Does Results-Based Aid Change Anything? Pecuniary Interests, Attention, Accountability and Discretion in Four Case Studies

Rita Perakis and William Savedoff

Abstract

This paper studies foreign assistance programs called Results-Based Aid (RBA) in which one government disburses funds to another for achieving an outcome. At least four theories are typically advanced to explain how RBA increases program effectiveness: by appealing to governments' pecuniary interests to shift domestic priorities, by drawing the attention of politicians and managers to results, by establishing accountability to constituents; and by giving recipients discretion to engage in local problem-solving. Using four case studies – from GAVI, the Amazon Fund, Ethiopian Secondary Education and Salud Mesoamérica – the paper analyzes program features to show which of these theories are being applied and what we can learn about the effectiveness of the RBA approach.

The four case studies show that concerns with corruption, unintended consequences, short-termism, and additional costs have not materialized. The analysis demonstrates that relatively few RBA programs are being piloted and that most do not rely on the assumption that a financial incentive will lead aid recipients to shift their priorities. RBA programs are also not typically designed to work through accountability and recipient discretion. Rather, most RBA programs seem designed to draw attention to results, making them more salient to politicians and managers. As relatively cautious adaptations of conventional approaches, these initiatives are unable to test the potential benefits from greater recipient discretion and public transparency. Future experimentation may test these other theories but for now, RBA remains a work in progress.

Rita Perakis is a program associate at the Center for Global Development and William Savedoff is a senior fellow at the Center for Global Development. The authors thank Tara Templin for her thorough work in compiling information about results-based aid projects. We are also grateful to the many people who took the time to share information about specific programs and debate concepts, including Belay Addise, Owen Barder, Nancy Birdsall, Paul Clist, Ellie Cockburn, Roger Drew, Alan Gelb, Amanda Glassman, Peter Hansen, Emma Iriarte, Donald Menzies, Ferdinando Regalia, Mead Over, and Frances Seymour. Any errors of interpretation or fact are our responsibility.

CGD is grateful for contributions from the UK Department for International Development and the William and Flora Hewlett Foundation in support of this work.

Rita Perakis and William Savedoff. 2015. "Does Results-Based Aid Change Anything? Pecuniary Interests, Attention, Accountability and Discretion in Four Case Studies." CGD Policy Paper 053. Washington DC: Center for Global Development.
http://www.cgdev.org/publication/does-results-based-aid-change-anything-pecuniary-interests-attention-accountability-and

Center for Global Development
2055 L Street
Fifth Floor
Washington DC 20036
202-416-4000
www.cgdev.org

Center for Global Development

Independent research and practical ideas for global prosperity

CGD Policy Paper 052
February 2015

Contents

Introduction

For decades, aid agencies have promoted the idea of disbursing payments in relation to results. Generally, such ideas were realized in programs that disbursed on the delivery of physical outputs or the enactment of laws and regulations. In the last decade, aid agencies have sought to bring this idea into practice by paying for less tangible outputs or outcomes in the spheres of environmental, social or governance objectives.

This paper surveys progress and draws lessons from a recent wave of foreign assistance programs – which we will call Results-Based Aid (RBA) – in which one government disburses funds to another for achieving an outcome or a proxy for such an outcome. RBA is usually promoted for a mix of reasons which encompass at least four distinct theories for how linking disbursements to results will make programs more effective: by appealing to governments' *pecuniary interests* to shift domestic priorities; by drawing the *attention* of politicians and managers to results; by establishing *accountability* to constituents; and by giving recipients *discretion* to engage in local problem-solving. To understand the implicit theory embodied in programs and consequently their prospects for success, this paper focuses on six program features: the results level (i.e., activities, outputs, or outcomes), the payment function, recipient discretion, credibility in how results are verified, transparency, and payment amounts.

After discussing these theories and program features, we analyze four RBA agreements to show which theories are actually being tested and how they function in practice. The analysis shows that concerns typically raised by critics over performance programs – such as corruption, unintended consequences, short-termism and high transaction costs – have not materialized. Of greater importance, the analysis demonstrates that relatively few RBA agreements paying for outcomes (or proxies of outcomes) are actually being piloted. Furthermore, most of these programs do not rely on financial incentives in the sense of using money to motivate recipients to shift their domestic priorities. Most RBA programs are also not testing whether the approach increases the accountability of the developing country government to its constituents or gives recipients discretion to engage in local problem-solving. Rather, drawing political and managerial attention to outcomes appears to be the most prominent theory being tested in RBA programs as they are currently designed and implemented. Most initiatives are cautious adaptations of conventional approaches and so the potential benefits from greater recipient discretion and public transparency remain untested. Future experimentation may test these other theories but for now, RBA remains a work in progress.

What Is Results-Based Aid?

There is no widely accepted nomenclature for programs that pay for results, though there have been numerous attempts to classify and distinguish them.[1] In part, this reflects different communities working on different issues. People in the private corporate world have developed their own terms of practice for performance contracts, results based management, and the like.[2] Similarly, the world of public administration experts has its own usage.[3] This paper reflects terminology applied in the world of international development organizations, including bilateral aid agencies, multilateral development banks, international NGOs, and the researchers who engage with these organizations.[4]

Aid agencies have implemented a wide range of programs which pay for results of one kind or another. Some of these programs pay providers for delivering services, while others offer prizes for technological innovations, or reward poor families directly through conditional cash transfers.[5] All of these programs use some form of incentive to improve results, but they vary greatly in the ways that incentives are used, the people or groups that are given incentives, and the level of the results chain at which they operate.

The focus of this paper is on a type of performance program called "Results-Based Aid" (RBA). The term RBA is used by the UK Department for International Development (DFID) to distinguish approaches which involve a contractual arrangement between a donor and national government from other performance-based approaches (Pearson 2011). This has become the most commonly used definition for results-based aid: a partnership between a donor and partner government that ties disbursement to results.[6] This paper uses that definition, but focuses on a *subset* of RBA programs in which the measured results are defined as closely as possible to development *outcomes*. In discussing RBA theories and case

[1] 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.
[2] One of the earliest of these is Drucker 1954.
[3] For an overview of different theoretical frames informing New Public Administration, see Gruening 2001.
[4] For a discussion of results-based management in relation to development agencies, see Meier 2003 and Vähämäki et al. 2011.
[5] See Savedoff 2011 for more on how a range of incentive-based development programs are defined and classified.
[6] See for example de Hennin and Rozema 2011; Pearson 2011; Klingebiel 2012; or DFID's Strategy for Payment by results 2014.

study examples, we are therefore referring to *agreements in which a donor government makes ex post payments to a national or local government for something that measures or is a proxy for an outcome.*

The two key features of this definition – who gets paid in the agreement and what they get paid for – distinguish this subset of RBA from a wide range of other programs that link payments to performance. *First, RBA agreements disburse funds to governments*. This distinguishes RBA from agreements that pay individuals, households, public facilities, communities, non-profit entities, or private businesses for some deliverable. Agreements involving this latter group of agents have sometimes been called "Results Based Financing" (RBF), a term which has been adopted by the World Bank for these kinds of health programs (www.rbfhealth.org) and Britain's program for "Results Based Financing for Low Carbon Energy Access."[7] RBA, which funds governments, is likely to require different designs than RBF because the nature of governments differs from other categories of recipients in terms of resources, behaviors, and dynamics.

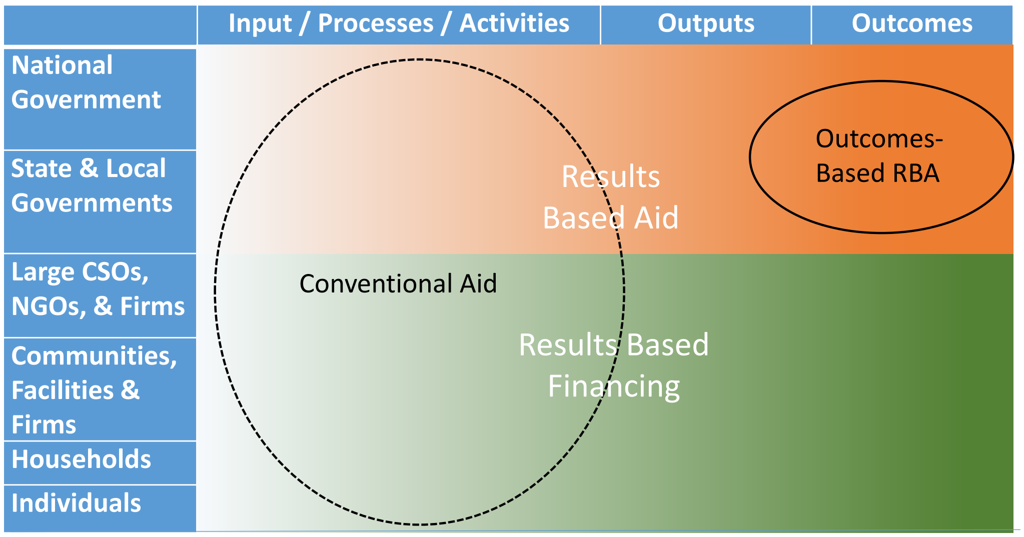
Second, we are exploring those RBA agreements which disburse funds against outcomes (or output indicators that closely approximate outcomes). This distinguishes outcomes-based programs from those that pay governments to purchase inputs, undertake activities, complete processes, or adopt policies which 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 but there is clearly a difference between programs that trigger disbursements upon passage of a new credit law, completion of teacher training courses, construction of a power plant, or malnutrition screening for children and those which disburse upon verifying that poorer people are getting access to credit, 12-year-olds can read and write, electricity outages are reduced, or child mortality has declined. We focus on the subset of RBA agreements that encompass the latter and not the former because we are interested in examining implementation experiences with those relatively new development programs which are paying for results "further along" on the chain.

RBA programs can be pictured on a two-dimensional figure that distinguishes different agents on the vertical axis and results levels on the horizontal axis (See Figure 1 and Figure 2). RBA programs are in the top of the diagram because they pay national governments. The subset of RBA agreements that interest us are located in the northeast quadrant because they aim to reward outcomes. This subset is different from programs that pay individuals for

[7] http://devtracker.dfid.gov.uk/projects/GB-1-202957/

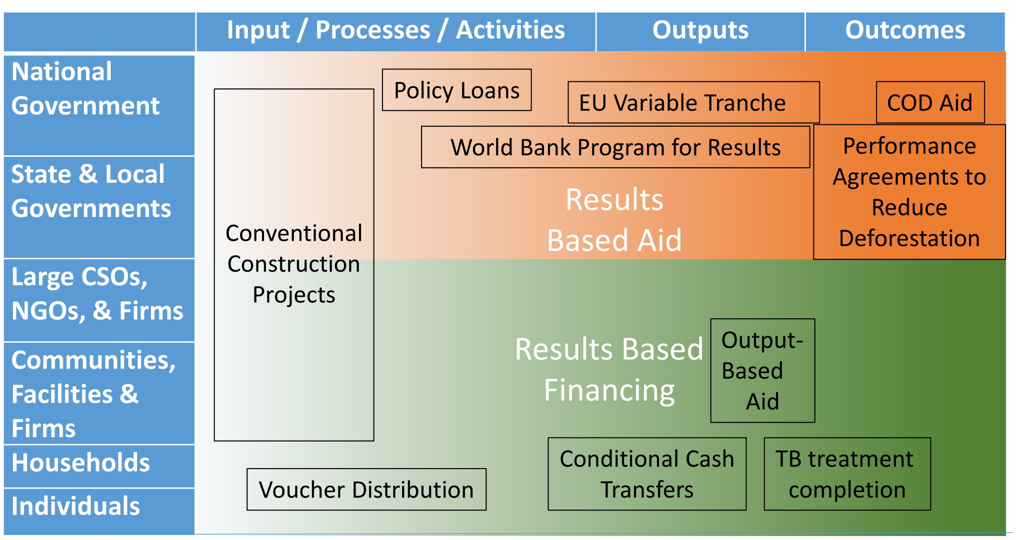
outcomes (e.g., successful completion of tuberculosis treatments); that pay individuals for inputs (e.g., provide vouchers to individuals for cook stove purchases); and that pay national governments for inputs (e.g., policy loan disbursements for enacting laws).

Figure 1: Performance programs by recipient type and results level



Source: Authors.

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



Source: Authors.
Notes: CSOs = Civil Society Organizations; NGOs = Non-Governmental Organizations.

Key Features of RBA Agreements

To fully characterize an RBA program, though, it is necessary to specify much more than the nature of the recipient (e.g., a government) and the results level (e.g., a measure of or proxy for an outcome). From our review, we developed a list of eleven design features that distinguish one RBA agreement from another (see Appendix A) and represent choices policymakers face when establishing performance agreements. In this paper, we chose to focus on six features that we judged particularly essential to predicting how an RBA program will function. They are:

- *Results level*: This paper focuses on agreements that pay in proportion to outcomes or proxies for outcomes. In practice, many agreements pay for changes at different points on the "results chain" – whether that means paying upon purchase of specific inputs, the completion of certain tasks or activities, enacting regulations or laws, establishing procedures, or producing goods or services.[8] Programs are likely to get more of what they pay for, whether those are inputs, processes, or outcomes.
- *Payment function*: Payments can be made in proportion to progress, in which case they are a continuous function. 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. Continuous functions (in progress units or time) create incentives for marginal improvements at any level of progress whereas thresholds and step functions create discontinuities that may undermine incentive effects at different achievement levels.[9]
- *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. Greater recipient discretion may increase the recipient's level of commitment and improve chances of success by giving recipients' flexibility in implementation, opportunities to innovate, and the ability to be more responsive to constituents.

[8] Perakis and Savedoff, 2014 offers examples of input, output, and outcome indicators for several development sectors.

[9] For a discussion of the implication of payment functions within performance payment programs, see Hallett and Over 2010 with an illustration for programs to reduce the prevalence of HIV/AIDS.

- *Credibility*: An agreement's credibility will influence whether the parties pay attention to it and whether it is likely to motivate change. The funder's credibility in terms of fulfilling commitments to pay and not to pay is important if the recipient is going to respond to the opportunity provided by the agreement. The credibility of the performance measurement is particularly important for RBA programs because indicators of performance determine disbursements. To avoid manipulation of performance measures, RBA agreements can contract an independent agent to verify results. They can also arrange for the collection of independent information to either generate results estimates or cross-check a governmental reporting system. Credible performance measurement can increase the chance that funders will abide by the agreement and that recipients will focus on achieving real progress.
- *Transparency*: The transparency of an agreement creates opportunities for more actors to influence a program. Agreements that are structured in ways that facilitate public dissemination are easier for officials themselves to understand and manage. The degree of transparency also introduces additional sources of feedback and accountability – from beneficiaries, civil society, employees, and peer governments among others – which may encourage better performance.
- *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. When payment amounts are small, relative to unit costs or to other funding sources, they still signal changes in outcomes that may be useful for management or accountability but their impact via pecuniary interests will be attenuated.

The six features emphasized here are also reflected in the results-based approach proposed by Birdsall and Savedoff (2010) called "Cash on Delivery Aid" (COD Aid). Birdsall and Savedoff argued that five key features are critical for effective agreements because these features can simplify the funder-recipient relationship, shift funders' attention to measuring and paying for development outcomes, and open space for recipients to focus on achieving progress (see Box 1).

Box 1: Five Key Features of COD Aid

- **Payment for outcomes, not inputs:** The program starts by defining an outcome related to an objective shared by both funder and recipient. Outcomes should be measurable and continuous so that progress can be rewarded over time in proportion to incremental improvements.

- **Hands off funders, responsible recipients:** Recipients assume full responsibility for the program. Funders do not specify or monitor inputs, set policy conditions, or track the use of COD Aid funds. Rather, they verify progress toward the shared objective, provide technical assistance only upon demand, and pay for outcomes according to the COD Aid agreement.

- **Independent verification of progress:** Results are independently verified by a third party to ensure that both the funder and the recipient have confidence in the way that progress is measured.

- **Transparency and public dissemination:** Both the contract and progress measures should be as simple as possible and made publicly available. This increases credibility and accountability, and makes it possible for constituents in each country to monitor and hold their own governments accountable for progress.

- **Complementarity with other aid programs:** COD Aid is intended to complement and not disrupt ongoing programs, whether funded by local or external sources. COD Aid should help the country use all available resources more effectively.

Source: Birdsall and Savedoff 2010.

To sum up, this paper discusses *a subset of RBA performance agreements that make ex post payments to a national or local government based on indicators that measure or are a proxy for outcomes*. Within this subset, substantial design variation remains in terms of results levels, payment functions, recipient discretion, credibility, transparency, and payment amounts. Differences in these features could have large implications for the effectiveness of an agreement depending on its underlying theory, an issue to which we now turn.

RBA Theories: Pecuniary Interest, Attention, Accountability and Discretion

Proponents of RBA assert that linking payments to outputs and outcomes will lead to change more effectively than conventional aid approaches for different reasons. Some proponents argue that when payments are linked to results, recipient governments respond because of their *pecuniary interest* – they will shift their domestic priorities or exert greater effort because they need the funding. Other proponents highlight how performance payments make results visible in a way that draws the *attention* of politicians and bureaucrats to outcomes or generate *accountability* to constituents. A fourth group of proponents emphasize how RBA allows funders to give recipients greater *discretion* during implementation, creating opportunities to adapt and learn. The particular features designed into an RBA program will reveal which of these theories is really being piloted, which may or may not align with explicit justifications in program documents.

Theories that focus on **pecuniary interest** argue that "you get what you pay for." People who emphasize such theories criticize conventional aid for rewarding governments when they complete tasks and paperwork independent of progress in improving outcomes. Paying for outcomes instead, it is argued, motivates a recipient government to deliver goods and services that will have a measurable influence on the outcome measure because the government needs or is mainly attracted by the money. Faced with this pressure, governments are expected to shift resources – whether physical, financial, managerial, or political – toward the desired outcome and to choose the most effective and least costly approaches for attaining that end.

The most common foundation for this approach is from institutional economics and, particularly, from principal-agent models.[10] When these models are used, funders are treated as "principals" who are trying to achieve an objective by contracting with a recipient who is conceived as an "agent." Fundamental to these theories is a divergence in objectives between principals and agents which principals "solve" by offering a contract that aligns agents' incentives with their own. Seen this way, an RBA program makes a recipient explicitly face a tradeoff between shifting resources toward the funder's objectives (in return for funding) and continuing to pursue their own objectives (and give up the funds).

[10] See Sappington 1991 for a general treatment of incentives. Clist and Verschoor 2014 use the principal-agency literature to elucidate RBA and RBF programs.

Principal-agent models may not be relevant to RBA agreements because they have been developed and tested primarily for interactions among individuals and firms, not governments or multilateral organizations. Furthermore, most aid relationships are conceived as partnerships between sovereign governments in which objectives are substantially overlapping even if there are areas of divergence. Of even greater importance, governments are not unitary actors with well-defined preferences or direct links between decisions and action.

For these reasons, political-economy approaches are likely to be more relevant for relationships between governments than principal-agent models, though few such theories have been advanced in the international aid literature to explain why financial incentives would motivate changes in public policy (Levy 1997). The only empirical literature of which we are aware that looks specifically at how national governments respond to external financial incentives relates to the debate over policy reform loans in the 1980s and 1990s (Killick 1998; Dollar and Svensson 2000). This literature generally finds little evidence that the conditionality in these incentive programs was able to induce domestic policy reforms. This could have been because the incentives were not attractive enough; because the international agencies lacked credibility since they commonly waived policy conditions; or because governments, as complex political organizations, do not respond in a direct and simple fashion to financial inducements.

If pecuniary interest is the key mechanism to induce change, then <u>outcomes, payment amounts, and credibility</u> will all be critical features for an RBA agreement. If you are likely to "get what you pay for" then you want to make sure you "pay for what you (really) want." This will generally mean paying for impact, if possible, but certainly for *outcomes* rather than inputs, activities or processes. The *payment amount* will make a big difference because when it is larger, it will have more power to motivate and mobilize action. Finally, *credibility of measurement* will be important because recipients need confidence that the measurement of results is accurate and that payments will only be forthcoming in relation to those measurements. Otherwise, they may discount the value of the financial incentive and exert less effort.

Other arguments downplay the importance of pecuniary interests and emphasize, instead, that RBA agreements increase aid effectiveness by bringing the **attention** of politicians, bureaucrats and managers to outcomes in a way that focuses them on achieving progress. This is not too far from the adage that "you can't manage what you don't measure." Works in the field of business management, psychology and economics (e.g., bounded rationality) recognize that people have limited capacities to absorb information and make decisions. As a result, they will generally act on the limited information available to them rather than expend efforts to find the information they need (Greenberg 2004; Kahneman 2003). Politicians, bureaucrats and managers are more likely to be faced with information about inputs – like budgets, disbursements and hiring decisions – than with information about what public programs are achieving. Similarly, conventional aid programs tend to spend more time monitoring inputs, activities and processes than they do documenting changes in outcomes and trying to relate them to different interventions. By paying for results, RBA agreements generate information that reaches policymakers because the funding linked to these results is a tangible signal – even when the payment amounts are small.

If drawing attention to outcomes is the key mechanism for change, then paying for outcomes is the critical feature for an RBA agreement. As with *pecuniary interest*, choosing results which come close to measuring real goals of public policy will be important because otherwise people will have their attention diverted to less important or less effective activities. The *payment amount* is less important here because even a low cost share – in the form of a prize, reward or subsidy – is sufficient to make performance visible and provide feedback to politicians and managers. Higher amounts would generate more attention but even small amounts provide a signal that funding is linked to progress on outcomes and no longer linked to inputs and activities.

A third theory argues that RBA programs induce change by generating government **accountability toward constituents**. Traditional aid arrangements typically hold recipient governments accountable to funding agencies. This can disrupt the important relationship between a recipient government and its domestic constituencies. RBA programs can exacerbate this problem if they are designed primarily to make recipients shift their priorities toward goals favored by funders. However, when RBA programs are designed around goals that are truly shared and openly reported, they can become an instrument for public accountability. An RBA agreement that is simple and public tells constituents exactly what outcome the recipient government has agreed to pursue (e.g., universal schooling, widespread access to energy), how much funding is forthcoming, and then allows them and/or civil society organizations to periodically monitor the recipient government's performance.

Theories in the fields of sociology, social psychology and political-economy are relevant to understanding how this model of change might operate. Such theories highlight the role of information and transparency in creating conditions for popular mobilization around key issues, formation of coalitions and identities, focusing public debate and ultimately rewarding or punishing political parties and politicians (Stiglitz 2002; Fox 2007). Politicians may be motivated to act more effectively if their performance is visible, either because they care about their reputation among peers (other governments or political elites) or because they believe it will influence their standing with key constituents. The accountability of developing country governments to their constituents is the essential one for achieving the agreement's development goals. However, this theory also has implications for donor governments who can be held accountable by their citizens for the use of public money by demonstrating what aid funds actually "bought."

If accountability toward constituents is the mechanism by which an RBA agreement is going to induce change, then <u>outcomes, recipient discretion and transparency</u> are going to be key features. *Outcomes* are important to this mechanism because unless they reflect goals of interest to the population or important stakeholders, the expected accountability relationship is unlikely to materialize. *Transparency* is particularly critical for this mechanism because regular dissemination of results and progress is necessary for peers and constituents to form opinions and influence leaders. *Recipient discretion* is also essential because recipient governments cannot be held accountable if they lack the autonomy to make choices and take action; without discretion, they can always deflect blame for lack of progress on the plans that were imposed on them by funders.

A fourth category of theories focuses on **recipient discretion**. Most aid programs are trying to achieve progress in highly complex, contingent and idiosyncratic settings (Andrews et al 2012; Barder 2012; Ramalingam 2013). In such contexts, conventional aid programs – which adhere to planned interventions – interfere in the normal process by which public policies are debated, negotiated, implemented, and modified. RBA programs give funders the opportunity to demonstrate that funds were used well by linking payments to results instead of tying them to pre-planned inputs and activities. This frees recipient governments to pursue a range of strategies and approaches, based on their local knowledge and embedded in local political and social dynamics, with greater chances of success.

Foreign aid agencies have long recognized the importance of recipient discretion to effective development. Since the 1990s, this has been manifested in general budget support programs which are designed to provide more predictable and flexible funding to recipient governments. General budget support programs channel funds through recipient countries' own financial, administrative, and political institutions rather than imposing external rules or financing projects in parallel systems. These programs exhibit inherent tensions between providing predictable funds and linking payments to performance, and between supporting domestic policies and responding to external priorities (Eifert and Gelb 2005; Cordella and Dell'Ariccia 2007). Unlike RBA agreements – which link funding explicitly to one or a few specific outcomes – general budget support gives recipients discretion over the use of funds within a more complex framework of negotiated disbursements (Koeberle et al 2006).

If recipient discretion is the key mechanism for change, then recipient discretion is really the only critical feature for an RBA agreement. Recipient discretion is important in several ways. Recipient governments may be more committed to programs that they themselves have designed and adopted (compared to those designed or heavily influenced by a funder). They may design more effective programs by relying on local knowledge and understanding of context. They will also be able to experiment, learn and adapt over time rather than focus on implementing a predetermined blueprint which may or may not be appropriate under changing circumstances.

The simplest way for funders to give recipients discretion is to provide lump sum payments (Clist and Verschoor 2014). But the justification for an RBA agreement under this line of reasoning is that it solves a problem for *funders* – it gives them a way to pay for progress (which they can demonstrate to their own constituents through *credible outcome measures*) without imposing conditions or rigid plans for the use of resources. Thus payment for outcomes and measurement credibility are important to theories that emphasize the role of recipient discretion because they can affect the *funder's* willingness to relinquish substantial control, not because they influence the recipient's behavior or affect the likelihood of outcomes being achieved.

All four theories regarding the way RBA could make aid more effective are plausible. They overlap enough that more than one theory may apply to the same agreement. However, they are also sufficiently distinct that a program designed to improve outcomes relying on one theory may lack features that are necessary for another. Thus in many cases it is possible to infer the implicit theory driving RBA agreements by looking at the features embedded in their design and in the way they are implemented (see). For example, an agreement that pays a small amount relative to unit costs and total funding available is unlikely to be relying on pecuniary interest as the main force for change; rather it is likely to be relying on the payment to draw political or managerial attention to outcomes. A program that is implemented with limited transparency would suggest that accountability is not going to play a significant role. Without recipient discretion, an agreement is clearly not relying on local innovation and experimentation; it may, however, still be relying on increased visibility of results to draw political and managerial attention to outcomes in a context of co-responsibility between funders and recipients.

Just because a feature is not required to test a particular theory (Table 1) does not mean that its value in designing a good program should be ignored. For example, credibility provided by independent verification of results is likely to be beneficial for almost any RBA agreement. Similarly, transparency is likely to strengthen the credibility, responsiveness and adherence of both parties even if the primary motor of change is pecuniary interests or attention. The role of a feature in testing theories is different than its role in improving the chances of success for a program.

Distinguishing these theories is not simply a conceptual exercise. It is critical to extracting the right lessons from the experience with RBA agreements. Whether successful or not, an RBA agreement tests the theory embedded in its design and implementation, not the one which is described in executive summaries and speeches. If recipients lack discretion, then the success or failure of an RBA program in terms of outcomes tells us little about the role of recipient discretion in achieving goals. If payments are small, they may tell us a great deal about drawing attention to outcomes or even accountability but much less about shifting priorities by satisfying pecuniary interests. In particular, the term "incentives" is used so broadly in describing performance programs – applying it to human as well as social behaviors and to marginal as well as significant payments – that it creates more confusion than clarity about the kind of approach that is really being piloted.

Table 1: Features of RBA agreements and related theories

Is this a critical feature of the RBA agreement…	**… in order to test this theory?**			
	Pecuniary Interest	**Attention to Outcomes**	**Accountability**	**Recipient Discretion**
Results Level: Outcomes	Yes	Yes	Yes	No
Payment Function: Incremental	Probably	No	Probably	No
Recipient Discretion	No	No	Yes	Yes
Credibility: Independent Verification	Yes	No	Probably	No
Transparency	No	No	Yes	No
Payment Amount	Yes	No	Probably	No

Understanding which theory is embodied in an RBA agreement is also important for evaluating the agreement's performance. Changes in the outcomes which are the subject of an RBA agreement are not, in themselves, evidence of the success or failure of the RBA approach. For example, an agreement designed to maximize recipient discretion cannot guarantee that the recipient will use that discretion in a way that improves outcomes; however, the design can be judged for whether or not the recipient, in fact, enjoyed and utilized that discretion. Subsequently, with enough experience, it would be possible to analyze whether outcomes improve more when supported through RBA agreements that operate through recipient discretion than conventional aid programs. Similar arguments could be made for evaluating RBA agreements designed with other theories in mind.

Four RBA Experiences

The research for this paper began with a search for the full universe of results-based payments in foreign aid programs. We used a broad definition of results-based payments to include any program that links funding to measurable results *ex-post,* and sought to categorize programs according to criteria such as *who is receiving the incentive* and *what kinds of indicators are being used to measure progress* (see Appendix B for further explanation and a list of criteria). It quickly became apparent that the universe of results-based funding was very big, but the number of programs with specific characteristics that interested us, that is, paying *governments* for progress on improving *outcomes* (as proposed in CGD's work on Cash on Delivery Aid) was very small. This finding contradicted what we typically heard from development agency staff and experts who told us that a large number of these RBA programs already existed. When we discovered so few examples, we chose to use a case study approach to delve deeply into examples we identified as (1) programs in which funders pay a *government* on the basis of results achieved where (2) results are defined as development *outcomes* or proxies for outcomes and (3) the agreement had been in place for at least a year and (4) information about the program was publicly available.

At the time of this exercise (August 2013), we identified only 6 programs that clearly met these criteria. Three of them were Norwegian programs to reduce deforestation (Brazil, Guyana and Indonesia) of which we chose to include the Brazilian program (Amazon Fund) because it was the longest running and best documented. We included all three other programs – The GAVI Alliance's Immunization Services Support (ISS) multi-country program, a secondary education program in Ethiopia financed by the UK Department for International Development (DFID) and a regional health program called Salud Mesoamérica 2015 administered by the Inter-American Development Bank (IDB).[11] At that time, DFID had initiated a second results-based aid program in the education sector in Rwanda that had some similar characteristics to the Ethiopia pilot (and additional programs since then with some form of payment for outcomes) but it was at an earlier stage of implementation.[12] The GAVI ISS program phased out in 2010 but we tracked the implementation of the other three programs, which are all approaching the end of their initial phases this year (2015).

[11] We tried to include the EU's MDG contracts but were unable to obtain the most basic documentation regarding the content of the contracts and procedures that were followed.

[12] In September 2014, DFID posted a list of pilot performance programs, some of which pay for outcomes and would meet our criteria today if they had been in the public domain: https://www.gov.uk/government/publications/dfid-guidance-on-payment-by-results-and-spreadsheet-of-pbr-projects

The programs are summarized in Table 2. All four make payments on the basis of indicators which are reasonable proxies for outcomes; three of them use incremental measures; two allow greater recipient discretion than the others. They vary with regard to independent verification, transparency and whether the payment is more of a reward or a subsidy. We will show that three of the programs seem designed to work primarily by making results more visible, drawing attention to outcomes and generating feedback for management. The Amazon Fund is the only case which is structured to work through accountability and recipient discretion – and in this case, the amount paid is unlikely to be driving change solely on the basis of pecuniary interests.

Table 2: Summary of Features for Four Case Studies

	GAVI	Amazon Fund	Ethiopian Secondary Education	Salud Mesoamérica 2015
Results level: Outcomes?	Close	Close	Close	Close
Payment function: Incremental?	Yes	Yes	Yes	No
Recipient Discretion?	Some	Yes	Yes	Some
Credibility: Is there independent verification?	Limited	Yes	Yes	Yes
Transparency?	Limited	Significant	Some	Limited
Payment Amount?[13]	100% of unit costs but small share of health budgets	10% of unit costs & small share of government budget	74% or 15% of unit costs & small share of government & aid budgets	25% of unit costs & small share of health budgets

[13] The figures for unit cost shares are extremely rough estimates given lack of data and are calculated as follows. GAVI initially paid $20 per vaccinated child which was based on an estimate of average costs. It is worth noting that although this figure was roughly expected to cover costs for GAVI overall, for any given country program, it might have been significantly more or less than actual costs. A later evaluation of the ISS program (Abt Associates, 2007) concluded that the average cost per additional immunized child was $53. Norway pays $5 per avoided ton of CO2 emissions which turned out to cover only 10% of the emissions actually avoided, although not necessarily of the cost to Brazil of avoiding those emissions, which is unknown. The UK's RBA program for secondary education completion in Ethiopia pays about US$100 per additional student compared to annual expenditures per student in Ethiopia of about US$68 (calculated for 2012/2013 with data in Ethiopian Ministry of Education 2013). The average UK payment therefore represents about 74% of the cost of educating a student for the two years of lower secondary school or 15% if the cost of the requisite primary school years are included. Salud Mesoamérica is structured to reimburse up to 25% of total program costs in relation to achievements. For more detail, see specific case discussion.

The following sections describe each program in chronological order, beginning with an overview and continuing with an analysis of the six highlighted features and related lessons.

Case Study 1: GAVI Immunization Services Support

The GAVI Alliance was established in 2000 to increase immunization coverage in developing countries. It is a partnership of public and private actors - including the World Health Organization, UNICEF, the World Bank, the Bill & Melinda Gates Foundation, bilateral aid organizations, governments of developing countries, research institutes, and vaccine industry representatives. GAVI started its Immunization Services Support (ISS) program to help countries expand routine immunization coverage using a results-based funding mechanism. ISS was an innovative, multi-country approach to funding vaccination programs, and one of the first programs to pay governments for units of progress toward a simply defined development goal.

ISS paid countries for increases in routine immunization, measured as coverage rates of the diphtheria-tetanus-pertussis vaccine (DTP3). A total of 62 countries received funding under the ISS program between 2000 and 2010. To be eligible, countries had to have per capita gross national income less than US$1,000 and immunization coverage rates of below 80% (although GAVI made some exceptions to these eligibility criteria) (Fan et al. 2013). Each country presented a plan for increasing routine immunization coverage and applied for five years of funding. This included three years of initial investment based on projections of the number of children to be immunized with DTP3. In subsequent years, countries were paid for increasing the number of children immunized with DTP3 over the number of children who had been targeted in the first year. The payments were US$20 for each additional child reached with three doses of DTP3. This figure was close to estimates at that time for the average cost of fully vaccinating a child with the basic series of six antigens. A key feature of ISS funding was that it was flexible, with GAVI giving governments the authority to choose how funds would be spent as long as they were spent within the health sector. The reward payments were intended to help countries continue to cover the costs of the immunization program.

Under the ISS program, performance payments were calculated from immunization data provided from countries' administrative reporting systems. External evaluators assessed the quality of these reporting systems before GAVI would accept them. The evaluators conducted a one-time Data Quality Audit (DQA) and countries which scored better than 80 percent on a series of standards were then eligible for the performance component funding. Countries that did not meet this eligibility standard were encouraged to improve data quality to meet the threshold for a future DQA.

Questions arose over the program's system for verifying country data when an independent study conducted by the Institute of Health Metrics and Evaluation (IHME) found discrepancies between household survey results and the results from the administrative reports (Lim et al. 2008). The study used modeling from household surveys to estimate immunization rates and concluded that countries had overestimated immunization coverage, meaning that GAVI had overpaid in many cases. According to IHME, on the basis of survey data that measured additional children immunized and a payment of US$20 per child, GAVI's payments up through 2008 should have amounted to US$150 million, while GAVI had actually disbursed US$290 million at that point. This led to a temporary suspension of ISS funding. More generally, the IHME study exposed the challenges of getting clear and accurate outcomes data when independent information is not used to cross-check self-reported performance.

GAVI responded to the criticism by conducting its own investigation which compared country administrative data to WHO/UNICEF estimated coverage rates. In 2009, GAVI reported that most countries reported data that were similar to the WHO/UNICEF estimates and concluded that there were fewer cases of "overpayments" than the study by Lim and co-authors had found (GAVI 2009a). According to statements from GAVI, eight countries may have received overpayments, amounting to between US$30 million and US$40 million, but argued that overall ISS helped to increase coverage in 62 countries from 65% to 78% (GAVI 2009b).

A more recent study by Sandefur and Glassman (2014) provides additional evidence that immunization coverage data was biased as a consequence of the incentive payments. This study showed that the gap between administrative reports and Demographic Health Survey (DHS) estimates of DPT3 immunization coverage rose significantly in 2000 when the performance scheme was introduced. By contrast, no significant change in this gap between administrative and survey information occurred for measles vaccinations, an output for which no performance scheme existed. This corroborates the view that independent verification of self-reported administrative data is needed whenever performance-based payments are introduced.

The ISS program is now being phased out and replaced with a new performance-based funding scheme under the Health Systems Strengthening (HSS) program. Under the new scheme, countries will receive an initial investment in the first year, and in subsequent years a portion of payments will depend on improvements in immunization coverage. The program has one set of incentives for countries seeking to maintain high immunization coverage and another incentive approach for those seeking to increase coverage (Fan et al. 2013 and GAVI 2011). The new HSS program emphasizes the need to improve data quality necessary to reduce discrepancies between administrative and survey reports. Countries can include investments in data quality and routine information systems as part of their HSS application (GAVI 2013). GAVI is moving towards the use of household surveys as a means of data verification in some countries but it is unclear how many will be independently verified because verification methods will be agreed with countries on a case-by-case basis.

GAVI Immunization Services Support: Key Features of RBA Agreement

Results Level: Outcomes? The GAVI ISS program paid for increases in the number of immunized children, which represents a development outcome: vaccination is a strong proxy for improved child health and may be correlated with the quality of a country's healthcare system (Murray and Evans 2006). The program used a simple structure by paying for a single results indicator - increases in the number of children immunized with DTP3.

Payment Function: Incremental? The ISS program provided a fixed payment of US$20 per additional child vaccinated. Although the cost per child of rolling out an effective vaccination program varies from one country to another, the ISS program offered a universal price, which kept the design simple and furthermore sent a message that the value of vaccinating a child is equal across contexts.[14] Paying for additional children, rather than reaching a target, created incentives to reach as many children as possible and avoided the dilemma of choosing a target that would be neither too low to be meaningless nor too high to be unreachable. Paying for additional children rather than a rise in the coverage *rate* also created incentives to reach more children even if the cohort size was declining and furthermore avoided errors that could have entered the calculations from poor estimates of cohort size.

[14] Further discussion about 'payment size' can be found in Birdsall and Savedoff 2010, p.53 and Vivid Economics forthcoming.

Recipient Discretion? The program took country ownership seriously and didn't impose specific requirements on how vaccination programs were to be rolled out or how reward funds were to be used. An evaluation of the ISS program reports that countries reacted positively to this aspect and that there was a lot of variation in how funds were used (Abt Associates 2007). The only requirement was that funds had to be used within the health sector which made it possible to adjust and reprogram funding during implementation. However, countries did have to provide detailed plans before they could access funds.

Credibility? The lack of a credible, independent source to verify government-reported data was perhaps the ISS program's most significant weakness. The GAVI case demonstrates that credible measurement is essential for programs that pay for results. Independent verification is necessary if funders and recipients are going to have confidence that results that are achieved and paid for are real. The key for independent verification is to collect information from independent sources, that is, from sources whose bias is unlikely to be correlated with the source of the recipient's own reports. Household surveys are one way to check administrative reporting but not the only way. Independent agents could also perform random checks on administrative data, for example, by randomly selecting individuals listed on administrative immunization records and visiting homes to confirm that they have indeed been vaccinated.

Transparency? The 2007 evaluation of the ISS program looked in detail at six countries and found variation in how informed countries were about the reward payment mechanism at subnational levels and the extent to which it created incentives for better performance. The study did find that awareness of the incentive program had generally increased over time, demonstrating that new aid modalities take time for funders and recipients to understand. Some countries passed incentives on to health staff at local levels, but overall few people seemed to link improved immunization coverage with increases in GAVI funding (Abt Associates 2007).

The program's transparency was not addressed systematically in the evaluation across all ISS countries. While GAVI publishes DTP3 immunization coverage and total disbursements made through the ISS program for each country, this data does not clearly distinguish which payments were made in the performance-based payment phase for each country nor does it demonstrate the link between payments and results. We were unable to assess the linkage using other data sources such as the World Health Organization (WHO) because they are not sufficiently disaggregated.

Payment Amount? GAVI established a payment of US$20 per immunized child which corresponded with estimates at that time of the average cost of immunization. The program may or may not have fulfilled its intention of covering full unit costs because immunization programs often rely on physical infrastructure and human resources financed through other channels. The GAVI program operates alongside other domestic and foreign-funded health sector programs, but it is still one of the biggest channels for financing immunization programs in the world (along with UNICEF). According to WHO, low- and low-middle-income countries depend on foreign aid to cover two-thirds of their immunization programs and a significant and increasing part of this comes through GAVI (WHO 2009). GAVI's spending is a significant share of domestic spending on immunizations which could constrain the ability to finance only ex-post if it were not for the fact that immunization programs in most countries are small relative to public health spending. Considering that immunization is so highly cost-effective, this reliance on external funding for a program that could be domestically sustained is problematic (Glassman et al. 2013).

Predominant Theory: Attention. GAVI's intention to cover unit costs indicates that pecuniary interests could be a factor in this case. However, the small size of the payment amount relative to total health budgets explains why governments did not in fact appear to have been driven by pecuniary interests. GAVI documents emphasize that the program gave recipients flexibility in how they spent their funds, but this is not consistent with the degree of planning, systems analysis, plan approval, and monitoring that occurred before and during implementation. The program also lacked transparency to the public in a way that would have generated accountability. On balance, the key channel of change induced by this program was increased attention on the performance of countries' immunization systems in order to improve program implementation and management.

Case Study 2: Amazon Fund[15]

The Amazon Fund is the first large-scale international effort to promote forest conservation with a performance-based financing mechanism. It was proposed initially by the government of Brazil at the time of the annual meeting of the United Nations Framework Convention on Climate Change (UNFCCC) in 2007 and the Norwegian government promptly responded with interest in piloting an agreement. The subsequent agreement transferred official development assistance from Norway to Brazil on the basis of verified reductions in

[15] This section draws heavily upon Forstater et al. 2013 and Birdsall et al. 2014.

carbon emissions from slowing the rate of deforestation. It quickly became a model for international financing to reduce emissions from deforestation and an early example of a national results-based aid program. Norway subsequently negotiated and signed similar agreements with Guyana, Indonesia, Liberia and Peru.

Before the Amazon Fund was even proposed, Brazil had already significantly reduced its rate of deforestation as a result of domestic political processes and international attention. Civil society organizations had been clamoring for coordinated action between federal, state, and municipal governments to tackle deforestation; many businesses in sectors responsible for deforestation were shifting toward land-intensive rather than land-extensive practices; and the administration of Lula da Silva (2003 to 2011) took significant steps to enforce environmental protection laws after the appointment of Marina Silva as Environment Minister in 2003 (Abranches 2014).

Forest conservation was originally excluded from international agreements to address climate change for a variety of reasons (La Viña and de Leon 2014). Nonetheless, international discussions continued and the issue gained traction under the framework of REDD+ (Reducing Emissions from Deforestation and Forest Degradation, "plus" forest conservation, sustainable management of forests, and enhancement of forest carbon stocks). By 2007, it was clear that a substantial amount of human-caused greenhouse gas emissions could be attributed to deforestation – 17 percent according to the International Panel on Climate Change (IPCC 2007).

At the 2007 United Nations Climate Change Conference in Bali, Brazil's Environment Ministry proposed a fund to be managed by the National Bank for Economic and Social Development (BNDES) which would enhance domestic efforts to promote forest conservation. At the same conference, the Norwegian government launched its International Climate and Forest Initiative (NICFI) to support the REDD+ agenda and encourage action to reduce global greenhouse gas emissions (GHG), including through performance-based bilateral agreements. This propitious timing meant that Norway and Brazil were both prepared to enter negotiations with complementary positions and quickly reached agreement on creating The Amazon Fund the following year.

Norway pledged US$1 billion to the Amazon Fund for the period 2008-2015. Norway committed to contribute to the Fund on the basis of Brazil's success in reducing greenhouse gas emissions from deforestation.[16] Payments by Norway to Brazil are based on a fixed price of US$5 per ton of avoided CO_2 emissions from deforestation. Deforestation rates are calculated from data collected by Brazil's National Center for Space Research (INPE) using satellite imagery. The area of 'avoided deforestation' is the difference between the current year's deforested areas and the average between 1996 and 2005. The agreement calls for this 10-year average to be shifted forward every five years, creating a new reference level (baseline). The associated greenhouse gas emissions are estimated using a single, conservative emissions factor of 100 tons of carbon per hectare. The amounts reported by Brazil are reviewed by a Technical Committee appointed by Brazil which includes both international and domestic experts. Public access to data also assures that independent researchers and civil society groups can develop their own estimates and detect discrepancies in the official reports.

Brazil designed the Amazon Fund explicitly to recognize *past* performance – that is, reductions in deforestation that they had already achieved – precisely to avoid the political appearance of changing environmental policy to suit international donors. The international funds, however, were to be channeled through the Amazon Fund, managed by BNDES, into environmental programs in the Amazon (and up to 20% in forest biomes in other parts of Brazil or other tropical forest countries) that would further contribute to forest conservation. Linking payments to the Amazon Fund was useful for both domestic and international political and fiduciary reasons. It also made it possible to share the benefits of reducing deforestation with local groups and gave political prominence to Brazil's national environmental efforts, domestically and internationally.

But linking Norwegian payments to the Amazon Fund also resulted in delays. Based purely on the massive reduction in deforestation, Norway's pledge of about US$1 billion could easily have disbursed the first year, but several obstacles arose. First, Brazil had to pass a law to exempt the Norwegian contributions from a financial tax. Then, BNDES had to develop procedures for applications to the Fund along with criteria for approval. Brazil and Norway also had to address a chicken-and-egg problem: BNDES does not review proposals for operations unless it has money available to fund them; whereas Norway was planning on disbursing in concert with the pace of approvals. As a result, Norway paid about US$127 million between the beginning of the agreement in 2008 and June 2013. Once these issues

[16] Details about the arrangement can be found on the web pages of the Norway's International Climate and Forest Initiative (NICFI). See Norad 2014b.

were fully resolved, Norway disbursed an additional US$627 million in the last quarter of 2013. As of December 2013, approximately US$750 million of the original $1 billion pledge had been disbursed.[17]

The Amazon Fund is managed by BNDES but is governed by a Guidance Committee which is chaired by the Environment Ministry and includes representatives of local government, other national ministries and civil society (Forstater et al 2013). The Guidance Committee establishes priorities for project funding although BNDES decides which projects will be funded.

The Amazon Fund: Key Features of RBA Agreement

Results Level: Outcomes? Of the cases analyzed here, the Amazon Fund most fully realizes all of the six design features of interest. Transfers from Norway to Brazil are designed to reward an outcome – reduced emissions from deforestation – based on satellite imagery that measures changes in deforested area.

Payment Function: Incremental? The agreement provides for Norway to pay Brazil a fixed amount for each "unit" of progress: US$5 per ton of CO_2 emission reductions. It is clearly designed to pay incrementally in proportion to progress. Given Brazil's rapid progress and administrative issues, the actual disbursements are justified by large reductions in emissions but, in practice, are not actually proportional to those achievements.

Recipient Discretion? Brazil had (and has) full discretion over the way it reduces deforestation. This was a key feature of the agreement with Norway which explicitly recognized that payments would be rewarding actions Brazil had already undertaken. Once emission reductions are verified, the use of funds that Norway disburses is governed by the agreement in a way that preserves substantial Brazilian discretion. This is possible largely because Norwegian concerns about fiduciary risk and governance were addressed by Brazil when it designed the Amazon Fund. The Agreement between the two governments requires that they abide by agreed general principles, including for resolving problems should they occur.

Some issues arose related to the allocation of payments once they were rewarded. While Norway was ready to make payments to Brazil on the basis of verified reduced emissions, the Amazon Fund – which was set up to receive those contributions – was not prepared to accept them because BNDES wanted to ensure that proposals for funding from local

[17] Amazon Fund 2014.

governments and NGOs met their fiduciary and technical standards.[18] It is unclear whether the procedures for local groups to secure funding are necessary or simply burdensome, but in terms of the functioning of the RBA program, Norway has fulfilled its commitment to pay Brazil for results at the national level, while discretion in the subsequent allocation of those reward funds is governed by the Guidance Committee whose members are entirely Brazilian.

Credibility? The Norway-Brazil agreement enjoys the advantage of an outcome indicator that is easy to verify through satellite technology. The Technical Committee is appointed by Brazil, which could compromise its independence but the remote sensing data is in the public domain and any significant deviations would be readily highlighted by Brazilian and international civil society groups. In practical terms, the magnitude of the reductions in deforestation has made precise measurement less important when certifying emissions reductions for the sake of determining payments.

Transparency? The Norwegian International Climate and Forest Initiative (NICFI) has made information about its REDD+ programs and total disbursements readily available on its website and through published evaluations.[19] A 2007-2013 "real-time evaluation" of NICFI in particular includes information about the process, issues that arose, and overall results. The Brazilian government has also posted considerable information about the agreement on the web. The Amazon Fund website managed by BNDES currently has copies of the formal agreements between Norway and Brazil, a full schedule of when disbursements were made, a list of projects in various phases of screening, review, approval, and implementation. Brazilians working on forest conservation issues are familiar with the existence of the agreement and its linkage to the Amazon Fund. However, details of the linkage between disbursements and emissions reductions are not widely understood, even among this well-informed group (Birdsall, Savedoff and Seymour 2014).

Though information on annual deforestation rates, emissions reductions and payments can be found on both Norwegian and Brazilian official websites,[20] greater transparency in terms of making information accessible to all citizens would require a table or shorter document that specifically lists the key performance and payment information, something which we were unable to find in the public domain. However, satellite data made public by Brazil and

[18] This led to a debate about whether these funds should count as official development assistance (ODA). Birdsall 2013 lays out reasons why external finance to reward slowing deforestation should be counted as ODA.

[19] Norad 2014a and Norad 2014b.

[20] Norad 2014a; Amazon Fund 2014.

other countries is still a powerful source of information – once analyzed by researchers and civil society organizations – that influences the domestic and international politics of deforestation.

Payment Amount? Norway's funding is complementary to Brazil's own spending on reducing deforestation, making it more of a reward rather than a subsidy or reimbursement. While US$1 billion over five years is a lot of money in relation to specific forest conservation programs in Brazil, it is dwarfed by the government's overall spending on environmental programs and enforcement of environmental laws, and certainly relative to total public sector spending. Brazil has the resources and capacity to reduce deforestation rates without the Norwegian funding, yet Brazilian officials also maintain that the Amazon Fund agreement is useful to them by providing financial resources and political support (Birdsall, Savedoff and Seymour 2014).

Predominant Theory: Attention, Accountability and Recipient Discretion. The Norwegian agreement with Brazil to finance The Amazon Fund seems designed to operate through attention to outcomes, accountability and recipient discretion. The amount paid is low relative to the true value of total emission reductions from reduced deforestation and Brazil had its own domestic political process to motivate progress, making pecuniary interests an unlikely motivating factor. The design was driven initially by Brazil rather than Norway, accounting perhaps for the large degree of autonomy preserved in the design. The fact that Brazil had already begun to reduce deforestation and made a proposal which addressed fiduciary, social and environmental concerns facilitated Norway's willingness to come to the table (Hermansen and Kasa 2014). The agreement draws attention to Brazil's performance in reducing deforestation in ways that strengthened the salience of deforestation policies within the government as well as by holding the government accountable to civil society.

Case Study 3: Secondary Education in Ethiopia

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. Over the last two decades, British governments have also given increasing attention to improving accountability to taxpayers by paying for results, whether in the National Health Service or the programs of the Department for International Development. This trend toward demonstrating results and increasing the "value for money" from public programs was given added impetus after 2010 by the Cameron government which – inspired in part by Cash on Delivery Aid – proposed to pilot a number of results-based programs. They argued that paying countries in proportion to outcomes would make aid more effective by giving recipient countries the discretion to choose the investments they wanted to make for their own development and increase accountability by clearly demonstrating to British taxpayers the results of British aid spending (One World Conservatism n.d. and DFID 2010).

DFID has since piloted a new series of performance payment programs. DFID negotiated an RBA program with the government of Ethiopia, making it the first program to be directly inspired by the idea of Cash on Delivery Aid.[21] This RBA pilot, launched in 2012, aims 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). It is a £30 million three-year program, with disbursements of up to £10 million each year. Funds are transferred directly to the Ethiopian Ministry of Education, and are additional to existing support that DFID provides to the education sector.

The program provides additional incentives payments for girls who sit and pass the exam compared to boys and for students in Ethiopia's emerging regions[22] to address equity. Payments are 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 (See Table 3).

[21] DFID published a business plan in May 2011 announcing its intention to pilot three results-based aid or Cash on Delivery Aid programs. A detailed description of its first pilot, an education program in Ethiopia, including the full business case can be found at: http://devtracker.dfid.gov.uk/projects/GB-1-202989/

[22] Ethiopia is a federated country with substantial authority exercised by nine Regions and 2 special districts. Ethiopia's poorest Regions, known as the emerging Regions are Afar, Somali, Benishangul, Gomuz and Gambella.

Table 3. DFID Ethiopia RBA Program Payment Structure

In emerging Regions DFID pays:

	… for each additional student sitting the exam	plus an additional … for each of these students who pass the exam
Boys	£75	£75
Girls	£100	£100

In non-emerging Regions DFID pays:

	… for each additional student sitting the exam	plus an additional … for each of these students who pass the exam
Boys	£50	£50
Girls	£85	£85

Source: Birdsall and Perakis, 2012.

The idea of paying for students who take the exam is based on the concept of "assessed completers" as a suitable outcome indicator for COD Aid programs in the education sector (Birdsall & Savedoff 2010). The rationale is that paying for students that take an exam in the final year of school is a good proxy for completion, and resulting test scores provide information about the quality of the education. Furthermore, paying for students that sit an exam, rather than students who pass, minimizes perverse incentives to 'teach to the test' or manipulate exams to inflate scores.[23] DFID initially proposed paying for assessed completers but during negotiation, the Ethiopian government requested that the agreement include an incentive for improved performance on the grade 10 national examination by providing an additional payment for students who actually pass the exam.[24]

The Ethiopia RBA pilot falls under a larger umbrella of programs implemented by the UK government which they call Payment by Results (PbR). Its key features include: payments based on results; recipient discretion to decide how results are achieved; and verification of

[23] This idea was initially proposed and explained in Crouch, L and J Mitchell, 2008. It was applied in Birdsall and Savedoff, 2010.

[24] Authors' interviews.

results as the trigger for disbursement.[25] DFID's agreement with the government of Ethiopia is that funds are linked to sitting and passing results of the grade 10 exam, and the Ethiopian government is in charge of determining strategies to improve performance as well as how ex-post payments will be used, with DFID providing support only when asked. As such, the only conditions in the agreement between DFID and the government of Ethiopia are that funding is linked to results which are independently verified and that both parties must adhere to the Partnership Commitments that apply to all UK aid. These Partnership Commitments establish basic common principles, such as a commitment to respect human rights and a provision that DFID may interrupt planned disbursements in exceptional circumstances if DFID staff were concerned that the partner government is not fulfilling those commitments.

DFID contracted an independent consulting agent to verify results reported by the government. The agent analyzes data on enrollment, retention and pass rates from Ethiopia's National Agency for Educational Assessment and Examinations to check for consistency. It also visits a sample of schools in each region to compare national reports with the schools' reports and verifies that nationally reported results are accurate (Birdsall and Perakis 2012).

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. According to a review conducted by an independent consultant for DFID, these results amounted to a reward payment that was below what the Government of Ethiopia had expected; the reward payment was about £900,000, reflecting that some positive results had been achieved, but was far below the £10million ceiling established for the project's first year (DFID 2013a).

The lower than hoped for results were not surprising given the limited time frame of the project and because the Regions of Ethiopia did not actively respond to the RBA agreement during the project's first year.[26] However, these results presented a challenge to DFID due to "[the Ethiopian] government perception that unspent funds from DFID's allocation is a

[25] From DFID, 2013b. A subsequent strategy document has broadened this definition of PbR to incorporate any ex-post payments for pre-agreed results (DFID 2014b).

[26] The final RBA agreement between DFID and the government of Ethiopia was signed in February 2012, and students took the grade 10 exam, which was the basis of the first year's results payments, in May/June 2012. This gave limited time for any new strategies to affect education sector results.

loss/punishment" (DFID 2013a). DFID staff debated whether or not to adjust the per-pupil payment amounts in light of the first year's lower than expected results, but ultimately did not make any changes to the RBA payment structure in response to these concerns.[27] DFID disbursed the first year reward payment according to the original commitment in early 2013. DFID agreed, however, to roll over unspent funds into future years of the program.

In the second year, results improved significantly. In 2013, 38,490 more students sat the grade 10 exam than in 2012 and 37,669 more students passed. These figures resulted in a payment of approximately £5.6 million, much more than the £900,000 of year 1, but still less than the £10 million per year ceiling. The Ethiopian government chose to combine year 1 and year 2 results payments and allocate them to regions, partly based on each region's incremental improvements in the number of sitters and passers and partly based on each region's population.

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. As the program is nearing completion of its third year, indications are that interest in the program has continued to increase as people at the regional and school levels have become 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.[28]

In addition to the £30 million allocated to pay the Ethiopian government for results achieved, DFID has allocated £1.1 million to cover the costs of the independent results

[27] Authors' interviews with DFID staff and consultants.

[28] For example, we learned of a case in which a previously existing program paid for plasma screens for secondary schools, but in one region the screens had not been used due to lack of electricity. Some schools chose to use reward funds from the RBA agreement to buy generators so that they could operate the screens. This would not have happened if discretionary resources had not been made available first at the national level through the RBA agreement. From authors' interviews.

verification and of evaluating the RBA pilot, which are two separate functions.[29] An ongoing evaluation of the program is assessing changes in education access and quality, as well as the effect of the RBA modality on the relationship between DFID and the government of Ethiopia.

Secondary Education in Ethiopia: Key Features of RBA Agreement

Results Level: Outcomes? The indicator of 'sitters and passers' of the lower secondary exam is not a perfect outcome indicator since it doesn't completely capture "improved student learning" but - particularly given available alternatives - it is close. It provides a clearer measure of progress in the education system than more typical performance measures like repetition, dropouts or enrollment. Linking disbursements to sitters and passers differentiates this program from almost all other education aid programs which tie disbursements to specific activities and inputs or sector programs that measure performance by enrollment.

The outcome indicator could have been more straightforward if it had paid only for exam sitters, particularly because passing scores for the exam are not consistent from year to year and are not directly linked to competency. Instead, Ethiopia's grading system is norm-referenced to determine who can attend upper secondary schools and the number of passers depends on openings in the next level of schooling. Despite this, the indicator is still closely linked to the number of students reaching the end of lower secondary school and, as such, is paying out against a well-defined outcome that is simple to understand and explain.

Payment Function: Incremental? The program pays for each additional boy or girl who sits the lower secondary exam and for each boy or girl who passes the exam. Therefore, it rewards the Ethiopian government in proportion to the progress made on these measures. For practical reasons, DFID has established an annual ceiling for the program but this is not a performance target, per se. The ceiling has not affected the program yet insofar as progress in Ethiopia has not been large enough to make it binding.

Recipient Discretion? DFID is committed to testing a program which has 'recipient discretion' as one of its key features. DFID has put in place the minimal necessary safeguards via its Partnership Commitments but has not imposed any other conditions on the Government of

[29] DFID breaks down its annual budget commitments and disbursements on the Development Tracker website, 2014b.

Ethiopia or plans for how services should be delivered or funds spent. This does, however, appear to be the aspect of the program from which most questions and challenges have arisen. For example, the Government of Ethiopia did not initially take full advantage of this discretion, showing that even after a donor agency has committed to this feature, partner governments will not automatically respond. This is a good reminder that implementing a novel aid modality takes time to understand and can only then begin to affect behaviors and results. Moreover, it has been difficult to test the extent to which recipient discretion is making a difference because the RBA program is a small program among much bigger aid programs, supported by the UK and others, which are not results-based (see below under "Payment Amount").

Credibility? DFID has contracted an independent group to verify results of the program each year and is linking disbursements to these independently verified results. The verification primarily involves a review of the quality and reliability of government reports by a consulting firm which is *institutionally independent* of the British and Ethiopian governments. The verification process could have been stronger if it had incorporated sources of *independent information.* For example, the verification could re-test a random sample of schools and compare results or the verifying agents could go to a sample of households to check whether students listed as taking the exam actually existed. In this specific case, the verifying agents checked a sample of student names listed as having sat the exams against the lists of students enrolled in their schools and found enough correspondence to make a judgment that the government reports were accurate.[30] This appears to be a credible approach for obtaining unbiased results.

Transparency? In terms of accountability to British taxpayers, DFID has provided key information on its website, especially in terms of the program's budget and design. However, considering that this is a flagship program which has drawn a lot of interest from other donors, it is difficult to find publicly available information about its progress. For a pilot program like this, it should be possible to find a single-page summary of when the program began, annual results and annual payments. However, this information can only be found with difficulty embedded within documents like the business case.[31] Our experience is that annual reviews and evaluations are not updated regularly on DFID's website and often show up, if at all, with significant delays.

[30] Authors' interviews with consultants.
[31] Project documents are available on the Development Tracker website, DFID 2014b.

In Ethiopia, the government had disseminated limited information about the program at the time we compiled this information. When first year results were reported and the RBA payment was calculated, the Government decided the payment (about £900,000) was too small to distribute to the regions and chose not to publicize the fact. There appeared to be little knowledge about the program outside of the education ministry at that point. The Government made a more elaborate announcement after the second-year results, at an October 2013 Annual Education Sector Review conference, where it communicated total reward amounts from the program's first two years. According to more recent reports, regional representatives now appear to be aware of the program and are engaged in related activities.[32]

At this early stage, it does not appear that domestic transparency has contributed to increasing accountability to citizens. For citizens to hold their governments accountable, they need information about the agreement and how the program is being implemented. However, the context is not conducive to such accountability. While Ethiopia's government is recognized for expanding social services and contributing to economic growth, it is also criticized as autocratic and politically repressive (Epstein 2010). The ability of this results-based education program to improve transparency and accountability in such a context would be difficult at best.

DFID staff have debated just how far the funding agency should go in encouraging the government to disseminate results to its citizens. This includes questions around what specific mechanisms governments can use to promote transparency and accountability; how funders can know whether these mechanisms are working; and whether funders can impose requirements for transparency without compromising the "hands-off" nature of the RBA agreement.

Payment Amount? The program clearly complements other financing, both foreign and domestic. It was additional to more than £60 million (US$102 million) that DFID spent on education in Ethiopia in 2010-2011 through two multi-donor aid programs, the General Education Quality Improvement Program (GEQIP) and Protecting Basic Service Grant (PBS). Total foreign aid commitments for the education sector in Ethiopia were US$385 million in 2012 (OECD 2014), and the Ethiopian government spends a high percentage of its own total expenditures on education (24% in 2010) (Nganwa 2013 and World Bank

[32] According to DFID Annual Review undertaken December 2013 and presentation by Solomon Shiferaw, Director of Education and Management Information System, Ethiopian Ministry of Education (Shiferaw, 2014).

2014). Even if the program were to disburse its full allocation of £10 million (or about US$17 million) per year, it would still be only a small supplement to those other funding channels. The program effectively functions as a small bonus for results the government achieves. The program's importance resides less in its total amount than in the attention it draws to results and the way that it gives the government discretion over implementation and use of the funds.

Predominant Theory: Attention. DFID's business case for the Ethiopian program refers to all four theories but drawing political and managerial attention to results and recipient discretion are most prominent. The RBA program seems too small (capped at £10 million per year) to be driving change through pecuniary interests, especially relative to other external sector funding. Although schools were beginning to demonstrate a pecuniary interest in getting a share of reward funds towards the end of the program's initial three years, we have not seen indications that the education ministry, with whom DFID has the RBA agreement and for whom the total amount was small, has shifted priorities or actions in order to get more funding. The Ethiopian government has disseminated very little information about the program domestically, so accountability to constituents also has not been significant.

Recipient discretion was featured prominently in DFID project documents and in high-level comments by former Secretary of State Andrew Mitchell. However, given the context and timeframe, the value of greater recipient discretion has not really been tested and it may be impossible to do so. Ethiopia's main strategy for expanding schooling and increasing secondary school completion is part of a sector-wide program funded by the UK along with the World Bank and other aid agencies. Evidence that the Ethiopian government was exercising its discretion only began when the program was already half over.

Thus, it appears that the RBA agreement is driving change primarily by making the outcome more visible to the government. The fact that payments are *linked* to the number of sitters generated high-level attention when payments were lower than expected but did not create significant financial difficulties for Ethiopia. This is supported by evidence of Ethiopian government reactions to lower than expected results as described in the concurrent evaluations.

In 2014, DFID and the government of Ethiopia agreed on plans to extend the RBA pilot by two years to determine if additional actions and results will be spurred by the RBA program and to evaluate the process over a longer time period. This will provide an opportunity to assess whether, in addition to generating visibility and high-level political attention, the Ethiopian government exploits its discretion under the agreement and takes actions which contribute to progress.

Case Study 4: Salud Mesoamérica 2015

Salud Mesoamérica 2015 (SM2015) is a five-year initiative that aims to reduce equity gaps in maternal and child health across Central America. It was launched in 2010 and is being funded by the Gates Foundation, the Carlos Slim Health Institute and the Spanish government and is managed by the Inter-American Development Bank (IDB). It is a results-based aid pilot in which a portion of funding to governments is contingent upon progress toward health-related development goals. The program operates in seven countries – Belize, Costa Rica, El Salvador, Guatemala, Honduras, Nicaragua, Panama – and Mexico's southeastern state of Chiapas.

SM2015 aims to create sustainable improvements to health systems by drawing attention to problems of health equity and creating incentives to reprogram domestic resources towards key services for vulnerable populations (Glassman 2013). Countries receive grants, but are also required to finance part of the program costs. If goals are achieved, countries are reimbursed for half of the amount they contributed.

The program has three core features: (1) funding can only be used to implement certain evidence-based, cost effective interventions in the areas of nutrition and maternal, neonatal, and child health care; (2) activities and performance targets are directed only to the poorest 20% of the population; and (3) country programs must address supply- and demand-side barriers to equitable health access. [33] In addition to direct interventions, measures to address barriers can include policy reforms, changes to the distribution of health spending, or technical support. SM2015 also includes a policy dialogue component to identify constraints and encourage sharing of evidence and best practices at the national and local levels. The program emphasizes learning through policy dialogue and impact evaluation, and feedback to improve current interventions or new programs.

The program provides conventional investment funds alongside a performance component. Funders committed about US$150 million to the program of which approximately US$25

[33] Information about the program's design comes from the Project Document for SM2015, March 2011 and interviews with IDB staff.

million could be paid to reward performance. Countries receive an initial grant, the investment tranche, to cover input costs. They are required to provide counterpart funding, which is negotiated for each country and reflects its ability to contribute. Typically, the counterpart contribution is 50% of the total financing for each country program, or roughly equal to the investment tranche. At the end of each funding period,[34] the country can receive a performance tranche equal to 50% of its counterpart contribution if targets are met. The IDB pays these funds directly to the finance ministry of each country, which must then show that the funding will enter the budget of the Ministry of Health. Once reward payments are made, they can be used for any legitimate purposes in the health sector.

The performance targets are negotiated for each country taking into account health trends in the country and projecting the expected effect of planned interventions. The country receives the performance tranche if a weighted average of 80% of targets are reached. About 8-12 indicators are selected for each country, five of which are common across countries in order to make regional comparisons. The indicators include such things as contraceptive use, anemia prevalence, attended deliveries, DTP3 vaccinations, and antenatal care visits.

Independent verification is an important part of the program, not only to assure credibility but also as part of its aim to increase the use of evidence for pro-poor health policies and increase the health system's accountability for expanding coverage and quality. The first performance payments will be based on the findings of a facility survey; while the second round of payments will be disbursed on the basis of results measured in a facility survey and a household survey. By relying on surveys, the funders and recipients will not only have greater confidence that progress was accurately measured but they will also have an independent check on the quality of administrative data for the same indicators. The total cost of monitoring and evaluation for the program – including the independent verification process, an evaluation of interventions, and a process evaluation – amounts to approximately 12% of the program's total budget. At the end of 2014, the verification process for the first funding period of the project had begun. An article in *The Economist* reported that three countries met their targets (Honduras, Nicaragua and El Salvador), three did not (Guatemala, Belize and Chiapas, Mexico), and results were not yet available for Costa Rica and Panama (The Economist 2014). The SM2015 website maintained by the IDB has begun to publish country-level information about program objectives and total disbursements, but

[34] The first funding period of 18 months was extended to 24 months, according to interview with IDB staff.

no information about results, verification, or performance tranche disbursements was publicly available when this paper was completed.

Salud Mesoamérica 2015: Key Features of RBA Agreement

Results level: Outcomes? SM2015 is a large regional initiative, for which a portion of program funding is paid on the basis of measured results (up to 50% of counterpart funding, which is at least 50% of input-based funding provided by the project's donors). The indicators directly measure outcomes (for example, reductions in prevalence of anemia) or outputs that are associated with desired outcomes (for example, increases in neonatal care provided within 24 hours of birth). The program does not pay for a single outcome but instead uses 8 to 12 outcome and output indicators which measure important features of the countries' health systems in the areas of prenatal, neonatal, and maternal health.

Payment function: Incremental? SM2015 payments are made on the basis of targets rather than in proportion to progress. The performance tranche is binary; countries only receive it if they achieve a weighted average of 80% of the targets for all indicators. This simplifies the payment structure and diversifies risks across several indicators. However, this model still exhibits the normal problems associated with targets. First, SM2015 set targets for each country based on research but are likely to be biased by the normal tendency of development programs to overestimate likely gains. Second, if countries fail to reach the threshold, anyone interested in making disbursements happen is likely to question the precision of the measures or the reasonableness of the original targets and could potentially exert pressure to waive conditions. Finally, the lack of payment for countries that miss the target can also reduce interest in the program in ways that would not happen if some payment, in proportion to progress, were made. In practice, however, concurrent evaluations and interviews suggest that none of these problems have emerged in the program.

Recipient Discretion? SM2015 was developed collaboratively between the funders and the recipients; however, public documents suggest that recipients have no more discretion in this operation than in conventional IDB grants and loans. Countries are in charge of implementing programs and are required to contribute initial financing as a part of the initiative's commitment to country ownership. The goals and activities in project documents are aligned with national health policies. Beyond this, it is not clear what flexibility countries have to deviate from the predefined strategies, activities and plans. The short timeframe for the funding period of the project (18-24 months) also makes it less likely that countries would be able to develop and execute their own strategies, even though the program did allow for adjustments within programs and particularly between the first and second funding period. Overall, the performance component seems designed not so much to provide room for recipients to experiment and innovate as it provides information, focus, and attention to increase the likelihood of full implementation.

Credibility? The performance component of SM2015 pays for progress on indicators that are measured by independently administered facility and household surveys. The results from the first round of facility surveys were not complete until the end of 2014 and were not publicly available so it is too early to assess implementation. Nevertheless, as designed, this program has very strong independent verification, both in terms of the institutional independence of the organization responsible for the surveys (IHME) and the independence of the information collected.

The investment portion of the program still operates like conventional input-based funding. As a result, measuring results has been laid on top of conventional reporting requirements for investment grants and loans. In this regard, the program may have lost an opportunity to reduce the reporting burden on countries.

Transparency? Information about the program is generally difficult to find. The IDB maintains web pages with country level information, including loan agreements and initial budgets, but this information is not consistent across countries and does not include accessible information about project design details or implementation progress. A strong test of the program's transparency will occur once initial performance tranche disbursements are made, at which time it will be possible to judge how well or poorly these key pieces of information are disseminated. Web access to information in a workshop at which results were presented in November 2014 was password protected at the time this paper was being completed.

The program's regional nature generates an additional channel for dissemination and accountability between countries. By holding regular meetings to report progress on specific indicators, the program draws the attention of high-level policymakers to how they are faring relative to their peers in other countries. This might be an effective mechanism for generating peer pressure among countries but it is not going to generate the same kind of accountability created by disseminating information to the public.

Payment Amount? The program is complementary to other health sector funding available in Mesoamérica, particularly countries' own resources, and therefore functions like a performance subsidy. The program requires most countries to contribute substantial funding from their own domestic budgets and the performance tranche is a share of these contributions. In the case of the lower-income countries (Honduras, Guatemala and Nicaragua), highly subsidized IDB lending is used for much of the counterpart contribution.

Predominant Theory: Attention. Documentation for Salud Mesoamérica stresses the involvement of each country in developing its own plan and coherence with domestic national health strategies which would suggest a great deal of recipient discretion. In practice, the ex-ante planning constrains countries to spend on the kinds of health sector activities that are established in the programs and co-financed up-front by funders. The amount committed is a fraction of total costs, and we have no information that would suggest pecuniary interests are the primary motivating factor. Information regarding the targets and progress are accessible to the public but only with difficulty and not in a timely manner, suggesting that change will not be induced by accountability to constituents.

Instead, the results-based element of this program puts greater attention on results and provides feedback to improve management and implementation. It is generating new sources of information from survey data to assess health care services and health conditions. Regional mechanisms for sharing information about progress seem to have played an important role, in part by creating some pressure to demonstrate improvements among peers, reinforcing our conclusion that greater visibility of results has driven progress. The shared data, along with a strictly monitored action plan, are providing information and focus for participating countries to meet deadlines and increases the likelihood of full implementation.

Cautious Implementation and Few Concerns but What Really Changed?

Reflecting on these four cases suggests, first of all, that common concerns about RBA – such as corruption risks and policy distortions – do not appear to have materialized. The case studies also demonstrate that overall implementation of RBA has been cautious. In part as a consequence of this caution, the main theory embodied in these programs is making outcomes visible and drawing attention to them. The Amazon Fund is exceptional for being designed to work through accountability and recipient discretion along with using visibility to focus political and managerial attention. For the most part, *ex post* payments for outputs and outcomes are being introduced on top of other conventional aid mechanisms rather than representing a distinct and dramatically new way of doing things.

Concerns That Did Not Materialize

Though these RBA agreements are few in number, none of them display the kinds of problems typically raised by critics. In particular, none of the agreements have been the subject of accusations regarding corruption. Nor do they appear to have distorted public allocation decisions, sacrificed long-term goals in the pursuit of short-term gains, or entailed costly verification and monitoring.

First, no allegations have emerged regarding *diversion of funds to improper uses*. In the case of GAVI, researchers have presented evidence that countries may have been overpaid relative to the results they achieved but no one has accused governments of misusing those overpayments. This may partly be a consequence of GAVI's restriction that reward payments be applied within the health sector. Similarly, for the Amazon Fund and the Ethiopian Secondary Education program, requirements that funds be applied within particular sectors or in line with the recipients' internal fiduciary mechanisms may be providing adequate assurance and control to avoid such problems. RBA programs may, in fact, be less vulnerable to corruption than conventional programs because payments only occur when outcomes improve – something which is difficult to achieve unless funds are used appropriately (Kenny and Savedoff 2014).

Another concern that is commonly levied at RBA programs is that high-powered (and perverse) incentives meant to appeal to pecuniary interests will *distort public policy decisions*; yet none of these cases suggest this has occurred. Countries do not seem to have diverted funds from other worthwhile public programs in order to capture RBA funds, whether that meant reducing spending on other health programs to capture GAVI ISS payments, taking funds from primary schooling to encourage secondary completion in Ethiopia, or reducing forest conservation efforts in difficult areas to concentrate on combatting deforestation in places that are easier to control. The related question of unintended consequences – such as violating human rights or dispensing with basic social or environmental standards – has also not materialized in any of these cases.

A third common concern is that RBA programs are *too short-sighted*; that funders and recipients are likely to pay undue attention to short-term gains at the expense of longer term sustainability. The easiest way to address such a concern in an RBA agreement is to make sure it is a long-term agreement (e.g., at least 10 years); however in practice, most agreements are quite short. GAVI's program and SM2015 were both programmed to last five years, and the Ethiopian secondary education agreement was originally for only 3 years, although it has been extended. The Amazon Fund was originally pledged for a five-year period but has also been extended.

Despite these relatively short time frames, governments do not seem to have sacrificed long-term development in the pursuit of short-term gains. Ethiopia continued to follow its long-term education sector strategies, which are articulated in national policy documents and supported by a mix of foreign aid programs. SM2015 was explicitly negotiated with participating countries on the basis of their own national health plans and, despite setting reward payments on a short time frame, is supporting activities that are embedded in broader health sector strategies. The Amazon Fund was created to reward Brazil for ongoing progress in reduced deforestation which was part of a broader domestic political shift favoring sustainable development; therefore, again, the short-term payment structure does not seem to have created a distortion in policy toward shorter-term actions. The GAVI program also does not seem to have created short-term distortions, although in many cases it also has not been able to assure the long-term financial sustainability that domestic immunization programs require.

A fourth concern is that RBA programs involve *large transaction costs* – in terms of design and negotiation as well as generating baselines and verification procedures for outcome indicators. People involved in implementing these programs have confirmed that designing and implementing an RBA program does require significant amounts of staff time and resources but not necessarily more than those required by conventional aid programs. The biggest "investment" is often related to explaining and convincing participants that the RBA agreement is really a different way of implementing a program. This is not true of the Amazon Fund – which was proposed by Brazil – but is certainly true of foreign aid programs like SM2015 and the Ethiopian secondary education project which are negotiated and signed like other international aid agreements. The direct costs of measuring and verifying outcome indicators has not turned out to be large in the cases of The Amazon Fund or the Ethiopian Secondary Education project. SM2015 explicitly incorporated facility and household survey data not only to have independently verified progress measurements but in order to have reliable data for public policy decisions.

An issue that often arises with independent verification – and certainly arose in the GAVI case – is the perception that a funder who requires an independent check on government self-reporting is implying a *lack of trust* in the government or undermining the development of the country's own data systems. These misconceptions can be avoided by making clear that the objective of independent verification is to get the most accurate and reliable data possible on development outcomes. The process of understanding any discrepancies between government and independent reports helps improve overall data quality and strengthens administrative capacity while avoiding overpayment *or* underpayment. We see no reason for a tradeoff between obtaining clear independently verified measurements of results and the goal of improving countries' capacity to monitor results. If anything, the two processes complement one another.

Cautious Implementation

One of the most important and surprising findings from this research was the limited number of RBA agreements that could satisfy two basic features – (1) *ex post payments to a national or local government for (2) something that measures or is a proxy for an output or outcome*. In particular, conventional provisions for addressing fiduciary risks in programs with *ex ante* funding are being applied to programs that pay *ex post*. That is, instead of verifying that an outcome was achieved to determine whether a payment was appropriate, these agreements still require recipients to document how funds are spent *after outcomes are achieved*. This approach implicitly requires recipients to work twice for the same payment – once to achieve the outcome and once to demonstrate the use of the reward payments. Applying conventional approaches of addressing fiduciary risk to the use of performance payments increases transaction costs, limits recipient discretion and, perhaps most critically, takes attention away from the credibility and importance of the outcome measure.

Another hindrance to the adoption of more innovative versions of RBA agreements might be termed bureaucratic inertia. Staff in funding and recipient agencies who are assigned to design and implement RBA agreements tend to have experience with conventional aid approaches. While staff may understand the potential of results-based approaches, they are still working within a set of procedures and managerial decisions that facilitate familiar programs and only adopt new approaches with difficulty. One reason that the Amazon Fund was framed and adopted so quickly was that key policy decisions in Brazil and Norway were made outside of traditional aid agencies (Hermansen and Kasa 2014). The Ethiopian Secondary Education pilot owes its existence to Cabinet-level political interest in Britain and identifying staff who were willing to experiment – similar proposals for other recipient countries failed to materialize when DFID's country staff had other priorities.

Despite the small number of RBA agreements, many funders continue to claim that they are making dramatic progress on such performance initiatives. Such approaches seem to have strong political appeal among constituencies who contribute funding. The demand for such results-based projects has understandably led funding agencies to re-characterize their more conventional programs as incorporating such performance rewards. So, for example, USAID describes its Fixed Amount Reimbursement Agreements (FARA) as pay-for-performance contracts, when they are used primarily to purchase small infrastructure (e.g., schools, clinics, and roads) and not outputs or outcomes. A similar re-characterization can be found in DFID's Payment by Results strategy which reports that 71% of contracts incorporate performance incentives (DFID 2014b). While this is probably true, it suggests that the definition of performance incentives has been broadened to incorporate very conventional contracts – such as those which pay for on-time delivery of research reports or supplies – rather than those which pay governments *ex post* for children vaccinated or hectares of protected forest.

Transparency and dissemination seem to have been a greater challenge for these programs than anticipated. One of the appeals of RBA agreements is that they can provide greater clarity to participants and external constituencies about just what is being paid to whom and for what. The clarity in explaining such programs has certainly resonated with those who fund the programs. GAVI, Norway, Britain and the IDB all emphasized the payment structure of their programs when seeking approval and funding from contributors. When it comes to information about the content of the agreements, most of these programs have also been forthcoming. But when it comes to annual reporting of results and payments, information is somewhat difficult to find. It has been difficult to find GAVI reports on country-by-country immunization coverage related to ISS payments. The results of the Ethiopian Secondary Education pilot could be summarized each year in a single page, yet they need to be extracted from within lengthy annual reviews. The Amazon Fund's results and payment information is the most transparent but this is facilitated by the nature of the verification process (remote satellite imagery) and the rapid pace of progress. Given the limited amount of information available, none of these programs (except possibly the Amazon Fund) could be generating change through accountability to the public; however, they could be motivating change by making outcomes visible to managers and implementers or through line-management accountability and peer effects.

Which Theory Is Being Tested?

In most cases, the pecuniary interest theory is the one most discussed at the time advocates are proposing RBA approaches; but in the design and implementation phases, those involved tend to give much greater emphasis to attention to outcomes, accountability and discretion.

In all four case studies, the link between payments and results has been used to distinguish the program from others of its kind. GAVI's ISS program was initiated at a time when financial incentives were rare among health care programs and had never been applied at the national level. Prior to the Amazon Fund, environmental aid programs had been focused on capacity-building and improved policies and plans, but never linked to environmental outcomes per se. Similarly, the Ethiopian Secondary Education pilot was the first RBA program approved by DFID and is distinct from sector budget support by linking payments so directly to independently-verified national level-indicators. SM2015 also used the feature of performance rewards to be part of the package that attracted contributors that included private philanthropies as well as bilateral support.

Though pecuniary interest was prominently discussed when characterizing and advocating for these programs, the direct impact of the financial incentive has been relatively unimportant in explaining the programs' dynamics. While the Amazon Fund supports programs that contribute to forest conservation, no one is arguing that the Norwegian money channeled through the Amazon Fund is the main cause of Brazil's rapid decline in deforestation. Rather the international debates over REDD+ and the prospects of international payments fed into and supported domestic political initiatives aimed at reducing deforestation. Evaluations of GAVI credit a wide range of factors for increases in immunization coverage – including logistical support, assistance in procuring vaccines, as well as technical assistance in planning and implementation – with relatively little mention of the pecuniary interest in the ISS payments. Ethiopia and the Mesoamerican initiative countries are also engaged in changes that are only marginally attributable to the pull of the money itself.

Attention to outcomes seems to be the most important feature of these RBA programs when looking at evaluation reports and talking to people involved in implementation. The link to funding made results visible and salient in ways that conventional funding approaches lack. For GAVI, SM2015 and Ethiopian Secondary Education, linking payments to outcome indicators drew the attention of politicians and officials, was useful for improving management decisions, and focused the attention of implementers on the purpose of their activities. Ironically, payments are not necessary to achieve this visibility of outcomes. In principle, politicians and managers could instruct staff to gather such information and respond to it. However, in each of these cases, staff and officials describe the linkage of payments to the outcome indicators as a novel and appreciated mechanism for drawing attention to program aims.

Accountability to domestic constituents is not prominent in most of these programs; it cannot be, given the lack of public information and discussion of annual results. The Amazon Fund is unique among these four cases because it does have such transparency and the deforestation rate is prominent in domestic environmental debates. In Brazil, the creation of the Amazon Fund and the Norwegian commitment to transfer funds for reduced deforestation is seen as providing political support to domestic groups that are working for forest conservation. The funds also represent a significant boost in resources for particular environmental initiatives.

Recipient discretion has been less of a hallmark for these RBA agreements with the exception, again, of the Amazon Fund. GAVI and SM2015 provide performance payments as a complement to upfront investment tranches. Each of these programs involves fairly detailed diagnostic work and planning before committing to pay for outcomes. During implementation, funders are also heavily engaged with recipients. The Ethiopian Secondary Education project is very hands off; however, it is implemented within the context of large education sector-wide programs which are themselves quite detailed in terms of action plans, reporting standards, and ongoing engagement with funders. The Amazon Fund is the only one of the four cases where the recipient has full flexibility to achieve progress as it sees fit – Brazil's main efforts to reduce deforestation are related to national and state policies, enforcement of laws, and public investment decisions. The programs supported by grants from the Amazon Fund are also run according to criteria and governance arrangements established by the Brazilians themselves. Consequently, from these four cases, it is impossible to say how useful recipient discretion is to success for RBA approaches.

Final Reflections on RBA Experiences

RBA focused on outcomes is a relatively underutilized and untested way of providing financial assistance to governments despite its substantial promise. The mix of theories which underlie RBA may have contributed to this timid start. RBA agreements are often advocated as providing strong financial incentives when, in practice, agencies and foundations value them more for their ability to draw attention to goals and progress. Recipients are often more interested in RBA programs because of the promise of greater flexibility when, in practice, funders' practices may be limiting recipient discretion. As a result, agencies and foundations tend to adapt their conventional approaches rather than design RBA agreements according to a new internal logic – the Amazon Fund being the one exception.

RBA agreements are less likely to be approved and less likely to be distinctive without substantial upfront effort to demonstrate that they are a fundamentally different way of working. This upfront effort involves not only designing good indicators, gathering baseline data, and establishing clear procedures for verification. It also requires that funder and recipient staff adopt new mindsets that significantly alter their relationships. In an RBA program that provides full recipient discretion, it is recipient government staff who initiate problem-solving, diagnostics, design and specification of technical assistance requests. In an RBA program that pays for independently verified outcomes, it is the outcome indicator that constitutes proof for funders that money was spent appropriately (rather than receipts for input purchases). The cases discussed in this paper all required time for the distinct logic of the program to be articulated as well as understood by participants.

The pressures to modify RBA agreements – once initiated – are quite strong. Because the modality is new, funders and recipients are likely to be surprised by the dynamics generated when results are first revealed. Sometimes progress exceeds expectations (as with the Amazon Fund); more commonly progress falls short of expectations (as initially occurred with GAVI and Ethiopian Secondary Education). If the potential benefits of RBA are to be realized and tested, it is important to stick with the basic principles of the agreement for a reasonable period of time.

Finally, if RBA agreements are going to be more successful than conventional aid programs, it will not result from "making" recipients do anything differently. The effect of pecuniary interests within something as complex as a recipient government tends to be oversold when advocating for results-based approaches. The visibility that RBA gives to outcomes – either to draw political and managerial attention or generate accountability to constituents – is more likely to distinguish these approaches from conventional aid programs. To the extent that RBA programs are designed to promote domestic accountability and full recipient discretion, they could significantly improve the chances for greater aid effectiveness and sustained development progress.

In sum, RBA focused on outcomes is still a promising approach for aid agencies and philanthropies. Indeed, initiatives to pay governments in relation to improved outcomes are continuing and expanding. DFID has promoted RBA within its new "Payment by Results" strategy and has published a list of its performance program pilots.[35] Norway is expanding its use of performance-based payments for reduced deforestation in partnership with other donors. The World Bank is linking payments to outcomes in more of its Program for Results operations (Gelb and Hashmi 2014). European donors continue to introduce different performance components in their budget support programs. Aid agencies and private donors are also experimenting with new ways of paying for outcomes, such as through Development Impact Bonds.

We know enough from the few experiences available to see that concerns with corruption, unintended consequences, short-termism, and additional costs have not materialized. We can also see that most initiatives are cautious adaptations of conventional approaches such that the potential benefits from greater recipient discretion and public transparency remain unrealized. Future experimentation may provide answers to these questions but for now, RBA remains a work in progress.

[35] https://www.gov.uk/government/publications/dfid-guidance-on-payment-by-results-and-spreadsheet-of-pbr-projects

References

Abranches, Sergio. 2014. "The Political Economy of Deforestation in Brazil and Payment-for-Performance Finance," CGD Climate and Forest Paper Series #10, Washington, DC: Center for Global Development.

Abt Associates. 2007. "Evaluation of the First Five Years of GAVI Immunization Services Support Funding." Prepared for the GAVI Alliance. Bethesda, MD. file:///C:/Users/cgd/Downloads/GAVI%20Alliance%20evaluation%20of%20ISS.pdf

Amazon Fund. 2014. "Donations". http://www.amazonfund.gov.br/FundoAmazonia/fam/site_en/Esquerdo/doacoes/. Retrieved January 13, 2014.

Andrews, Matt, Lant Pritchett, and Michael Woolcock. 2012. "Escaping Capability Traps through Problem-Driven Iterative Adaptation." CGD Working Paper 299, Washington, DC: Center for Global Development. http://www.cgdev.org/publication/escaping-capability-traps-through-problem-driven-iterative-adaptation-pdia-working-paper

Barder, Owen. 2012. "Complexity, Adaptation and Results." Views from the Center, Center for Global Development, September 7, 2012. http://www.cgdev.org/blog/complexity-adaptation-and-results

Birdsall, Nancy. 2013. "OECD Dispute with Oslo Has a Bad ODA." Views from the Center, Center for Global Development, September 26, 2013. http://www.cgdev.org/blog/oecd-dispute-oslo-has-bad-oda

Birdsall, Nancy and Rita Perakis. 2012. "Cash on Delivery Aid: Implementation of a Pilot in Ethiopia." CGD Note. Washington: Center for Global Development.

Birdsall, Nancy, William D. Savedoff and Frances Seymour. 2014. "The Brazil-Norway Agreement with Performance-Based Payments for Forest Conservation: Successes, Challenges, and Lessons." CGD Climate and Forest Paper Series #4. Washington: Center for Global Development.

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

Clist, Paul, and Arjan Verschoor. 2014. "The Conceptual Basis of Payment by Results." School of International Development, University of East Anglia. http://r4d.dfid.gov.uk/pdf/outputs/misc_infocomm/61214-The_Conceptual_Basis_of_Payment_by_Results_FinalReport_P1.pdf

Cordella, T. and Dell'Ariccia, G. (2007), Budget Support Versus Project Aid: A Theoretical Appraisal. *The Economic Journal*, 117: 1260–1279.

Crouch, Luis and Jonathan Mitchell. 2008. "Audit Options to Certify Results for a 'Cash on Delivery' Contract in the Education Sector." http://international.cgdev.org/doc/progressbasedaid/CrouchMitchell.pdf.

de Hennin, Carlo de and Harm Rozema. 2011. "Study on results-based programming and financing, in support of shaping the multi-annual financial framework after 2013." European Union Development Cooperation Program.

DFID (Department for International Development). 2010. "Full transparency and new independent watchdog will give UK taxpayers value for money in aid." Speech at Oxfam and Policy Exchange at the Royal Society. London. 3 June 2010.

DFID. 2013a. Annual Review Pilot Project of Results-Based Aid in Ethiopia. http://devtracker.dfid.gov.uk/projects/GB-1-202989/documents/. January 2013.

DFID. 2013b. "Payment by Results." Policy paper. Published on October 15, 2013 and archived on June 26, 2104. https://www.gov.uk/government/publications/payment-by-results.

DFID. 2014a. "DFID's Strategy for Payment by Results: Sharpening incentives to perform." Policy paper. Published on June 26, 2014.

https://www.gov.uk/government/publications/dfids-strategy-for-payment-by-results-sharpening-incentives-to-perform
DFID. 2014b. Pilot Project of Results Based Aid in the Education Sector. Retrieved November 20, 2014.
Dollar, David, and Jakob Svensson. 2000. "What explains the success or failure of structural adjustment programmes?" The Economic Journal 110(466): 894-917.
Drucker, Peter F. 1954. *The Practice of Management*. New York: Harper Brothers.
The Economist. 2014. "Misleading Means: An alliance of billionaires and bureaucrats makes a difference." November 22, 2014. http://www.economist.com/news/americas/21633944-alliance-billionaires-and-bureaucrats-makes-difference-misleading-means
Eifert, Benn and Alan Gelb. 2005. "Improving the Dynamics of Aid Towards More Predictable Budget Support," World Bank Policy Research Working Paper 3732, Washington, DC: World Bank. October.
Epstein, Helen. 2010. "Cruel Ethiopia," *New York Review of Books*, April 24.
Ethiopian Ministry of Education. 2013. Education Statistics: Annual Abstract. EMIS, Planning and Resource Mobilization Directorate. Addis Ababa: Ethiopian Government. http://www.moe.gov.et/English/Resources/Documents/eab05.pdf
Fan, Victoria Y., Denizhan Duran, Rachel Silverman, and Amanda Glassman. 2013. "Performance-based financing at the Global Fund to Fight AIDS, Tuberculosis and Malaria: an analysis of grant ratings and funding, 2003–12." *The Lancet Global Health* 1(3):e161-e168.
Forstater, Maya, Smita Nakhooda and Charlene Watson. 2013. "Understanding the effectiveness of climate finance: the Amazon Fund." Working Paper 372, London: Overseas Development Institute (ODI).
Fox, Jonathan. 2007. "The uncertain relationship between transparency and accountability." Development in Practice 17(4-5): 663-671.
GAVI. 2009a. "Immunisation Services Support (ISS) update." http://www.gavi.org/library/news/statements/2009/immunisation-services-support-(iss)-update/
GAVI. 2009b. Update on Immunisation Services Support (ISS). http://www.gavi.org/library/news/statements/2009/update-on-immunisation-services-support-(iss)/
GAVI. 2011. Report to the GAVI Alliance Board, 16-17. November 2011. file:///C:/Users/cgd/Downloads/13%20-%20Performance%20Based%20Funding.pdf
GAVI. 2013. "GAVI Health Systems Strengthening (HSS) Cash Support, HSS and PBF Guidance Note" www.gavialliance.org/library
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. http://www.cgdev.org/sites/default/files/anatomy-program-for-results.pdf
Glassman, A., V. Fan, M. Over, R. Silverman, K. McQueston, and D. Duran. 2013. "More health for the money: A practical agenda for the Global Fund and its partners. A report of the Center for Global Development Working Group on Value for Money in Global Health. Consultation draft." Washington, DC: Center for Global Development.
Glassman, Amanda. 2013. "New PPP Tests Innovations in Health Aid." Views from the Center, Center for Global Development, March 18, 2013.
Greenberg, Jerald. 2004. *Managing Behavior in Organizations*. 4th ed. Upper Saddle River, NJ: Prentice Hall.
Gruening, Gernod. 2001. "Origin and theoretical basis of New Public Management," International Public Management Journal, 4:1–25.

Hallett, Timothy B. and Mead Over. 2010. "How to Pay 'Cash-on-Delivery' for HIV Infections Averted: Two Measurement Approaches and Ten Payout Functions." CGD Working Paper 210. Washington, DC: Center for Global Development.

Hermansen, Erlend A. T. and Sjur Kasa. 2014. "Climate Policy Constraints and NGO Entrepreneurship: The Story of Norway's Leadership in REDD+ Financing," Working Paper 389, Washington, DC: Center for Global Development.

IPCC (Intergovernmental Panel on Climate Change). 2007. "Climate Change 2007: synthesis report." http://www.ipcc.ch/pdf/assessment-report/ar4/syr/ar4_syr.pdf

Janus, Heiner. 2014. "Real Innovation or Second-Best Solution? First experiences from results-based aid for fiscal decentralisation in Ghana and Tanzania." Discussion Paper. Bonn: German Development Institute/Deutsches Institut für Entwicklungspolitik (DIE). http://www.die-gdi.de/uploads/media/DP_3.2014.pdf

Kahneman, Daniel. 2003. "Maps of bounded rationality: Psychology for behavioral economics." *American Economic Review*, pp. 1449-1475.

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

Killick, Tony. 1998. *Aid and the Political Economy of Policy Change*. London: Routledge/ODI.

Klingebiel, Stephan. 2012. "Results-Based Aid (RBA): New aid approaches, limitations and the application to promote good governance." Discussion Paper. Bonn: German Development Institute / Deutsches Institut für Entwicklungspolitik.

Koeberle, Stefan, Zoran Stavreski, and Jan Walliser, eds. 2006. *Budget Support as More Effective Aid?: Recent Experiences and Emerging Lessons*. Washington, DC: World Bank.

La Viña, Antonio G.M. and Alaya de Leon. 2014. "Two Global Challenges, One Solution: International Cooperation to Combat Climate Change and Tropical Deforestation," Working Paper 388, Washington, DC: Center for Global Development.

Levy, Jack S. 1997. "Prospect theory, rational choice, and international relations." *International Studies Quarterly* 41(1): 87-112.

Lim, Stephen S., David B. Stein, Alexandra Charrow, and Christopher JL Murray. 2008. "Tracking progress towards universal childhood immunisation and the impact of global initiatives: a systematic analysis of three-dose diphtheria, tetanus, and pertussis immunisation coverage." *The Lancet* 372(9655): 2031-2046.

Meier, Werner. 2003. "Results-Based Management: Towards a Common Understanding Among Development Cooperation Agencies," Discussion Paper Prepared for the Canadian International Development Agency, Performance Review Branch, for consideration by the DAC Working Party on Aid Effectiveness and Harmonisation. October 15.

Murray, Christopher JL, and David Evans, eds. 2006. *Health systems performance assessment: debates, methods and empiricism*. Geneva: World Health Organization.

Musgrove, Philip. 2010. "Financial and Other Rewards for Good Performance or Results: A Guided Tour of Concepts and Terms and a Short Glossary." Results-Based Financing for Health. Washington, DC: World Bank. http://www.rbfhealth.org/rbfhealth/system/files/RBF%20Glossary%20long.pdf

Nganwa, Peace. 2013. "Ethiopia. Resources for poverty eradication: A background paper." Development Initiatives: Africa Hub. February 2013. http://devinit.org/wp-content/uploads/2013/08/Ethiopia-Resources-for-poverty-eradication.pdf

Norad (Norwegian Aid Agency). 2014a. "Real-Time Evaluation of Norway's International Climate and Forest Initiative. Synthesising Report 2007-2013." August 2014. http://www.norad.no/en/tools-and-publications/publications/evaluations/publication?key=415169

Norad. 2014b. Norway's International Climate and Forest Initiative (NICFI): Brazil. https://www.regjeringen.no/en/topics/climate-and-environment/climate/klima--ogskogprosjektet/kos-innsikt/brazil-and-the-amazon-fund/id734166/. Retrieved January 13, 2014.

OECD. 2014. "Development Aid at a Glance: Statistics by Region, Africa 2014 edition." http://www.oecd.org/dac/stats/documentupload/2.%20Africa%20-%20Development%20Aid%20at%20a%20Glance%202014.pdf

One World Conservatism. n.d. "A Conservative Agenda for International Development." Policy Green Paper No. 11. https://www.conservatives.com/~/media/Files/Green%20Papers/OneWorldConservatism.ashx?dl=true

Pearson, Mark, Martin Johnson and Robin Ellison. 2010. Review of major Results Based Aid (RBA) and Results Based Financing (RBF) schemes: Final report. London: DFID Human Development Resource Centre. March. http://www.hlsp.org/LinkClick.aspx?fileticket=GWmdM_6ynhk%3d&tabid=2499

Pearson, Mark. 2011. "Results based aid and results based financing: What are they? Have they delivered results?" London: HLSP Institute. January. http://www.hlsp.org/LinkClick.aspx?fileticket=tdqKrWX321Q%3D&tabid=1648

Perakis, Rita and William Savedoff. 2014. "An Introduction to Cash on Delivery Aid for Funders." CGD Policy Brief. Washington, DC: Center for Global Development.

Ramalingam, Ben. 2013. *Aid on the Edge of Chaos: Rethinking International Cooperation in a Complex World.* Oxford University Press, United Kingdom.

Sandefur, Justin and Amanda Glassman. 2014. "The Political Economy of Bad Data: Evidence from African Survey & Administrative Statistics." CGD Working Paper 373. Washington, DC: Center for Global Development. http://www.cgdev.org/publication/political-economy-bad-data-evidence-african-survey-administrative-statistics-working

Sappington, David EM. 1991. "Incentives in principal-agent relationships." *The Journal of Economic Perspectives,* pp. 45-66.

Savedoff, William D. 2011. "Incentive Proliferation? Making sense of a new wave of development programs." CGD Essay, Washington, DC: Center for Global Development. http://www.cgdev.org/sites/default/files/1425405_file_Savedoff_Incentive_Proliferation_FINAL_0.pdf

Shiferaw, Solomon. 2014. "Results Based Aid Pilot." Presentation at the OECD-German Development Institute Technical workshop on results-based funding. May 19-20, 2014. Paris, France. http://www.die-gdi.de/fileadmin/user_upload/pdfs/veranstaltungen/2014/20140519_results_based_funding_16_Ethiopia_Education_-_Shiferaw.pdf

Stiglitz, Joseph E. 2002. "Participation and development: Perspectives from the comprehensive development paradigm." *Review of development economics*, 6(2): 163-182.

Vähämäki, Janet, Martin Schmidt, and Joakim Molander. "Review: results based management in development cooperation." Riksbankens Jubileumsfond, Stockholm (2011).

Vivid Economics. Forthcoming. "Results-based aid in the energy sector." Report prepared for the World Bank. London: Vivid Economics.
Washington, DC: Center for Global Development.

WHO (World Health Organization). 2009. *State of the World's Vaccines and Immunization*. Geneva: World Health Organization. http://www.researchamerica.org/uploads/factsheet27vaccines.pdf

World Bank. 2014. Public spending on education, total. Retrieved June 2014. http://data.worldbank.org/indicator/SE.XPD.TOTL.GB.ZS

Appendix A

Studies that identify key features of RBA include Pearson et al. 2010 and Klingebiel 2012. This list of features is based on our review of RBA designs and our literature review on the theories behind them. The accompanying table is a heuristic device to show how complicated (and perhaps amusing) it would be to try to classify a program in a single sentence.

- *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. On this latter point, if the program extends over a long period then errors in payment one year are likely to be offset by compensating errors in subsequent years on average. Short-term programs must be more precise because there is less chance of averaging out errors over time.
- *Renewability*: Renewable and non-renewable agreements will differ in terms of the incentives they provide for collaborating, sharing information, and enforcement. One-time agreements provide less incentive to collaborate than for each party to simply maximize the gains they can extract from the contract. When agreements are renewable and attractive to both parties, then interests in continuing the arrangement will affect each party's willingness to offer benefits to the other, even at the expense of some short term personal benefits.
- *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 manage. The degree of transparency also introduces additional sources of feedback and accountability which may affect the design and performance of programs.
- *Recipient*: Under performance-based funding agreements, recipients range from individuals to governments. This paper focuses on agreements with governments, as RBA is typically defined. Governments could be finance ministries, sector ministries, or any level of sub-national government. These agreements should be distinguished from agreements with other kinds of agents (see Figure 1) whose resources and decision processes follow different theories of change and behavioral models.

- *Results level*: This paper focuses on agreements that aim to pay in proportion to outcomes, but agreements can pay for changes at any point in the "results chain" – whether that means paying upon purchase of specific inputs, the completion of certain tasks or activities, enacting regulations or laws, establishing procedures, or producing goods or services. Depending on the indicator chosen, agreements can be structured to pay for impacts, that is, changes in outcomes attributable to the recipient's actions – but this requires establishing a counterfactual reference level of what would have happened otherwise.
- *Result complexity*: Agreements differ over the complexity of the results for which they pay. For example, agreements might include payments for one, a few or many results with implications for such things as managerial attention, the strength of incentives, and ease of measurement.
- *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. When payment amounts are small, relative to unit costs or to other funding sources, they still signal changes in outcomes that may be useful for management or accountability but their impact via pecuniary interests will be attenuated.
- *Up-front payments*: The presence of up-front payments may support or undermine performance agreements. Up-front payments related to preparation for measuring results are completely consistent with ex-post performance payments. Up-front payments for activities or inputs considered necessary for achieving progress may be successful if program designers are fortunate to select the right plans but may undermine the results-based agreements; these payments would be open to all the criticisms associated with conventional aid approaches that prompt interest in performance payments.

- *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. An important element of credibility is the verification of results through the use of independent information.

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

Projects considered for inclusion in survey of RBA Programs:

The research for this working paper began with a search for the full universe of results-based funding in foreign aid programs. We used a broad definition of results-based funding to include any program that links funding to measurable results *ex-post,* and sought to categorize programs according to criteria such as *who is receiving the incentive* and *what kinds of indicators are being used to measure progress* (see full list of criteria below). It quickly became apparent that the universe of results-based funding was very big, but the number of programs with specific characteristics that interested us, that is, paying *governments* for progress on improving *outcomes* as proposed in CGD's work on Cash on Delivery Aid, was very small. This finding contradicted what we heard from development agency staff and experts who believed a large number of these RBA programs already existed. When we discovered so few examples, we chose to use a case study approach to delve deeply into the four examples we identified as (1) paying a government on the basis of results achieved where (2) results were development outcomes (3) the agreement had been in place for at least a year and (4) information about the program was publicly available.

Many of the programs we initially investigated fell under the category of *results-based financing* (payments going directly to service providers, households or individuals). We excluded these programs because we expect the incentive structures and underlying dynamics of paying NGOs, firms, or individuals to be quite different from those associated with making payments to a government.

Among programs that paid governments, some were excluded because they measured performance in terms of compliance with predefined plans rather than progress in achieving outcomes. For example, the Global Fund to Fight Aids TB and Malaria provides grants to governments and links payments to performance but not in a way that clearly relates to health system outputs or health outcomes. Instead, Global Fund programs measure performance with a long list of indicators that seem designed to assess compliance with pre-determined plans –much more like conventional aid. The European Commission's budget support programs have included variable tranches which are ostensibly paid out against performance on specific output and outcome indicators. However, in this case, the lack of transparency and limited public information on basic design features and performance made it impossible to include in our analysis.

The following list of projects in the results-based funding "universe" represents the sample that we identified in August 2013, from which we selected our four RBA case studies and (with the exception of the GAVI Immunization Services Support program which had begun to be phased out in 2010) tracked implementation progress through the end of 2014:

Projects excluded because payments are not to governments

- Conditional cash transfers (6 programs)

  We looked at conditional cash transfer programs, whereby governments pass on incentives to households, in 6 different countries. These programs are described in the book "Performance Incentives for Global Health: Potential and Pitfalls" by Rena Eichler, Ruth Levine, and the CGD Performance-Based Incentives Working Group, 2009.

- Pay for Performance health programs under USAID-funded Health Systems 2020 scheme (16 programs)

  We found information for 7 out of 16 listed programs, funded upfront by multiple donors, in which incentives were passed down from governments to patients or health care providers, typically according to output or process indicators.

- Health Results Innovation Trust Fund (HRITF)

  The HRITF (rbfhealth.org), funded by Norad, DFID, and complemented by funding from the World Bank's International Development Association, has committed or disbursed grant funding for 36 country pilot programs. Funding is used to support the design and implementation of RBF in low income countries and incentive payments are directed to health facilities.

- Global Partnership for Output-Based Aid (GPOBA)

  GPOBA is a multi-donor trust fund administered by the World Bank that funds and documents learning from output-based aid projects that aim to improve the delivery of basic services. Incentives are typically directed to private providers.

Projects that pay governments but were excluded because payments are not for outcomes (or for proxies of outcomes)

- Global Fund (900+ projects)

  We found information for over 900 projects on Global Fund's website, which included a mix of government ministries and non-governmental organizations as recipients. Initial funds were approved by an independent technical review panel and, in order to continue funding after two years, recipients had to meet a certain performance grade on a series of indicators. They were graded against mainly process indicators.

- European Commission MDG Contracts and budget support programs

  The EC's budget support programs include a fixed tranche, and a variable tranche, which links funds to performance. We were particularly interested in the MDG contracts which linked at least a portion of the variable tranche to the achievement of clearly defined indicators of progress that correspond to the Millennium Development Goals. For both MDG Contracts and other EC budget support programs, funds are often linked to the adoption of certain policies or reforms which may or may not lead to improved results. In either case, we were not able to find - and did not receive responses to our inquiries about – country-level information on how results indicators were defined and how programs were progressing, and thus were not able to include the EC's programs.

- World Bank Program for Results (PforR) (5 programs)

  We considered the first 5 projects that had been approved by the World Bank under the new PforR financing mechanism in 2013. These programs generally linked funding to a list of process and output indicators, not to outcomes. Moreover, they were still too new for us to have enough information about how they were progressing. As of November 2014, there were 22 approved projects on the Bank's PforR website. Gelb and Hashmi 2014 reviews the design of the initial operations.

Data collected on results-based programs (criteria for exclusion/inclusion indicated by italics)

Category (RBA or RBF)

Name of scheme

Program Dates

Sector

Sub-Sector

Country

Funder

Recipient of Incentive

Progress Indicator(s)

Type of Progress Indicator (outcome, output, or input)

Type of Progress Indicator (incremental or binary)

Payment rate

Total Commitment

Total Disbursement

Transparency: Is contract available online?

Level of Funder Involvement

Results to date

How is program being verified?

How is program being evaluated?