the Education Commission

Background Paper

The Learning Generation

# A Global Offer for Learning (GOL): Based on Experiences With Paying for Results

**William Savedoff**
Center for Global Development

# A Global Offer for Learning (GOL): Based on Experiences with Paying for Results

By William Savedoff

Center for Global Development, Washington, DC

October, 2016

**Acknowledgments**

This CGD Policy Paper was prepared by William Savedoff with inputs from Nancy Birdsall, Barbara Bruns, Justin Sandefur and Janeen Madan. It is based on a proposal solicited by the Secretariat of the International Commission on Financing Global Education Opportunities, and gratefully acknowledges the funding which supported development of this paper. The paper benefited from comments, ideas and suggestions from Luis Crouch, David Ferreira, and Rita Perakis. The contents are the sole responsibility of the author.

## Contents

# 1. Introduction

The world has long recognized the importance of education for social well-being and national development. Nevertheless, foreign assistance for improving education in low- and middle-income countries came rather late in the history of foreign aid. Initially, aid agencies focused on supporting educational infrastructure – schools, university libraries, research laboratories. In the 1980s, however, foreign aid to education began to grow with a push for expanding primary schooling. Most countries have made rapid progress in enrollment and completion of primary schooling but another problem has emerged: poor learning outcomes. Consequently, almost all initiatives aimed at improving education in low- and middle-income countries today incorporate attention to improving learning, typically through activities that seek to raise the quality of schooling. Inequities, exclusion of marginal groups, and low secondary enrollment are still problems. But for the most part, enrollment is no longer the main issue. Rather, the low overall quality of primary and secondary education is the chief impediment to learning, social well-being, and productivity.

Though countries have made historically unprecedented progress on enrollment, the pace of *learning* improvements has been remarkably slow. At current rates, it would take Ghana 100 years for its primary completers to have the competencies expected for learners coming out of British schools today. In India, it is unclear whether there is any progress in learning at all (Pritchett 2013). If countries are going to make substantial gains in learning over the next 20 years, it will take more than additional schools, teachers, and textbooks delivered through existing school systems. Rather, it will require reforms, improved practices, and innovation throughout the education system. Changes need to be systemic because even the best classroom innovations cannot thrive without a supportive context. Teaching children well requires more than resources; it also requires systematic feedback and positive reinforcement for good performance.

Yet a key feature of education around the world is its limited capacity to innovate at the system level (IDB 1996; Pritchett 2013). Education has been the subject of numerous innovations over the years in terms of curricular development, teaching methods, classroom management, child-centered strategies, computer-assisted learning, etc. However, the basic structure of most education systems[1] around the world is remarkably similar: the public sector is responsible for most of the funding for primary and secondary education and is also responsible for most provision. Systems vary with regard to structure and governance, whether centralized or decentralized, and whether funds are raised locally or nationally. Some systems incorporate large non-profit segments which still receive public funding (e.g., Chile's voucher schools, Fé y Alegría schools in many Latin American countries). These systems sustain themselves (in the sense of obtaining funding to continue operation) by demonstrating the need for inputs, not by demonstrating their responsiveness to feedback on performance as indicated by student learning. Without greater attention to learning and without the ability to respond to feedback with experimentation and innovation, most education systems in the world will not substantially accelerate student learning.

[1] This paper focuses on primary and secondary levels of the "education system." Other segments of education – such as early childhood development, tertiary, job training, and continuing learning – face other challenges.

**Box 1: A Summary of the Global Offer for Learning (GOL)**

*A detailed explanation of GOL can be found in section 5.*

**The basic idea**: The international community establishes a US$1 billion fund which will pay eligible countries an Assessment Award and an Achievement Award. *The Assessment Award* pays US$1.5 million each year that a country applies a qualified test to assess learning and publishes results (for up to 7 years). *The Achievement Award* pays US$4 for each 9-year-old who has learned basic skills (for up to US$2 million each year over the same time period). Participation is completely voluntary; the entire US$1 billion will be disbursed by 2027; and procedures will be clear, objective, and simple.

**Advantages:** The Global Offer for Learning (GOL) is structured to draw international attention to learning; to reward improvements in the education system as a whole; to create feedback loops that stimulate domestic policy innovation; and to encourage full payout and financial sustainability through its innovative pooling and disbursement mechanism.

**Design features:** The organization managing GOL will sign agreements with any country that (1) has income below a certain threshold, (2) has a qualified test, and (3) agrees to public disclosure of test scores. A qualified test uses internationally-recognized methodologies (Item Response Theory) to generate nationally representative estimates of learning by school-age children that are equivalent over time and can be reliably benchmarked to other countries.

**Funding mechanism**: Donors would not contribute directly to the fund but would, instead, sign binding agreements with a Financial Intermediary (FI) to reimburse the FI after Assessment Awards and Achievement Awards are disbursed. The FI would be an organization (such as the World Bank) which has sufficient liquidity and borrowing capacity to pay participating countries when payments are due. Donors would program funds through their normal budgetary process in the year that they need to fulfill their obligation to the FI.

This paper analyzes how one particular aid instrument can facilitate and encourage system-level innovation: results payments to governments.[2] Other kinds of results payments – to firms, schools, teachers, or individuals – can only have system-level impacts if they expand at sufficient scale. By contrast, paying *governments* for results holds the possibility of real system-level change. The Global Offer for Learning (GOL) proposed in this paper provides funding to governments to assess and improve learning in their countries in a way that explicitly opens opportunities for them to undertake such system-level change (see Box 1).

The next section of this paper characterizes the wide range of programs that pay for results and identifies the subset of programs that are relevant to this proposal: those that pay

[2] The idea of paying governments for high-level outcome measures and implementing the approach as a global offer to developing countries is explained in Barder and Birdsall 2006 and then extended and detailed in Birdsall and Savedoff 2010.

governments for outputs or outcomes. The third section discusses the experiences of Multilateral Development Banks (MDBs) and bilateral agencies with programs that pay for results and the implications for designing good programs. The fourth section describes specific experiences in paying for results in education. The fifth section details the proposal to create a Global Offer for Learning (GOL) that would reward national governments for progress in learning outcomes while ensuring they have the information and the flexibility to pursue their own system-level reforms and innovations of their own design.

## 2. Why Pay for Results and How?

All foreign aid pays for something—domestic inputs that are delivered in-kind (drugs, food, training); foreign inputs that are purchased abroad or obtained locally (construction, supplies); outputs (completion of water connections, completion of a road, completion of an electricity power plant); or outcomes (improved health, job placement, higher incomes) (see Table 1). The choice of what foreign aid pays for affects (1) what is measured (2) how progress is reported (3) who is accountable to whom for what and (4) what kinds of flexibility and discretion recipients have over the use of funds.

**Table 1: Examples of What Aid "Buys" by Results Level**

| Result Level | Examples |
|---|---|
| Inputs (in-kind) | Food, medicines, training |
| Inputs (purchased with aid funds) | Construction, vehicles, local services, supplies |
| Activities and processes | Completing a planning process, completing training courses, establishing a new institution |
| Policies | Enacting a regulation, changing a law |
| Output | Water connections, kilometers of road, construction of a generating plant, student completion |
| Outcomes | Improved health, job placement, higher incomes |

Conventional aid programs largely finance the distribution of in-kind goods and services (including technical assistance), the purchase of inputs in recipient countries, or support for policy changes. If conventional aid were roundly successful, aid agencies would not be looking for alternative instruments. A prominent alternative is to pay for results (rather than inputs). The current trend of experimenting with result-based payments represents an effort to complement conventional aid by designing programs that additionally give attention to outputs and outcomes.

Proponents of paying for results make a number of different arguments for how it might perform better than conventional aid. Four distinct theories can be found in aid agency literature (Perakis and Savedoff 2015):

- **Pecuniary interests** will lead recipients to put out greater effort or shift their priorities away from other interests and toward the results that are being compensated.
- **Attention** of managers and politicians will be focused on results, leading to better decisions and more effective implementation.
- **Public dissemination of results will improve accountability to constituents**.
- **Recipients will have more discretion**, by delinking payments from inputs, giving them greater ownership and the ability to achieve progress in more efficient and innovative ways.[3]

Of these four reasons, aid agencies are most comfortable with the idea that paying for results will draw attention to program goals and can therefore be added on top of conventional aid approaches to diagnosis, planning, monitoring and evaluation. By contrast, greater recipient discretion is resisted by most agencies for political and bureaucratic reasons and has been utilized only rarely.

These different theories of action are not exclusive but the way a pay-for-results program is designed will determine whether it works more through one or the other of these channels. Note that fully characterizing a conventional aid program for its impact on measurement, reporting, accountability and flexibility requires a large number of details related to the exact triggers for disbursements (e.g., expenditure statements or full invoices), liquidity accounts, project units, fiduciary controls, requirements for review of bidding documents and no-objections, etc. Similarly, it is impossible to fully characterize results payments without important details related to such things as the unit of measurement, amounts, and portion of funding up-front. Perakis and Savedoff (2015) provide a list that distinguishes 13 features of results-payment programs:[4]

- **Time frame**: The time frame of an agreement affects such things as: the amount of time funders and recipients have to understand how this type of agreement works; the ability to try, test, and adapt strategies; choices regarding approaches which differ by pace, efficiency, and sustainability; and the required level of precision in measuring results.

- **Renewability**: Renewable and non-renewable agreements will differ in terms of the incentives they provide for collaborating, sharing information, and enforcement.

- **Transparency**: The transparency of an agreement has implications for design as well as accountability. Agreements that are structured in ways that facilitate public dissemination are necessarily easier for officials themselves to understand and

[3] Clist and Verschoor (2014) argue that if the recipient has ownership and wants the same results as the donor, then a lump sum payment would be preferable to a payment for results. This would be the case if the donor had no need to justify its spending. In this regard, paying for results solves an accountability problem facing donors which is to show what their foreign assistance funds achieve.

[4] Adapted from Appendix A of Perakis and Savedoff 2015.

manage. The degree of transparency also introduces additional sources of feedback and accountability from constituents and peers which may affect the design and performance of programs.

- **Recipient** (see Figures 1 and 2): Performance-based funding agreements vary in terms of who gets paid. Recipients may be individuals, households, community groups, non-profit organizations, for-profit firms, public service providers, subnational governments, or national governments.

- **Results level** (see Table 1, Figures 1 and 2): Programs described as paying for results vary regarding the point in the "results chain" for which they are disbursing funds. Thus, programs may pay upon proof of purchasing inputs, completing certain tasks or activities, enacting regulations or laws, establishing procedures, producing goods or services, or contributing to outcomes like health status, employment, or income.

- **Result complexity**: Agreements differ over the complexity of the results for which they pay. For example, programs may focus on only one indicator or on many. They may also define directly observable indicators or indirect ones requiring statistical techniques to define a counterfactual.

- **Recipient discretion**: Agreements can give recipients varying degrees of authority to decide for themselves how they will achieve results and how they will spend payments. The level of recipient discretion has implications for the recipient's level of commitment, flexibility in implementation, ability to innovate, and responsiveness.

- **Payment amount**: The payment amount may be significant in relation to unit costs or relative to other funding sources. Some agreements fully reimburse unit costs and even provide bonuses, while others only subsidize costs or provide rewards and prizes that are a small share of costs. Some agreements represent potential payouts that are large relative to domestic budgets or other foreign aid opportunities while others are marginal.

- **Up-front payments**: Up-front payments related to preparation for measuring results are completely consistent with ex-post performance payments. However, up-front payments for activities or inputs considered "necessary for achieving progress" run counter to the logic of results-payments. The more up-front funding provided as part of the program, the more it will suffer from the criticisms associated with conventional aid approaches.

- **Payment function**: Payments can be made in proportion to progress, in which case they are incremental. Alternatively, payments may be made in tranches for meeting a target or for passing a predetermined threshold. Sometimes payments are triggered

by whatever results are achieved at a certain date, while in other cases the date of the payment depends upon when targets are achieved.

- **Credibility**: An agreement's credibility will influence whether the parties pay attention to it. Agreements which lack credibility are unlikely to motivate change. One important element of credibility is the verification of results through the use of independent information.

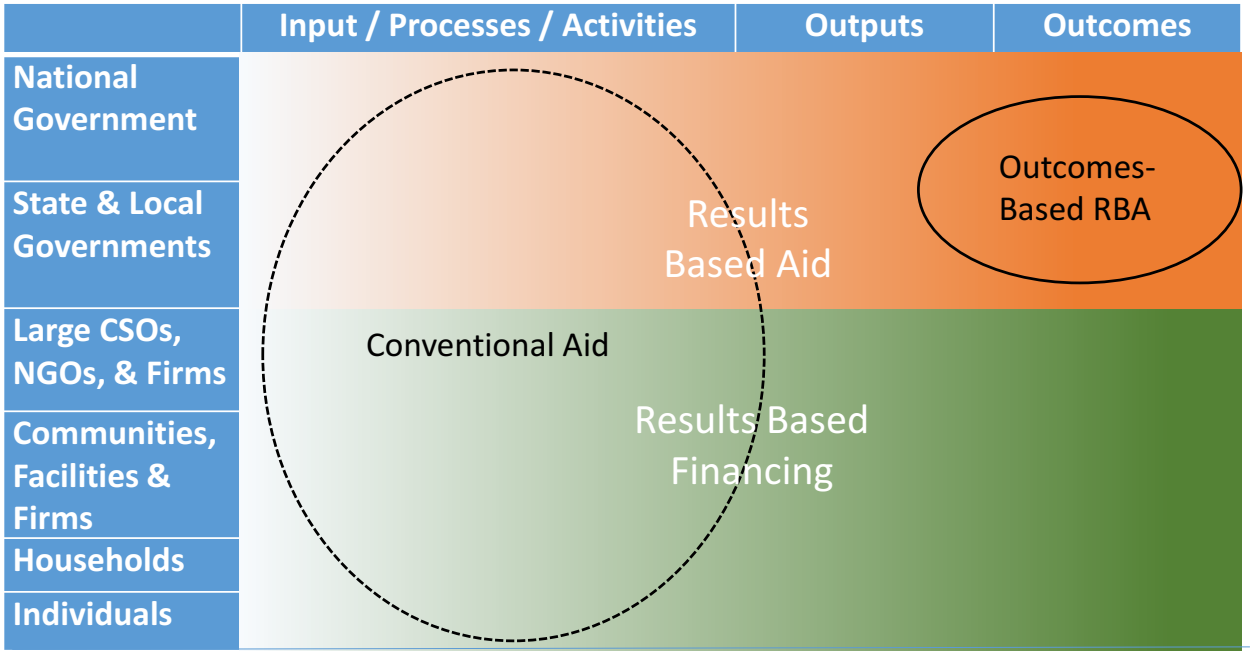
Of these features, the *recipient* and the *results level* are particularly important for understanding, designing, and evaluating results-payments. Programs can be distinguished visually in a figure that places recipients on the vertical axis and results levels on the horizontal axis (see Figure 1 and Figure 2).

First, results-based payments have to be designed differently and will function differently depending on **who receives the funds**. This note follows the convention used by Britain's Department for International Development (DFID) to call programs that pay governments "Results-Based Aid" (RBA), pictured in the top of the diagram, and programs that pay other agents – individuals, households, public facilities, communities, non-profit entities, or private businesses – "Results Based Financing" (RBF).[5] Specifying who will receive payments is important because different models of behavior and change must be used to design programs for entities that vary so substantially in terms of resources, decision-making procedures, accountability, and interests.

Second, programs that use results-based payments will vary significantly by **the character of "result"** that they pay for. Results are commonly understood as the final goal of a program, which would normally be described as its expected "outcome" or even "impact". However, in practice, aid agencies have defined results to mean many things, including purchasing inputs, undertaking activities, completing processes, or adopting policies which are primarily means to ends. The dividing line along the typical "results-chain" from inputs, activities, and processes to outputs, outcomes and impact is not exact. Nevertheless, programs that disburse funds upon proof of purchasing an input (e.g., textbooks) are largely indistinguishable from conventional aid programs. Therefore, if results-payments are to be an innovative mechanism, it must mean analyzing, designing and implementing programs which define results as something that is closer to an output or, preferably, an outcome. This requires distinguishing programs that pay upon passage of a new credit law from those which pay upon verifying that poor people have expanded access to credit; those that pay for completion of teacher training courses from those which pay for the number of 9-year-olds who can read; or those that pay for construction of a power plant from those which pay when electricity outages are reduced.

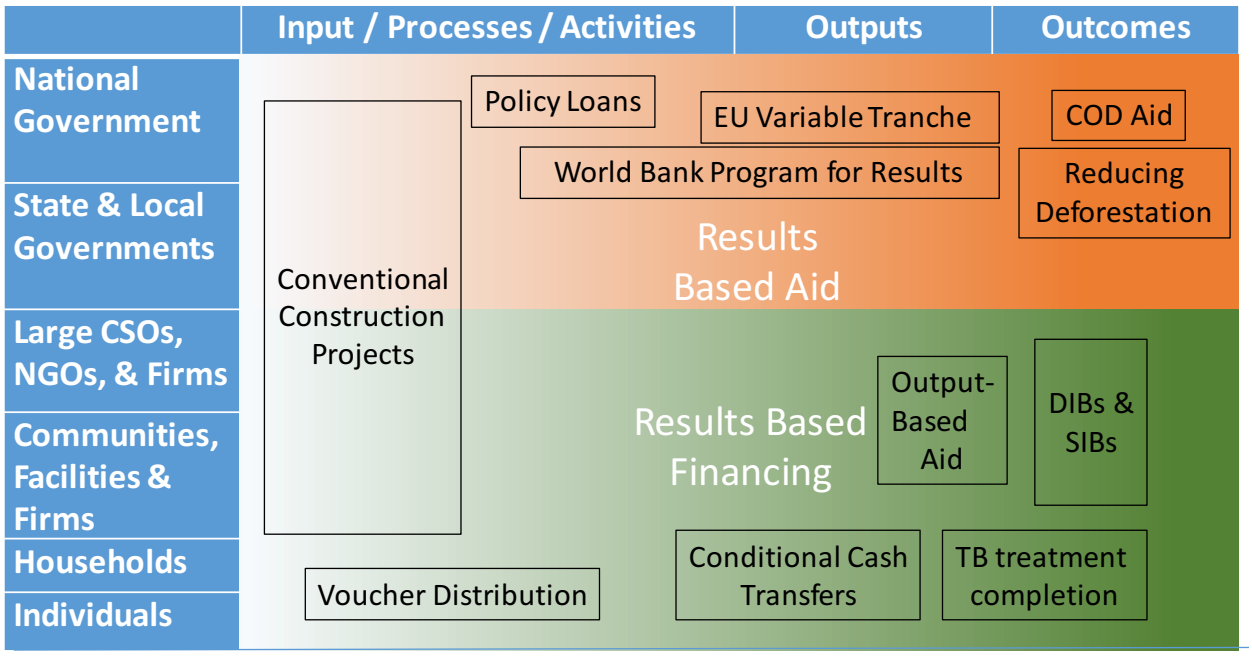
[5] See for example: Pearson 2011; Musgrove 2010; Savedoff 2011. For a review of results based aid and results based financing, see Pearson, Johnson and Ellison 2010 and Janus 2014.

**Figure 1: Performance programs by recipient type and results level**




*Source*: Adapted from Perakis and Savedoff 2015.

**Figure 2: Performance programs classified by recipient type and results level**




*Source*: Adapted from Perakis and Savedoff 2015.
*Notes:* CSOs = Civil Society Organizations; NGOs = Non-Governmental Organizations; DIBs = Development Impact Bonds; SIBs = Social Impact Bonds.

Thus, any discussion of paying for results has to identify (1) who is being paid and (2) what they are being paid for. MDBs, bilateral agencies, and developing country governments have more experiences with paying for inputs and activities than outputs and outcomes. They also have more experience with paying for results achieved by firms, communities, household, and individuals than by national, state, or local governments.

**Implications for a Global Offer for Learning (GOL):** Education is delivered locally, but testing is typically a national or state-level function. Furthermore, local delivery of education is conditioned by state and/or national agencies that define the content of education, eligibility for educational institutions, licensing of teachers, and, frequently, personnel management and contracting. Therefore, paying governments for progress in applying tests and promoting system-level innovations to improve learning would introduce a new instrument for promoting progress. To complement existing conventional approaches to introducing national assessments and improving learning, a new pay for results program could be offered to pay governments for outputs (national assessments completed) and for outcomes (children who demonstrate competency in particular skills).

## 3. Experiences with Paying for Results in MDBs and Bilateral Agencies

MDBs and bilateral agencies have worked for many years on initiatives to link payments to performance. When applied to aiding governments, these initiatives take one of three paths. *Policy lending* rewards countries for reforms with payments that are triggered by enacting particular laws, adopting new procedures or creating new institutions. *Budget and sector support programs* operate through policy dialogues informed by a large number of "soft" indicators and a great deal of discretion by funders to insist upon or waive performance measures. *Investment programs* have found ways to pay for inputs or outputs that were countable and had well-defined cost-structures. Agencies rarely, if ever, tied payments to outcomes like student learning, declines in mortality, or lowering transportation costs.

The major objections to paying for results have been addressed many times and, in fact, most agencies have no fundamental legal obstacles to doing so. For example, the World Bank formally adopted a policy for its Program for Results (PforR) instrument in 2012 but the underlying logic had already been applied in many conventional loans (e.g., Plan Nacer paid for provincial level expansions of health insurance in Argentina, see Cortez and Romero 2013). As another example, the United States Agency for International Development (USAID) has "Fixed Amount Reimbursement Agreements" which allow it to pay governments a fixed amount for infrastructure expansion, like schools, as long as unit costs can be verified.[6] USAID has gone further in its programmatic aid to governments, as in its current Afghanistan New Development Partnership (NDP) which will disburse when the government implements a Citizen's Charter and improves its ranking on the Doing Business Index, among other indicators. The European Union has more than 10 years of experience

[6]For a recent example, see Dunning and McGillem, "Country Ownership at USAID: Enabling and Empowering Liberia's Ministry of Health." CGD Blog 4/28/2016.

with variable tranches in its budget support programs, which disburse in proportion to progress on a number of indicators. Norway has demonstrated that it can offer payments for an outcome – reduced deforestation – on a large scale with Brazil and that it can credibly withhold paying when performance is lacking – as has occurred in Indonesia (see Box 2).

DFID has piloted a number of programs paying for results. Two of them are paying for outputs or outcomes in education at the national level. Ethiopia and Britain signed an agreement in 2012 for a three-year program that paid Ethiopia for increases in the number of students above a baseline who take or pass a national exam at the end of lower secondary school. The disbursements ranged from £50 to £100 for each additional child who sat the lower secondary school exam, with higher payments for girls than boys and for children in disadvantaged regions than in more prosperous regions. DFID committed a total of £30 million to the program over three years (see Box 3).

Another program signed between Rwanda and Britain in 2012 has a similar structure. It pays the Rwandan government for each additional child taking exams at the end of primary school and the end of lower secondary school, with the unit payments ranging from £10 to £100 depending on the schooling level and baseline. The program also pays £50 for every additional teacher competent to teach in English—calculated as the difference between a baseline assessment in 2012 and a follow-up assessment in 2014. DFID committed a total of £9 million to the program over three years.

Britain's experience with these RBA programs was the subject of two concurrent evaluations (Upper Quartile 2015; Cambridge Education 2015). DFID staff worked hard to make sure the program would be "hands off," yet this aspect of the program took time to understand such that the recipient government could take advantage of its autonomy. Both programs were also embedded in larger education sector programs—involving intensive planning, monitoring and engagement by numerous donors—which constrained flexibility and made it difficult to assess the effectiveness of the pay-for-results mechanism.

**Box 2: Norway Pays in Proportion to Reduced Deforestation**

In March 2004, the Brazilian government initiated a range of policies and enforcement actions (under the Action Plan for Preservation and Control of Deforestation in the Legal Amazon) that brought sharp reductions in the rate of deforestation. Subsequently, in 2008, Brazil and Norway completed an agreement under which Norway would contribute to a Brazilian environmental fund if the annual rate of greenhouse gas emissions from deforestation were brought below the average rate of the 1996–2005 period—a "performance-based payment" agreement. Norway pledged up to US$1 billion for this purpose. Under the Agreement, donations go to the Amazon Fund (Fundo Amazônia), which had been created for this purpose at the Banco Nacional de Desenvolvimento Econômico e Social (BNDES) to finance projects that contribute to reducing deforestation and promoting sustainable development in the Amazon. Up to 20 percent of the funds could also be applied in other biomes, whether in Brazil or other tropical countries.

The Agreement with Brazil was the first of many bilateral performance-based agreements established by Norway with forest-rich countries in the context of broader international negotiations and experimentation with Reducing Emissions from Deforestation and Forest Degradation (REDD+). In December 2013, Brazil and Norway expanded and extended the Agreement to December 2021.

Source: Adapted from Birdsall, Savedoff, and Seymour 2014.

**Box 3: Britain Rewards Ethiopia for Increasing Secondary School Completion**

Britain has been providing substantial financial and technical assistance to Ethiopia since the 1990s. Much of this support has been channeled through sector budget support programs and multilateral institutions like the World Bank. In 2012, DFID launched a "Results Based Aid" (RBA) program to improve secondary education in Ethiopia by rewarding the government for increases in the number of students above a baseline that take or pass a national exam at the end of lower secondary school (grade 10). The commitment was to pay out up to £10 million each year for three years. The program paid more for girls than boys, and more for children in poorer areas than richer ones. Payments were either £85 or £100 for each additional girl who sits or passes the grade 10 exam, depending on her location. For additional boys, payments are either £50 or £75 depending on the region. Funds were transferred directly to the Ethiopian Ministry of Education, and were additional to existing support that DFID provided to the education sector.

After the first year of the project, the independent verification confirmed that there were 6,316 additional sitters (3,326 girls) and 4,383 additional passers (1,512 girls) of the grade 10 exam, with almost all of these gains occurring in the emerging regions of Ethiopia. The associated reward payment was about £900,000, reflecting that some positive results had been achieved, but was far below the £10 million ceiling established for the project's first year. In 2013, the following year, 38,490 more students sat the grade 10 exam and 37,669 more students passed. These figures resulted in a payment of approximately £5.6 million. In the final year, DFID disbursed almost £9 million for a net increase of 104,770 students who sat the exam and 16,155 who passed.

Experience with the new aid modality over the first two years led Ethiopian government officials and DFID staff to a fuller understanding of the RBA program's principles. In the first year, Ethiopian government officials did not seem to fully absorb the aspects of the RBA program that made it different from other types of aid programs. For example, the government initially requested DFID to provide guidelines on how to spend funds as if it were a conventional DFID grant. In the second year, the Ethiopian government developed a plan for allocating funds to regions and, though it consulted with DFID, the decision was ultimately the government's own. In its third year, people at the regional and school levels became aware of the program and the associated opportunity to secure additional flexible funding. Regions developed their own plans for allocating reward funds to schools. In some cases, schools received a considerable amount of discretionary funding from the national-level RBA payments. This allowed them to purchase things they needed but which could not necessarily be obtained under other aid programs.

DFID commissioned an evaluation of the program that argued the educational outcomes could not be attributed to the Ethiopian government's actions and therefore could not be attributed to the RBA program. While the evaluation acknowledged that conventional approaches to assessing attribution might not be applicable, it still contributed to a decision by DFID not to continue the program. As a result, we have lost an opportunity to learn whether a longer term program might have significantly altered the aid relationship and fostered better national policy toward education.

Sources: Perakis and Savedoff 2015; Cambridge Education 2015.

Critics of paying for results tend to raise three questions about the approach:

- Where will up-front money come from?
- What about corruption?
- Isn't it unfair to punish countries for things that affect outcomes which are outside their control?

**Where will up-front money come from?** This question assumes that the foreign aid will cover the full cost of achieving the outcome and that countries are not already committed to the goals with their own public revenues and any resources they can mobilize by increasing efficiency. Furthermore, countries are not starting from a blank slate. In addition to their own domestic revenues, most countries have access to a range of international programs either as grants from bilateral agencies or loans from MDBs. Some even have access to private capital markets to finance investments until results payments begin to be paid. Once the first period is concluded and initial results payments are disbursed, countries can use those funds to accelerate further progress since they are not tied to any particular plan or budget category. There may be cases where up-front investment requirements are so large relative to the existing resource base that initial funding is required. But this needs to be assessed on a case by case basis and is not an inherent obstacle to paying for results.

**What about corruption?** This question assumes—in contrast to the preceding one—that recipient governments can achieve outcomes without spending money. When funders pay for the achievement of an output or outcome, they are essentially reimbursing recipients for the effort expended to achieve those results. Corruption that diverts spending from the activities necessary to achieve results will necessarily reduce the amount paid under the program. By contrast, corruption that diverts funds from conventional programs can easily lead to an increase in aid flows when faced by lack of results. Corruption can be a problem in programs that pay for results when the indicators of achievement are manipulated or when windfalls generate payments without requiring much effort. Well-designed programs that pay for results minimize these problems by independently verifying results and by choosing indicators which are mostly responsive to government action (Kenny and Savedoff 2013).

**Isn't it unfair to punish countries for things that affect outcomes which are outside their control?** This question assumes that once a program is approved that the recipient has, in some sense, an entitlement to the full amount of money. This is a strange notion for conventional aid but even stranger for projects designed to pay for results. When a country makes progress in terms of educating more children, lowering mortality rates, or increasing access to energy, a payment in proportion to that progress is a smaller or larger reward, not a punishment.[7] The question also presumes that we can never find indicators for outcomes that are closely related to government effort. Paying for results is not the answer for every goal, but for the subset of programs in which outcomes are responsive to government programs, it is certainly an option that can be offered. Well-designed programs will minimize

[7] See Birdsall and Savedoff, "A Critical Moment for COD Aid or The Trouble with Targets," CGD Blog, 2/7/2013.

the degree to which payments are affected by windfalls and shortfalls by choosing indicators which are less affected by factors outside governmental control, by establishing payment formulas which explicitly control for such external factors, and by committing to long terms so that windfalls one year are likely to be offset by shortfalls in other years.

Paying for results is certainly not the right approach for every aid program. But as demonstrated by these responses, the scope for paying for results is reasonably large and responses exist to address most concerns. Therefore, today, aid agencies have three options for how they approach recipient governments. The first option treats recipients as wards of rich countries and imposes particular forms of management, public institutions, and project design. The second option treats low- and middle-income countries as partners but assumes that they value intensive support and engagement from funding agencies and are willing to cede some level of their autonomy. The third option, which remains very rare, treats low- and middle-income countries as active and capable agents of their own development, offering them more choices. Under this third option, choosing whether to use conventional aid approaches tied to specific plans and inputs or less conventional aid approaches paying for results is in the hands of the recipient governments themselves. Yet this latter option is strictly limited for most countries.

So if paying for results offers the aid community an instrument that allows recipient governments to solve their own problems in new ways and there are few obstacles, why isn't this instrument being offered more widely? The two main reasons that agencies have been reticent to pay governments for outcomes is that the approach questions their primacy in directing development assistance and exposes them to unfamiliar accountability risks.

First, foreign assistance has evolved on the premise that deep engagement with developing countries is necessary for those countries to understand their own problems, develop plans, and take action. A large literature has demonstrated that country ownership (i.e., commitment and responsibility) is the most critical factor for any program's success. Yet aid agencies continue to insist on diagnostics, programming, planning, and supervision that may undermine this very ownership. In some cases, the partnering of aid agencies with recipients is valued by the recipients—either for technical expertise, external political support, or the creation of financial commitments to completing a particular course of action. But in other cases, recipients are frustrated by the degree of agency involvement and prefer greater scope of action. With a results payment, committed countries that prefer this avenue of autonomy can pursue the goal and obtain funding; while those that prefer conventional partnering can continue with existing arrangements.

A second source of resistance is the fear of exposure to fiduciary, environmental or social risks.[8] The conventional approach to such risks has been to insist on detailed implementation plans which can then be subjected to risk assessments. While this is compatible with conventional aid programs and with results payments that are "added on"

[8] When the World Bank approved its PforR modality in 2012, most of the associated policy documents focused on how these "safeguard" issues would be handled, with almost no discussion of the expected benefits of the new modality or the characteristics of its indicators and payments.

to conventional projects, it is not compatible with results payments that are meant to encourage country ownership and increase recipient autonomy. This is an understandable response to the visibility of these risks compared to the invisibility of results failures. However, it interferes in adopting an approach that would actually open space for country ownership, responsibility and innovation.

As a result of these concerns, aid agencies have an easier time paying for results defined as inputs, activities, and policy changes than outputs or outcomes. This can be seen in the sectors where paying for results is most common such as water and sanitation, energy, and health. It is also apparent in the way agency staff talk about the success of these modalities.

Of the four theories of change described above, most agency staff *and counterparts* agree that the pay-for-results programs **improve implementation and management** (see section 2). People who see the benefits of drawing attention to results and using it to improve intensive engagement between aid agencies and recipient countries view this positively. This approach has been likened to the positive role private creditors can play in improving corporate governance in countries like Germany and Japan (Gelb et al 2016). Thus, paying for results can be useful even when it is linked to processes and activities if the recipient country values that level of engagement and if the production function is relatively well-known.

By contrast, the instances in which recipients would prefer a "hands off" approach are not being exploited. Very few programs are operating through any of the other three theories listed above: creating an effective **pecuniary motive** for governments, increasing **accountability** or establishing **country ownership** and discretion (Perakis and Savedoff 2015; World Bank 2015). For those who, in particular, see country ownership and discretion as critical to addressing complex forms of learning and institutional development, the limited use of paying for results with a "hands off" approach is a serious problem.

Reviews of results payments in MDBs and bilateral agencies demonstrate that most programs "add" results requirements on top of their conventional program requirements. Most pay-for-results programs require conventional preparation and monitoring such as: diagnostics, pre-approved plans, implementation schedules, monitoring & evaluation of activities, and reimbursements based on statements of expenditures or receipts. They also typically include pre-clearance for or plans to address fiduciary, environmental and social risks—something which cannot be done without detailed plans. In cases where programs pay for results, a portion of disbursements is linked to outputs or outcomes which may or may not be independently verified. As a result of layering the results payment on top of conventional requirements, most evaluations of pay-for-results programs conclude that the approach is more complicated and difficult than conventional aid. Few programs are designed to dispense with the conventional requirements which are unnecessary for or antithetical to the logic of paying for results.

**Implications for a Global Offer for Learning (GOL)**: Current programs that support national assessments and improving learning focus on the technical side of education but offer recipient governments little opportunity to solve complex political and institutional

problems in their own ways. It is this side of the "production function" about which the international community lacks knowledge. Furthermore, although the development of a robust national assessment program requires sophisticated technical expertise that usually must be recruited internationally, the scale of upfront funding required, relative to existing education budgets and external funding, is not large. Paying for results in a way that offers recipient governments autonomy while creating objective signals regarding performance would open space to address local political constraints while giving funders the assurance that disbursements are related to real progress. For countries which are committed and want greater freedom in choosing their own education policies, a pay-for-results program can make outcomes and progress more visible and, arguably, make sustained efforts more likely.

## 4. Paying for Results in Education

Foreign aid for education has fewer experiences with paying for results than health, infrastructure or environment, though the number of programs has increased in recent years. The programs that are included in various surveys differ in terms of their focus and definitions.

Results for Development (R4D 2016b) has produced the most complete survey of education projects that pay for results in low- and middle-income countries and it finds only 20 programs. Of these, two are operations that pay governments in proportion to a result; the rest involve payments to schools, training institutions, NGOs, teachers, or students. Of the two government programs, one pays for something close to an educational outcome: DFID's agreement with Ethiopia pays a fixed amount for each student who completes lower secondary school and takes a test (see Box 2).[9] The other government program disburses US$40 million to a Pakistani social protection agency for meeting a range of performance indicators, most of which are process or activity indicators.

The World Bank has 37 PforR operations with 23 governments of which only two (Tanzania and India-Bihar) are specifically about education.[10] However, education is partly addressed in four general public services programs or state modernization operations in Burkina Faso, Ethiopia, Morocco, and Mozambique.

The first of the two World Bank education programs is "Big Results Now in Education" for Tanzania, approved in 2014. It is a US$416 million operation of which DFID and SIDA are providing US$130 million as grants and the World Bank is lending US$122 million as PforR. The operation lists six Disbursement Linked Indicators (DLIs) which follow conventional program logic. The first DLI involves payments for completing "foundational activities." Other DLIs pay for producing reports, deploying teachers, and distributing grants to schools. The sixth DLI is for improvements in student learning as demonstrated by a reading assessment. If student learning increases sufficiently, the World Bank would disburse

[9] The R4D database did not include a similar DFID program in Rwanda which also pays for results in education, see Upper Quartile 2015.

[10] As of March 31, 2016.

a maximum of US$16 million to Tanzania, representing 4 percent of the entire operation or 13 percent of the World Bank loan. To our knowledge, this sixth DLI is the only true outcome indicator in any education program of the World Bank or any MDB.

The second education program is "Enhancing Teacher Effectiveness in Bihar," approved in 2015. It is a US$357 million operation of which the World Bank is lending the Indian Government US$250 million; US$225 million of the loan is PforR (the rest is regular technical assistance). The program ostensibly has six DLIs but these are actually groupings of a number of indicators which are called "results," all of which are inputs to the education process. The DLI for "Infrastructure" disburses in relation to the number of schools that are properly equipped. "Quality Improvement" is focused on the qualifications of teachers; it disburses on the basis of indicators related to preparation of training materials, commissioning a study, and implementing a teacher learning program. The final portion of this DLI is triggered by the number of teachers passing an exam and enrolling in professional development.

Of the other four PforR programs, only one pays for something that is like an output or outcome. The Ethiopian "Enhancing Shared Prosperity through Equitable Services" operation involves a US$600 million loan for a wide range of activities related to improving public sector function and service delivery. One component provides up to US$60 million for increasing net enrollment in grades 5-8 outside of Addis Ababa. Over the three-year project, Ethiopia will receive disbursements of up to US$20 million each year that they increase this net enrollment by 50,000 students.[11]

The remaining PforR programs pay for the share of new primary school teachers who are officially appointed (Burkina Faso), the share of girls residing in educational dormitories who progress to the next grade (Morocco), and the share of primary schools that comply with standards for transparency and accountability (Mozambique).

Beyond these programs which pay governments for results, R4D's database (R4D 2016b) lists 18 programs that are *not* paid to governments. Twelve of these programs pay for a mix of indicators which are poorly defined or not specified. However, six are clearly paying for outputs or outcomes like students passing tests, trainees getting jobs, and enrollment. In general, these programs can be roughly grouped into five categories (see Table 2):

> **Enrollment or attendance**. Five programs pay an educational institution for student enrollment or attendance. Four of these are government programs that pay a voucher or a direct subsidy to private schools to enroll children. The fifth is a government program in the Democratic Republic of the Congo that provides (unspecified) awards to public schools for improved enrollment.

[11] For a sense of scale, net enrollment in grades 5–8 outside of Addis Ababa grew from 3,979,791 in 2012/13 to 4,127,078 in 2013/2014.

**Passing tests**. Three programs pay providers for students who pass tests. Two of these pay private providers and one is a pilot experience which pays public school teachers on the basis of student learning (in Tanzania).

**Post-training employment**. Two programs pay providers of vocational training programs at least in part on the basis of post-training employment.

**Broad performance measures**. Four programs pay providers rewards on the basis of broadly-defined performance measures such as improvements in standardized test scores, teacher attendance, enrollment, and retention rates.

**Unspecified**. The remaining eight programs did not specify either the amounts that would be paid or the indicators on which payments would be made.

**Table 2: Pay-for-Results Projects and Corresponding Indicators from R4D Database**

| Indicator category | Projects |
|---|---|
| Enrollment or attendance | Bogotá Concession Schools Program, Colombia; Programa de Ampliación de Cobertura de la Educación Secundaria (PACES), Colombia; Stimulating School Performance, DRC; Balochistan Education Support Project, Pakistan; Punjab Education Foundation – Foundation Assisted Schools Program, Pakistan |
| Passing tests | Secondary Education Quality and Access Improvement, Bangladesh; Chile Lifelong Learning and Training Project, Chile; KiuFunza – Thirst to Learn, Tanzania |
| Post-training employment | Employment Fund Nepal, Nepal; Enhanced Vocational Education and Training Project, Nepal |
| Broad performance measures | Female Secondary School Assistance Project II, Bangladesh; Skills for Employment Project, Nepal; Lagos Eko Secondary Education Support Project, Nigeria; Big Results Now in Education (BRNEd), Tanzania |
| Unspecified | Bangladesh Education Development Programme (BEDP), Bangladesh; Female Secondary School Assistance Project I, Bangladesh; Secondary Education Finance Reform, Belize; Improving the Quality of Education in Bubanza, Burundi; Education and Care for Children, Central African Republic; Contracting Primary Schools for Performance, Malawi; Girls' Education Challenge Fund (GEC), Global; Skills Development Project, Nepal |

*Source*: Author's categorization of projects listed in the R4D 2016b.

Thus, experimentation with paying for results in education is still quite limited and often poorly specified. It is less likely to be found in public primary and secondary schooling than in other parts of the education system. As R4D 2016a notes, it is more commonly found in:

- Private for-profit or non-profit schooling (primary and secondary),
- Early childhood development / pre-school,
- Vocational training,
- University level, and
- Special education

Initiatives which face the least resistance are those that provide scholarships (merit-based or otherwise) and conditional cash transfers to families whose children are attending school. Those which face more skepticism but are increasingly being tried include performance pay for teachers, school reward systems and a mix of Social Impact Bonds, Development Impact Bonds and Output-Based Aid. By contrast, current trends suggest that programs that pay governments for improvements in performance are likely to remain rare.

If the world thinks it needs to go beyond conventional aid approaches and pay for results to encourage the significant reforms needed to accelerate progress in learning outcomes, then it will have to design and implement programs that depart more clearly from conventional aid. Such programs will push toward the end of the results chain (i.e., outputs and outcomes), insist on independent verification, drop pre-approved planning, and manage risks differently.

MDBs in particular could create a real alternative to conventional aid by offering their clients education programs that pay for outputs and outcomes and which eliminate the transaction costs associated with preapproval, diagnostics, planning and monitoring. However, the MDBs are constrained by their governing principles and policies, particularly with regard to fiduciary, environmental, and social safeguards. Addressing these concerns *ex ante* makes sense for programs that predefine a program's processes and activities. However, the need for such *ex ante* checks is unclear for programs that hold few risks and which are intended to build more fully on country ownership. As a result, innovative programs are subject to delays and alterations that undermine their effectiveness—as demonstrated by experiences with the Fast Track Initiative (Bermingham 2011) and the Forest Carbon Partnership Facility.[12]

**Implications for a Global Offer for Learning (GOL)**: This review demonstrates a range of reasons why MDBs and bilateral agencies find it difficult to offer programs that maximize country ownership and recipient governments' flexibility. Therefore, success is more likely if a program like GOL were administered by a new organization dedicated to supporting learning assessments rather than an existing aid agency or development institution.

[12] The Forest Carbon Partnership Facility (FCPF) Carbon Fund began operating in 2011, yet five years later only 11 countries had completed the first of 8 steps needed to be eligible for results payments related to reducing greenhouse gas emissions (Savedoff 2016).

A new organization could pay for results in a way that offers countries the autonomy and flexibility they need to undertake innovations and systemic reforms. A new organization could minimize transaction costs because it would not have to address a wide range of environmental, social and fiduciary concerns that are not specifically relevant to learning measurement and education. Such an organization would require relatively few staff and could dedicate them to supervising the processes required to verify results, allowing countries to obtain technical assistance from other organizations as they saw fit.

## 5. The Global Offer for Learning (GOL)

This section proposes a program that would support and encourage countries to measure children's skill levels and to improve their learning over time.

As noted earlier, children are enrolling and graduating from schools with serious gaps in their learning. Furthermore, better data on learning outcomes is necessary for improving education system policies and programs, as recognized by the Sustainable Development Goals.

Institutionalizing learning assessments, assuring their financial sustainability and incorporating their results into national education policy requires substantial country ownership and commitment. This global offer maximizes country ownership by minimizing *ex ante* requirements and paying *ex post* for applying assessments and raising learning outcomes.

### The Basic Idea

The international community commits US$1 billion to pay eligible countries an Assessment Award and an Achievement Award. The Assessment Award is a payment of US$1.5 million each year that they apply a qualified test (for up to 7 years). The Achievement Award is a payment of US$4 for each 9-year-old[13] who has learned basic skills (for up to US$2 million each year over the same time period).

Details are provided below but key elements are:

- Participation is completely voluntary.
- The fund represents a firm commitment to disburse the entire US$1 billion by 2027.
- Eligibility criteria and procedures will be clear, objective and simple.

### Benefits

This Global Offer for Learning is structured to:

[13] An earlier version of this paper submitted to the Education Commission in May 2016, recommended a global test for 8-year-olds. This paper recommends testing 9-year-olds because this age: corresponds to grades 2 and 3 in many low-income countries; is the modal age for TIMSS and PIRLS (grade 4) tests; and is a cohort that will be tested when they are 15 years old as part of PISA's 3-year cycle.

- Draw attention to learning
- Promote technically sound learning assessments
- Reward improvements in the education system as a whole
- Signal which countries are making progress
- Provide results feedback for domestic innovation and improvement
- Provide fungible money for addressing any initiative – public or private, education sector or other – that the government deems useful for improving learning
- Assure full payout through an innovative pooling mechanism across many countries
- Encourage financial sustainability by paying *ex-post* and over an extended period of time

## Design Features

**Who will administer it?** The "**Learning Initiative Facility**" described in Birdsall et al 2016 would be an ideal organization for administering GOL. The proposed "Learning Initiative Facility" builds on the UIS Global Alliance to Monitor Learning and would have a 10 to 15-year fixed term, limited staffing, and an independent board elected by funders. The "Learning Initiative Facility" would manage GOL by convening a technical committee with the requisite independence, technical skills, and diversity to assess country eligibility and certify Assessment Awards and Achievement Awards. Alternatively, an existing non-profit organization or international agency could administer the program, but it would require the independence and other characteristics of the "Learning Initiative Facility" (see "implementation" below).

**Who will be eligible?** The "Learning Initiative Facility"[14] will sign agreements with any country which meets three eligibility criteria:

- its median per capita consumption is PPP$10 per day or less,[15]
- it has a qualified test (see below), and
- it agrees to make test scores publicly available and give researchers access to anonymized data.

In order to become eligible, countries without qualified tests can apply for funding and technical assistance (see "Up-front investments" below).

Up to three provinces or states within federated countries can participate in the Global Offer if the national government does not object and if each one's population is 3 million or more.

**What is a qualified test?** A qualified test (1) uses internationally recognized methodologies (2) to generate nationally representative estimates of (3) learning by school-age children that are (4) equivalent over time and (5) that can be reliably compared with test results from

[14] Or whatever organization is entrusted with GOL

[15] Purchasing Power Parity dollars from 2005. Alternatives are discussed later.

other countries. It is also (6) independently certified by an international technical committee (see "Who judges the results?" below).

The ability to compare the test to other countries will be automatic for countries that participate in qualified international assessment programs (e.g., TIMSS, PIRLS, PISA) or the proposed 9-year-old global literacy and numeracy assessment (Birdsall et al 2016). For countries that prefer to rely only on national assessments, comparability can be achieved by including an adequate subset of international "anchor" questions, as explained in Sandefur 2016.

**How does the program pay?** Each participating country could be eligible for an Assessment Award and an Achievement Award, depending on the type of test they use.

- The Assessment Award would pay US$1.5 million per year for up to 7 years. This payment would be disbursed each year after successful application of *any* qualified test and demonstration that anonymized results are publicly available.
- Countries that assess numeracy and literacy of 9-year-olds would be additionally eligible for an Achievement Award that would pay US$4 for each child who has learned basic skills (see definition below). For sample-based tests, the number of 9-year-olds with basic skills would be extrapolated from the point estimate of the nationally representative sample. This payment would be capped at a maximum of US$2 million per year (i.e., for up to 500,000 children) *unless* other participating countries earned less than US$2 million. Any 'unused' funds from countries with fewer than 500,000 children demonstrating basic skills would be available and distributed to countries in which more than 500,000 children demonstrate competency in basic skills.

Note that countries can be eligible for the Assessment Award without participating in the Achievement Award program if their qualified test is assessing different ages, grades, or content. By contrast, any country with a test targeting 9-year-olds that qualifies for the Achievement Award program will necessarily be eligible for the Assessment Award.

**What are basic skills?** For the purposes of this Global Offer only, a technical committee will develop a definition of "basic skills" that is comparable across countries and is *good enough* to measure the foundational literacy and numeracy skills that 9-year-olds are expected to acquire in well-functioning education systems. This definition will *not* be a universal standard of basic skills, something which other agencies and organizations are already developing. Rather this definition would be *only* used for the purposes of calculating the Achievement Award.

**Who judges the results?** The GOL program would convene an international technical committee to assess country eligibility, certify results, and inform arbitration over any eventual disagreements. This committee could be convened using models like The Amazon Fund's technical committee (see "Technical Committee" below).

**How do the donors guarantee payment?** The difficulties for donors of setting aside funding for long periods of time can be resolved by signing agreements with a financial intermediary (FI) that has legal recourse to assure payment. Donors would **not** contribute directly to a GOL fund but would, instead, sign binding agreements to reimburse the FI after the Assessment Awards and Achievement Awards are disbursed. The FI would be an organization (such as the World Bank) which has sufficient liquidity and borrowing capacity to pay participating countries when payments are due. Donors could then program funds through their normal budgetary process in the year that they need to fulfill their obligation to the FI. In this way, donors do not have to make any initial contributions and no new trust fund arrangements are required.

**Up-front investments**. Countries without a qualified national assessment may participate in existing global assessments or regional assessments that meet the criteria above. They may also apply for conventional technical assistance grants to improve their national assessments. They could seek this support from existing multilateral and bilateral agencies, from global assessment programs (e.g., TIMSS, PIRLS, PISA), or from new funds designed for this purpose as described in "Learning Data for Better Policy" (Birdsall et al 2016).

**Suspension**. The agreement could be suspended for technical reasons; for example, if a test is changed and no longer qualifies, or if test data are withheld from researchers. Provisions for arbitration would be included in the agreement for addressing any disagreements over the implementation of the test and its results.

## Further Explanation of the Global Offer for Learning

**Paying ex post and financial sustainability**. This proposal explicitly pays *ex post* for the test and this is a very important feature. It implies that participating countries initially pay for the test from their own budgets and that the spending is 'on-budget' from the start. Some countries might finance these costs with a small technical assistance loan and pay it back with the first year payment. Others might seek grants from existing funds like the Global Partnership for Education (GPE) or the "Learning Initiative Facility" if it is created. Such funders would probably provide funding for only the first year, knowing that countries will have access to Assessment Awards once the test has been successfully administered. By having a 7-year commitment, GOL gives countries time to institutionalize annual or biannual learning assessments in their education system. As an *ex post* payment, the Assessment Award focuses attention on getting the test implemented, analyzed, and published.

**Paying for test passers**. The idea of paying US$4 for each child that passes the basic competency level poses the well-known risk of gaming. However, the amount is set deliberately low to reduce such a temptation. Rather, the funding would be presented as and recognized as "icing on the top" and a useful signal. GOL pays for *all* children who pass the exam rather than some incremental amount. This lowers the stakes on the unit measure and makes it easier to get precise estimates (i.e., it is much easier to measure "how many kids know how to read and write" than "how many *more* kids know how to read and write than last year" with a reasonable sample size).

**Small countries**. The US$2 million cap for Achievement Awards assures small countries that they get their share and aren't crowded out by larger countries. For example, countries like Bolivia or Somalia have 9-year-old cohorts of about 200,000 each year. If they reached 100 percent literacy and numeracy, they would get a full award of US$800,000. This would leave at least US$1,200,000 in the pool of funding to pay Achievement Awards, if needed, in other larger countries.

**Large countries**. Large countries will probably get their entire US$2 million Achievement Award unless they have extremely poor performance. They are likely to get even more because of the inclusion of smaller countries in the pool of participating countries. For example, Tanzania's 9-year-old cohort is about 1 million. They would get the entire US$2 million Achievement Award if 50 percent of the children could read and write. If instead Tanzania achieved 75 percent literacy and numeracy, they might get as much as US$4 million but only if funds were left over from countries that were unable to earn the full US$2 million due to smaller 9-year-old cohorts or poorer performance.

**Assuring disbursement**. To disburse the entire US$1 billion requires that about 40 countries end up participating in the program over an 11-year period. Projections are based on assumptions regarding the number of countries entering and leaving the program (see Table 3). Note that this structure encourages countries to join in the earlier years because later participation would limit the number of years that payouts are possible.

**Table 3: Number of Countries Participating in the Global Offer for Learning**

| | Year 1 | Year 2 | Year 3 | Year 4 | Year 5 | Year 6 | Total partici-pating Countries | Maximum payment per country | Total disbursed |
|---|---|---|---|---|---|---|---|---|---|
| 2017 | 2 | | | | | | 2 | $ 3,500,000 | $ 7,000,000 |
| 2018 | 2 | 5 | | | | | 7 | $ 3,500,000 | $ 24,500,000 |
| 2019 | 2 | 5 | 10 | | | | 17 | $ 3,500,000 | $ 59,500,000 |
| 2020 | 2 | 5 | 10 | 10 | | | 27 | $ 3,500,000 | $ 94,500,000 |
| 2021 | 2 | 5 | 10 | 10 | 8 | | 35 | $ 3,500,000 | $ 122,500,000 |
| 2022 | 2 | 5 | 10 | 10 | 8 | 5 | 40 | $ 3,500,000 | $ 140,000,000 |
| 2023 | 2 | 5 | 10 | 10 | 8 | 5 | 40 | $ 3,500,000 | $ 140,000,000 |
| 2024 | | 5 | 10 | 10 | 8 | 5 | 38 | $ 3,500,000 | $ 133,000,000 |
| 2025 | | | 10 | 10 | 8 | 5 | 33 | $ 3,500,000 | $ 115,500,000 |
| 2026 | | | | 10 | 8 | 5 | 23 | $ 3,500,000 | $ 80,500,000 |
| 2027 | | | | | 8 | 5 | 13 | $ 3,500,000 | $ 45,500,000 |
| Subtotal | | | | | | | | | $ 962,500,000 |
| Administration | | | | | | | | | $ 18,250,000 |
| Financial Intermediary Fees ( 2.0% ) | | | | | | | | | $ 19,250,000 |
| **Total** | | | | | | | | | **$ 1,000,000,000** |

**Income threshold for eligibility**. Birdsall and Meyer (2014) argue that GNI per capita is not a good indicator for determining aid eligibility and propose using median per capita consumption instead. A cutoff of PPP$10 for the median per capita income would make approximately 95 countries eligible, containing all low-income countries; all lower-middle

income countries except Ukraine; and a number of upper-middle income countries such as Peru, Romania, Mongolia, and the Dominican Republic (see Diofasi and Birdsall 2016). An alternative would be to use GNI per capita to determine eligibility. A GNI per capita of less than US$12,735 would include all countries classified by the World Bank as low- and middle-income (98 countries). A lower GNI per capita income threshold of US$4,125 would include only low- and lower-middle income countries (61 countries).

**Implementation**. *The "Learning Initiative Facility".* GOL could be implemented by an existing organization but would face the risk described in this paper of turning an unconventional approach into a conventional one with additional burdens. Existing organizations tend to have technical staff accustomed to detailed planning and policies that are designed to address risks that have little to do with assessments and learning. In particular, it would be advisable to avoid mechanisms like World Bank Trust Funds which are subject to a range of policies that delay implementation without clear benefit (on education, see Bermingham 2011; on climate change, see Savedoff 2016).

The ideal organization would be something like the "Learning Initiative Facility" described in Birdsall et al 2016. The Initiative builds on the UIS Global Alliance to Monitor Learning to create a 10 to15-year organization (sunset in 2030) bringing together philanthropic, donor, civil society, countries, and other groups with a core mission to promote and improve country and global data on children's learning, and promote use of these data to inform education and other policies, practices, and programs in developing countries. The Initiative would be a non-profit entity with technical, fiduciary, and legal capacity to receive and allocate resources to agencies, programs, and countries. A small staff would work under the auspices of a governing body (a "board") comprised of individuals with a diverse mix of the skills, perspectives, and experience required by the organization. The board would be elected by funders to renewable three-year terms but would not directly represent any particular funding organization. Technical advisory groups drawn from existing assessment and technical bodies (OECD, IEA, SACMEQ, PASEC, LLECE, READ, REACH, RTI) would support the "Learning Initiative Facility" in the development of strategy, program oversight, and other technical areas. However, as they are expected to be major grant recipients under the Initiative, they would have no legal and fiduciary responsibility or representation in the Initiative's governance or any of its financing decisions.

**Administrative costs.** Administrative costs are low because GOL does not provide technical assistance or *ex ante* funding for learning assessments. GOL's core function is to certify country eligibility and results. It would do this by employing a small number of staff to nominate, convene, and support periodic meetings of the GOL technical committee which, in turn, would instruct the FI on when and how much to disburse for the Assessment and Achievement Awards. The cost of these program activities is estimated at about US$1.7 million per year for an organization like the Initiative which would have the capacity to contract staff, manage modest amounts of funds and assure appropriate governance. Additional payments to the FI would be required to cover its costs, which are estimated to be two percent of the award values. Financial administration fees would be low because the FI's function is restricted to effecting payments when instructed by GOL. Furthermore, it

provides liquidity during relatively short periods (less than one year) and the transactions benefit from sovereign guarantees.

**Double payment**. Some funders may object to paying an Assessment Award to countries which have conducted qualified tests with financial support from bilateral or multilateral agencies by arguing that this represents a "double payment." This concern ignores the complexity of achieving development goals and the intangible assets invested by recipient countries. People concerned with double payment are assuming that the problem of undertaking a national assessment and institutionalizing it as a mechanism for feedback in the education system is a simple problem when in fact it is quite complex. It requires that recipients apply a lot of political, managerial and financial inputs of their own in ways that are not recognized by most project designs except as "risks" to be managed. Thus it is appropriate for a country to receive an Assessment Award or Achievement Award in addition to up-front technical and financial support. The agencies providing such up-front funding will be aware that GOL awards are available and may scale back their financial support after one or two Assessment Awards are disbursed.

**GOL technical committee**. The GOL program would convene an international technical committee to assess country eligibility, certify results, and inform arbitration over any eventual disagreements. The people selected for the committee would serve in their individual capacities and not as representatives of any country or organization. They would be selected for their impartiality, technical expertise, experience and every effort would be made to assure diversity. The committee would direct GOL program staff to prepare the information necessary for the committee to make its decisions regarding eligibility and certification of results. The committee would probably meet twice each year, once to identify information needed for making decisions and once to review that information and issue its decisions. The committee's decisions would be binding on all parties. If disagreements arose, an arbitration committee could be convened to review the technical committee's decision but only if both the organization administering the GOL program and the recipient country were to agree. A discussion of some technical committee models in other organizations is provided in the appendix.

## References


Bermingham, Desmond. 2011. "The politics of global education policy: the formation of the Education for All–Fast Track Initiative (FTI)." *Journal of Education Policy*, 26(4): 557-569.

Barder, Owen and Nancy Birdsall. 2006. "Payments for Progress: A Hands-Off Approach to Foreign Aid." CGD Working Paper No. 102. Washington, DC: Center for Global Development.

Birdsall, Nancy, Barbara Bruns, and Janeen Madan. 2016. "Learning Data for Better Policy: A Global Agenda." Policy Paper 092. Washington, DC: Center for Global Development.

Birdsall, Nancy and Christian Meyer. 2014. "The Median Is the Message: A Good-Enough Measure of Material Well-Being and Shared Development Progress." Working Paper 351, Washington, DC: Center for Global Development. January.

Birdsall, Nancy, William Savedoff and Frances Seymour. 2014. "The Brazil-Norway Agreement with Performance-Based Payments for Forest Conservation: Successes, Challenges, and Lessons." CGD Brief. Washington, DC: Center for Global Development. August.

Birdsall, Nancy and William Savedoff. 2010. *Cash on Delivery: A New Approach to Foreign Aid.* Washington, DC: Center for Global Development.

Cambridge Education. 2015. "Evaluation of the Pilot Project of Results-Based Aid in the Education Sector in Ethiopia Final Report, EC 2004 – 2006." Cambridge: Cambridge Education. May.

Clist, Paul, and Arjan Verschoor. 2014. "The Conceptual Basis of Payment by Results." School of International Development, University of East Anglia.

Cortez, Rafael and Daniela Romero. 2013. "Argentina Increasing Utilization of Health Care Services among the Uninsured Population: The Plan Nacer Program." UNICO Studies Series 12. Washington DC: World Bank. January.

Diofasi, Anna and Nancy Birdsall. 2016. "The World Bank's Poverty Statistics Lack Median Income Data, So We Filled In the Gap Ourselves." CGD Blog and Dataset. February.

Gelb, Alan 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.

Gelb, Alan, Anna Diofasi, and Hannah Postel. 2016. "The Anatomy of Program for Results: The First 35 Operations." Working Paper 430. Washington, DC: Center for Global Development.

Holzapfel, Sarah and Heiner Janus. 2015. "Improving Education Outcomes by Linking Payments to Results." Bonn: German Development Institute.

IDB (Inter-American Development Bank). 1996. *Social and Economic Progress Report 1996: Making Social Services Work*. Washington, DC: Inter-American Development Bank.

Kenny, Charles and William Savedoff. 2013. "Can Results-Based Payments Reduce Corruption?" CGD Working Paper 345. Washington, DC: Center for Global Development.

Perakis, Rita and William Savedoff. 2015. "Does Results-Based Aid Change Anything? Pecuniary Interests, Attention, Accountability, and Discretion in Four Case Studies." CGD Policy Paper 052. Washington, DC: Center for Global Development.

Pritchett, Lant. 2013. *The Rebirth of Education: Schooling Ain't Learning*. Washington, DC: Center for Global Development.

R4D (Results for Development Institute). 2016a. "Paying for Performance: An Analysis of Output-Based Aid in Education." Commissioned by the Global Partnership on Output-Based Aid (GPOBA). Washington, DC: R4D. February.

R4D. 2016b. "Database of Output-Based Aid Education Projects." Commissioned by the Global Partnership on Output-Based Aid (GPOBA). Washington, DC: R4D. February.

R4D. 2016c. "Literature Review: Situating Output-based Aid in the Context of Results-based Financing in Education." Commissioned by the Global Partnership on Output-Based Aid (GPOBA). Washington, DC: R4D. February.
Rose, Sarah, Nancy Birdsall, and Anna Diofasi. 2016. "Creating a Better Candidate Pool for the Millennium Challenge Corporation." Policy Paper. Washington, DC: Center for Global Development.
Sandefur, Justin. 2016. "Linking Regional and International Assessments: Prospects for Creating a Global Learning Metric." A note for the Secretariat of the International Commission on Financing Global Education Opportunities. Washington, DC: Center for Global Development. April.
Savedoff, William. 2016. "How the Green Climate Fund Could Promote REDD+ through a Cash on Delivery Instrument: Issues and Options. CGD Policy Paper 072. Washington, DC: Center for Global Development. January.
Upper Quartile. 2015. "Evaluation of Results Based Aid in Rwandan Education - Year Two." Glasgow, Scotland: Upper Quartile.
World Bank. 2015. "Program for Results: Two-Year Review." Washington, DC: The World Bank.

# Appendix A: Features of Results-Based Programs and Associated Options

**Reproduced from Perakis and Savedoff 2015**

**Appendix Table: Features of results-based programs and associated options**

(Select the appropriate option under each heading in order to describe a program in a single lengthy but complete sentence. See examples below.)

| Time Frame | Renew-able | Transparency (5) | Payment Amount (6) | | Recipient Type | | Recipient Discretion (3) | | Result complexity | Payment Function (2) | | Result Level (1) | Credibility (4) | | Up-Front Payments |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | National Govt | | | | | | | | | | |
| Short term (3 years or less) | | Public (public) | Reward (small share of TC) | | Subnational Govt | | complete discretion | | one | Incremental | | Impacts | | | |
| Medium term (4-7 years) | Renew-able | Public simple (transparent) | Subsidy (large share of TC) | *.. payment to ..* | Public Corporation | *.. with ..* | significant discretion | *for achieving* | few | target | *.. indicators for..* | Outcomes | that are verified with Independent information | *.. and with..* | Up-Front Payments |
| Long term (8 years or more) | non-Renew-able | Public simple & usable (highly transparent) | Reimbursement (close to 100% of cost) | | Public facility | | limited discretion | | many | threshold | | Goods and Services | that are not verified without Independent information | | Only Ex-Post payments |
| | | | Bonus with Reimbursement (more than 100% of cost) | | Private corporation | | | | | | | Tasks | | | |
| | | | | | Non-profit entity | | | | | | | Activities | | | |
| | | | | | Community Group | | | | | | | Processes | | | |
| | | | | | Household | | | | | | | Inputs | | | |
| | | | | | Individual | | | | | | | | | | |

Examples:

- COD Aid is a reward (or subsidy) payment to a national government with complete discretion for achieving one incremental outcome indicator that is verified with independent information and with only ex-post payments.
- OBA is a reimbursement or bonus with reimbursement to private entities for goods and services with significant recipient discretion
- PforR is a subsidy payment to a national or subnational government for achieving many target indicators for any results level. Other features are unspecified.
- RBF is a reimbursement or bonus with reimbursement to public facilities or private entities for goods and services with significant recipient discretion.
- RBA is a payment to national governments preferably for goods and services or outcomes. Other features not specified.

# Appendix B: Models for Setting Up an International Technical Committee

### The Amazon Fund

The Amazon Fund, which uses a pay-for-performance model to reduce deforestation in Brazil, is managed by the Brazilian Development Bank and governed by a Guidance Committee. The Guidance Committee is responsible for outlining the specific guidelines and criteria that govern how the Fund's resources are used. It consists of three blocks of representatives from federal government, state governments, and civil society. Each block holds one vote on committee decisions, and each member holds one vote within the block.[16] This framework allows key non-governmental stakeholders and different levels of government to provide input in the Fund's decision-making processes. In addition, the Fund also includes a Technical Committee comprised of scientific experts that certify annual calculations on the emissions avoided.

### International Initiative for Impact Evaluation

The International Initiative for Impact Evaluation (3ie) is an independent non-profit organization that promotes the generation and use of impact evaluations to guide public policy with evidence. Its process of reviewing grant proposals involves a technical review panel.[17] As a first step, all proposals are reviewed by one internal and 2 to 3 external independent reviewers. The average score from these reviews is then used to produce a shortlist of proposals. A technical committee of about five individuals, with relevant expertise and no conflict of interest is convened by the Executive Director to review the shortlist against specific criteria determined by 3ie. Proposals are ranked and passed on to its Board of Commissioners, which makes a final decision on grant awards.

**Global Alliance to Monitor Learning** (in proposal stage at the time of preparation of this report)

The Institute for Statistics' proposed Global Alliance to Monitor Learning, which is currently in its infancy, would include a Technical Standing Group (TSG) to provide technical guidance.[18] The TSG would include stakeholders and country representatives who meet the following criteria: impartiality, technical expertise in learning assessments, as well as broad geographic and inter-sectoral representation. A subset of this group, solely comprised of technical experts, could serve as the technical committee envisaged for the GOL program.

---

[16] Amazon Fund, "Amazon Fund Guidance Committee," available at http://bit.ly/24nqm2y and Forstater et al 2013.

[17] Details are drawn from the "Founding Document to Establishing the International Initiative for Impact Evaluation," June 25, 2008; and 3ie's Request for Proposals (E.g.: http://bit.ly/24vRrNz and http://bit.ly/1W6Nv4R)

[18] Based on the "Global Alliance for Learning – Concept Note, Draft," March 2016.

the
Education
Commission

The International
Commission
on Financing Global
Education Opportunity

**educationcommission.org**