threshold for "significant" environmental impacts is much less stringent than that of the Bank. The ESMM tries to bring these definitions closer so that project screening can be done consistently. As with all other PforR operations, sub-projects with Category A impacts will be ineligible for financing. There are provisions to track spending and disclose information about the grants. Criteria regarding accountability and public consultation are included in the annual performance assessments.