Can Results-Based Payments Reduce Corruption?

Charles Kenny and William Savedoff

Abstract

A common objection to results-based programs is that they are somehow more vulnerable to corruption. This paper explains why results-based approaches to foreign aid may be less vulnerable to corruption than the traditional approaches which monitor and track the purchase and delivery of inputs and activities.

The paper begins by classifying different corruption costs and specifically distinguishes the problem of diverted funds from the costs associated with failing to generate benefits. It then characterizes the key differences between traditional input-tracking programs and results-based approaches in terms of how they are supposed to work, the implicit risks that preoccupy designers, how they function in practice, and what this means both for the scale of corruption and the realization of benefits. It then considers the conditions under which one approach or another might be more appropriate. The paper concludes that input-tracking approaches are vulnerable to corruption because they have high failure costs and a weak track record for controlling diverted funds.

By contrast, results-based approaches are less prone to failure costs and limit the capacity of dishonest agents to divert funds unless those agents first improve efficiency and outputs.

JEL Codes: D73, D86, F35, O20, O31

Keywords: Corruption, aid effectiveness, results-based aid, results based financing, forensic economics, cash on delivery aid

Center for Global Development
www.cgdev.org

Working Paper 345
September 2013

Can Results-Based Payments Reduce Corruption?

Charles Kenny
Center for Global Development

William Savedoff
Center for Global Development

CGD is grateful for contributions from the Bill & Melinda Gates Foundation, the Norwegian Ministry of Foreign Affairs, and the Swedish Ministry of Foreign Affairs in support of this work.

Charles Kenny and William Savedoff . 2013. "Can Results-Based Payments Reduce Corruption?." CGD Working Paper 345. Washington, DC: Center for Global Development.
http://www.cgdev.org/publication/can-results-based-payments-reduce-corruption-working-paper-345

Center for Global Development
1800 Massachusetts Ave., NW
Washington, DC 20036

202.416.4000
(f) 202.416.4050

www.cgdev.org

Contents

1. Introduction

The idea of improving foreign assistance by paying for "results," "outputs" and even "outcomes" has been around for decades, yet its use has been restricted. A relatively small number of programs actually pay for results; a larger number claim to be paying for results but are only doing so superficially. One of the reasons for its slow adoption has been criticism that programs which disburse against outputs or outcomes lack the procurement procedures and audit mechanisms necessary to avoid corruption. This paper argues, to the contrary, that results-based payment systems are likely to be *less* vulnerable to corruption than traditional input-tracking approaches because results-based approaches make the relevant effects of corruption – failure of programs to deliver results – visible.

What does it mean for a program to be "vulnerable" to corruption? Typically people ask whether funds are diverted from programs to inappropriate uses. In this sense, the standard for assessing the anti-corruption bona fides of an aid program is the degree to which we can assure that funds are used to purchase inputs and activities that are designed into the program. However, this perspective assumes that the inputs and activities will necessarily translate into the outputs and outcomes which are the true goal of the program. Such tracking of inputs ignores the most important question funders should ask of a program: does it yield results? In this sense, the appropriate way to measure the costs of corruption in a program is not to track how the funds are spent but whether abuses stopped the program from achieving its goals. We argue, therefore, that a program should be considered vulnerable to corruption if abuses are likely to reduce a program's development impact.

Discovering that funds disappeared from a program generates many scandals, but discovering that benefits have not materialized at the project level is comparatively rare and unremarked, however common it may be in practice. At the same time, the cumulative effect of neglecting the systematic measurement of project results is to have less effective aid programs. By shifting attention from controlling inputs to rewarding outputs and outcomes, results-based programs make benefits visible. In so doing, results-based programs can mitigate the impact of corruption on development.

This paper begins by discussing how corruption is defined and how it affects the impact of aid programs. It then proceeds to contrast traditional input-tracking programs and results-based programs in terms of how they are supposed to work, the implicit risks that preoccupy designers, how they function in practice, and what this means both for the scale of corruption and the realization of benefits. It concludes by considering the conditions under which one approach or another might be more appropriate.

2. What are the real costs of corruption in foreign aid?

Corruption is often defined as "the abuse of public office for personal gain" (Bardhan 1997) or "the abuse of entrusted power for private gain" (Transparency International 2001). For the purposes of this paper, the exact definition is less important than the point that foreign aid is public funding which enters a country for an express purpose and which can be diverted from that purpose through, for example, fraud, misrepresentation, embezzlement, theft, bribes, or collusion. While the motivation of such acts is often personal gain, it can also be directed toward political or social aims without involving individual enrichment. We will therefore use the term corruption with regard to aid broadly to refer to the act and impact of foreign aid funds diverted from their intended purpose.

Diverted funds are perhaps the most obvious way that people think about the costs of corruption in foreign aid. Diverted funds are simply the amount of money that is taken from a program through, for example, theft or bribes. So, for example, if half the money were stolen from a $10 million program designed to provide piped water to a community, the cost of this theft might be considered to be $5 million. This is a typical way in which international and bilateral agencies report corruption in their investigative reports.

Failure costs are not counted when corruption is measured in terms of diverted funds. In fact, the amount of funds that are diverted will generally underestimate the impact of corruption because it fails to count the significant cost of failing to achieve the intended project outcome. Even a small diversion of funds from a program can completely ruin its impact if it undermines a critical component; for example, when failing to reinforce concrete causes a building or bridge to collapse. In the example above of a water program, $5 million may have been stolen, but the more significant impact of that theft is to reduce the number of households who gain access to potable water. Typically, households in peri-urban areas without water connections can spend 10 times or more per liter than those who are connected to the public network when they purchase water in bottles or from trucks. Thus, the cost of diverting funds from this program is better measured as the *foregone benefits* that would have otherwise occurred, a failure cost that might exceed the stolen funds by orders of magnitude.

This is a major problem for development programs. Collusion, bribery or incompetence that raises contract prices is as a rule significantly less damaging to development outcomes than malfeasance which leads to building the wrong thing, building it badly, or utilizing and maintaining it poorly (Kenny 2006). For example, Olken (2004) estimated that the theft of an additional $1.00 of materials reduced the discounted benefit of a road program by $3.41 because substandard construction reduces road quality and durability.

It is possible to imagine cases where corruption occurs but doesn't harm results and, in such cases, a results-based program would disburse funds. This diversion would go undiscovered which is very troubling for obvious moral reasons. In fact, the ubiquity of input-tracking and fiduciary controls in public administration demonstrates that polities regularly adopt

approaches that provide assurances of legal probity (regardless of results) over approaches that assure delivery of results (regardless of legal probity). However, in the case of foreign assistance, there are several reasons why this revealed preference should hold less sway. First, foreign assistance is a transaction between one government and another. While recipient governments should be accountable to their citizens by demonstrating that funds are used as intended it is less clear why additional (and often ineffective) reporting requirements by funders should be laid on top of such domestic procedures, especially given that funders regularly state that their primary goal is to achieve results. Second, input-tracking requirements from those who finance aid programs have to be *effective* if they are going to fulfill the demand for legal probity. In practice, the evidence for such effectiveness is lacking. Finally, it is important to be explicit about the equally troubling morality of investigating and punishing corruption regardless of its impact on the intended purposes of the program (i.e. to generate benefits). Thus, a program in which individuals divert funds by producing roads, health services or water connections more efficiently than anticipated can be considered morally preferable within certain frameworks (e.g., consequentialist views) than a program in which benefits fail to materialize even if fewer funds were diverted. Populations that reward openly corrupt politicians with reelection after delivering results are, fortunately or not, demonstrating that this view is not uncommon.[1]

In addition to diverted funds and failure costs, a full accounting of the costs of corruption also requires consideration of: (1) the direct costs added to programs in order to prevent and detect corruption; (2) the efficiency costs caused by interference of fiduciary controls in cost-effective design or implementation; and (3) the dynamic costs caused by distorting institutions to facilitate corruption or undermining social trust.

Direct costs of preventing and detecting corruption are fairly easy to conceptualize. They include the staff time and associated expenditures required to document financial flows, receipt of funds, and delivery of goods and services. They also include the expenditures associated with internal and external audits, maintaining channels for grievances and complaints, and for investigation of allegations of malfeasance. These costs are covered by administrative budgets of funding agencies or built into programs as administrative overhead.

Efficiency costs occur when a mechanism designed to prevent or detect corruption introduces a change in program design that reduces its cost-effectiveness or a rigid procedure that makes it difficult to adapt a program to new information or changing conditions. Efficiency costs are essentially invisible. The only way to measure or reveal efficiency costs is to compare the design and implementation of similar programs responding to different fiduciary controls.

Dynamic costs occur when the existence of aid distorts domestic social norms or institutional development in ways that facilitate corruption and undermine accountability. For example, if

[1] Consider, for example, Ademar de Barros who ran for Governor of São Paulo, Brazil during the 1950s with a campaign slogan of "Roubo mas faz" ("I steal but I get things done").

fraud committed against aid programs goes unpunished, it may undermine social norms that discourage such abuse in other spheres of activity. The prospects of enrichment from aid programs can also distort local institutions when officials intentionally design programs to facilitate corruption by reducing transparency and weakening control mechanisms. The possibility of diverting funds from aid programs can also have dynamic effects when local individuals put their efforts into constructing elaborate shadow corporations or sham workshops (Berkman 2008).

The rest of this paper will focus on the *diverted costs* and the *failure costs* of corruption because diverted costs are the primary objection to results-based modalities and because failure costs are the least visible cost of corruption in input-based programs. In particular, we will demonstrate that traditional input-tracking programs are more likely to disburse funds without achieving their aims when corruption is present, as compared to results-based programs. A well-designed results-based program – by linking payments to verified outputs or outcomes – will only disburse funds when benefits are realized. By contrast, a well-designed traditional input-tracking program can easily disburse funds despite a failure to achieve its goals. The paper will also suggest that the direct costs and efficiency costs of results-based programs are likely to be much smaller than for input-tracking programs.

3. The traditional input-tracking model

Traditional input-tracking aid programs are designed to produce results but are actually managed in terms of completing activities and tasks. A traditional aid program, then, can fail to produce results for a variety of reasons related to poor design or poor implementation. While most funders have systems in place to monitor results, most monitoring effort is aimed at verifying activities like procurement and completion of tasks rather than verifying results in terms of outputs or outcomes. This is why we will refer to the traditional aid model as an "input-tracking" approach. This section explains the typical traditional aid program that utilizes an input-tracking model and illustrates the limitations in terms of controlling corruption even if it is implemented successfully.

In essence, the traditional model of assistance pays for inputs even though this is not necessarily its aim. The process begins with the funder[2] and recipient government agreeing on program goals such as educating children or improving transport infrastructure. Then, the funder and recipient government develop a design for a program that can realize those outcomes, as it might be a school construction program or a road rehabilitation program. This program design includes details for the technical approach to be used in constructing physical infrastructure or the activities (such as training) and equipment required for service delivery. The program design is also used to create tender documents for the provision of goods, works and/or services required to complete the program. The recipient or funder

[2] The term "funder" is used here to refer primarily to bilateral and multilateral agencies that provide financial resources to support programs or programs in low- and middle-income countries. Sometimes these funds are provided as grants and sometimes as loans.

issues these documents and a contractor is typically selected on the basis of a competitive bid process. The funder directly pays the contractor on the basis of contractually agreed milestones or monitors the recipient government when it is responsible for hiring and supervising the contractor or procuring and verifying delivery of particular goods. Once inputs are delivered, the program is complete. In the traditional model, monitoring focuses on processes and receipts. While in theory funders monitor outputs (roads built, technical assistance given), this is often a comparatively cursory focus of attention compared to process management – not least because, as financiers, funders are primarily concerned with payments linked to those processes rather than outcomes.

The primary claim regarding the probity of traditional input-tracking programs is that by tracking the money used to purchase inputs, the funder can judge whether or not funds have been applied as intended. If the process works as designed (e.g., monitoring and supervision are effective), inputs are presumably delivered at a competitive (low) price. If funds are monitored (or directly delivered) from the funder (or recipient government) to the contractor in a way that prevents leakage, then all funds are presumably used to purchase the inputs required by the program design. Whether the inputs are applied properly or have the qualities required to produce the desired benefits is a matter which is addressed only imperfectly in most cases.

The direct costs to funders associated with this approach are those of design, procurement (oversight), financial oversight and delivery oversight. Additional costs are incurred because the system is not in fact watertight –these include internal and external audits, mechanisms to receive complaints and investigations.

Traditional input-tracking programs face a number of risks that can keep them from achieving their goals. Design failures occur when the selected approach is not, in fact, an efficient mechanism by which given resources can achieve outcomes. Tender failures occur when associated documents mis-specify or under-specify the technical requirements for implementing the design. Procurement failures occur when the procurement process fails to select a qualified bidder at a competitive price. Oversight failures occur when those responsible for supervision do not adequately monitor the application of funds or the quality of deliverables; and operational failures occur when the delivered inputs are improperly applied or maintained.

These potential failures – in design, tendering, procurement, oversight and operation – can happen as a result of human error and be unintentional. However, these failures can also happen as a consequence of actions by individuals intending to defraud the program. Corruption can undermine a traditional input-tracking program in any of these stages. For example, a program's design can be skewed towards activities and procurement that a particular individual or group can more easily manipulate for their own ends. Collusion can undermine competitive procurement processes. Supervisors and auditors can be bribed to overlook poor quality deliverables or embezzlement.

This enumeration of risks in traditional input-tracking programs is not fanciful. Ample evidence is available from aid agencies to demonstrate that such failures occur. The following discussion of failures in traditional input-tracking programs draws from the experience of the World Bank – not because it is any worse than other aid agencies but simply because its problems are better documented. It shows that traditional input-tracking programs often fail to achieve their objectives; that associated procurement rules are not robustly controlling corruption; and that controlling such programs is costly.

First, we know that many traditional input-tracking programs do not achieve their objectives. Pritchett (2000) reports data from the World Bank's evaluation department from 1973-1991 showing an average rate of return for all countries of about 14 percent but also notes an example of one African country where 31 World Bank-financed programs produced a median rate of return of zero (Pritchett, 2000). Of the 1,151 World Bank investment projects approved in FY2000 or later and rated by the Independent Evaluation Group by February 2013, 296 were judged unsatisfactory or moderately unsatisfactory. A further 430 were rated only moderately satisfactory –suggesting 63% of all World Bank investment projects are graded less than fully satisfactory by IEG.

Second, tendering, procurement and audit procedures to control corruption are only weakly effective at best. Not least, one of the chief mechanisms by which tendering rules are supposed to limit corruption and assure value for money is by ensuring competitive bidding, but the extent of competition is often limited. World Bank rules for International Competitive Bidding (ICB) and National Competitive Bidding (NCB) require public advertising and selection on the basis of price after bids have demonstrated a technically satisfactory response to the specifications laid out in the bid documents. Yet the number of bidders is low and declining (see Figure 1) and the difference between international and national bidding competitions is relatively small. More than a third (37 %) of all internationally-bid and nationally-bid infrastructure procurements are held with 3 or fewer bidders, and 12% are held with a single bidder. The average number of bidders on internationally-bid infrastructure contracts fell from 6.4 to 4.3 between the two-year periods 1995-6 and 2006-7 (Kenny and Musatova, 2009).

Another way that the tendering and procurement process is supposed to control corruption is through attracting and selecting qualified firms. Yet for a competitive process to work in that regard, it is essential to have deliverables fully specified and oversight to ensure full compliance with those specifications. Otherwise contractors who intend to provide poor quality goods are the ones more likely to win by offering prices that are unrealistically low. Without adequate monitoring, such firms can deliver poor quality or, if necessary, avoid being penalized through the judicious use of bribes. In corrupt regimes in particular, 'successful' competition may merely shift where corruption takes place –from bribing to win the contract at a high price to bribing to cover up substandard works.

Some evidence of this tendency of 'bad contractors to drive out good ones' (to borrow from Gresham's Law) can be found in World Bank programs. The capacity to judge the quality of

contractors seems to be fairly limited. A survey of firms that bid on international contracts found that only fifteen percent of respondents thought that tender rules were an obstacle to corruption (Søreide, 2006). Perhaps as a result of this perception alongside strengthened OECD laws on bribery overseas, in 2009, there were fewer than 20 bids from the OECD on all World Bank financed internationally bid contracts, down from an average of around 100 in the last years of the 1990s (Figure 2).[3]

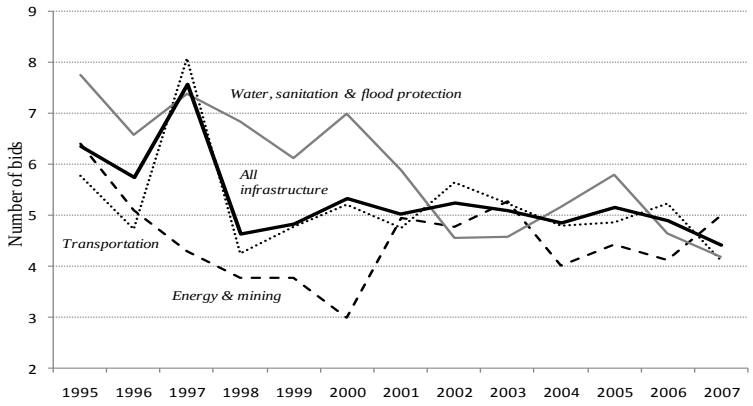
As a result of these weaknesses, procurement rules do not effectively insulate World Bank operations from corruption. For example, the World Bank applies similar rules for monitoring and supervising road programs across countries, but the average cost is substantially higher in countries where bribes are more common. The average cost for rehabilitating a two lane highway across eighteen countries for which we have good data on both bribes and costs was $\$36/m^2$. In countries where the average bribe for a government contract was reported to be below two percent of the contract value, this cost was $\$30/m^2$. For countries where bribes for government contracts were reported to be larger than two percent of their value, average costs were $\$46/m^2$.[4]

The third and final limitation that we will discuss is the high transaction costs involved in these control mechanisms. The procurement model at the heart of traditional input-tracking programs is burdensome for clients and funders alike. It requires multiple contracts with the government, consulting firms and contractors, none of which are explicitly designed to monitor outcomes. Indeed, the consequences for failing to disburse funds are often more serious than the consequences for failing to achieve outcomes.

[3] Other explanations are possible, including that legitimate and honest competitors from developing country firms have been able to win contracts because they have lower supply costs.

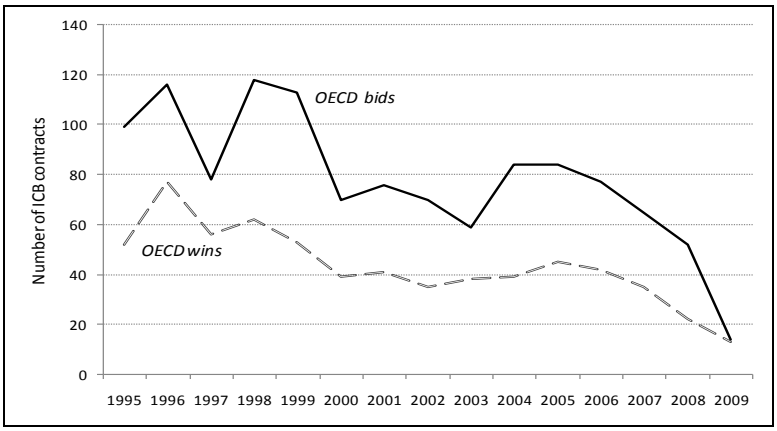
[4] Adding to concerns that World Bank oversight mechanisms appear to be only a partial defense against governance weaknesses is the fact that a number of programs declared satisfactory by staff and evaluators appear in hindsight to have been less than completely successful over the long term (Kenny, 2009).

Figure 1: World Bank Civil Works Infrastructure Procurement Bids by Sector



Source: Kenny and Musatova, 2010

Figure 2: Proportion of All ICBs Won or Bid On by OECD firms Over Time



Source: Kenny and Musatova, 2010

One indication of the high transaction costs associated with traditional input-tracking programs is time. A study of consulting services in World Bank financed programs found that the average selection process took 17 months, representing almost two-thirds of the average contract duration. Furthermore, negotiation times varied from 401 days in East Asia to 561 days in South Asia with no strong pattern relating to the difficulty of the context or the program (Casartelli and Wolfstetter, 2007).

The complexity and cost of this procurement approach is also demonstrated by the frequency with which procedural problems get flagged independent of apparent wrongdoing. Kenny and Musatova (2010) examined a sample of World Bank water and sanitation programs and found that almost every contract raised at least one "red flag" – such as failure to advertise properly, low number of submitted bids, or two almost identical bids being submitted – that might indicate an effort to manipulate the process. Yet, contracts that had been separately investigated and judged potentially tainted were no more likely to have "red

flags" than other contracts. The ubiquity and apparent randomness of these red flags suggests that they are linked to the overall complexity of procurement processes rather than systematic evidence of corruption risks.

High contracting and process monitoring costs divert resources that could otherwise be used to monitor outcomes, adjust strategies, test new approaches, and reward good performance. The focus on controlling the procurement of inputs takes the attention of government staff away from improving their own government's institutions in favor of satisfying the requirements imposed by funders in a parallel governance system. A (very) partial accounting is provided by the World Bank where the budgets of the internal audit and institutional integrity departments alongside the evaluation and suspension and sanctions boards is $30 million –only about $5 million less than the budget of the Independent Evaluation Group, tasked with measuring the development impact of World Bank programs. This does not account for the 417 full time procurement and financial staff and 200 procurement-accredited staff spread throughout the Bank who work to ensure projects follow approved financial management and procurement approaches –at a cost that dwarfs expenditure on evaluation.[5] Nor, of course, does it account for the staff in recipient country governments tasked to manage mandated procurement and financial processes and report on World-Bank projects.

In defense of World Bank task teams (and the institution as a whole) they *do* care about results. This is a major reason for the coping strategies put in place in an attempt to assure that the input-dominated official process produces results. These strategies include bundling investments with consulting contracts to provide design and project oversight alongside the use of technical assistance, programmatic aid, and sector wide approaches, all predicated on the idea that overall policy and institutional status are likely to have a considerable impact on the development return to aid financing.

But the technical assistance and contracting support involves additional procurements and yet more 'input oversight' in an area where it is even weaker at ensuring outcomes – consulting services. A recent analysis of procurements for consulting services found no correlation between cost and quality scores and limited presence of sector specialists in evaluation committees (only 38 percent of evaluators across all contracts were sector experts) (Casartelli and Wolfstetter, 2007). Another indication that qualification may be inadequately considered is that the lowest bidders won these contracts 67 percent of the time. Only 15 of the world's top consulting engineering firms and 4 percent of the top fifty management consultant firms bid on these contracts and when they participate "they win a minimal proportion of contracts" according to the analysis (Casartelli and Wolfstetter, 2007).

---

[5] http://siteresources.worldbank.org/INTPROCUREMENT/Resources/AR-FY11-final-2-29-12.pdf. The World Bank has about 9,000 staff and a budget of $3.8 billion. That suggests per-staff costs of a little more than $430,000. At a pro-rata rate, then, procurement and financial management costs the institution about (.43x417)=$180 million per year.

If traditional input-tracking programs were successfully implemented, then procurement procedures might eliminate or at least control corruption. But this still does not make them successful programs. Within this paradigm, the procedures required to control corruption also require predefining the inputs that will be purchased and expending resources on monitoring and supervision. Therefore, when the system works perfectly, a traditional input-tracking program still runs the risk that the chosen inputs will not achieve the desired outcomes and may generate transaction costs that exceed the benefits.

In reality, traditional input-tracking programs are not perfectly implemented. Errors and abuses in design, tendering, procurement, oversight and operations mean that collusion and other difficult-to-detect abuses may still occur. But more importantly, the traditional input-tracking model has no built-in mechanism to detect and respond when outcomes fail to materialize. Thus the costs of corruption in terms of foregone benefits are never revealed and the system is structured to spend money almost without regard to its impact.

4. The results-based model

Results-based programs disburse funds in relation to outputs or outcomes rather than inputs and activities. In a pure results-based model, the funder and recipient government agree on a goal and establish a fixed fee per unit of progress toward that goal. The agreement establishes the indicator used to measure progress as well as the process for measuring and verifying the amount of progress achieved. In principal, the funder need have no further involvement in terms of technical assistance or oversight, allowing the recipient government to have full ownership of its policies and full flexibility to pursue whatever strategy it deems most likely to succeed. Disbursements are then made *ex-post* upon confirmation that results were achieved.[6]

Traditional input-tracking programs also generally start by establishing a goal; however, they then proceed to design, in detail, how the recipient government is to go about achieving it. By contrast, the first step after setting goals for a results-based program is to specify the desired outcomes in measurable terms. This assures that attention to the program's *impact* is established from the outset.

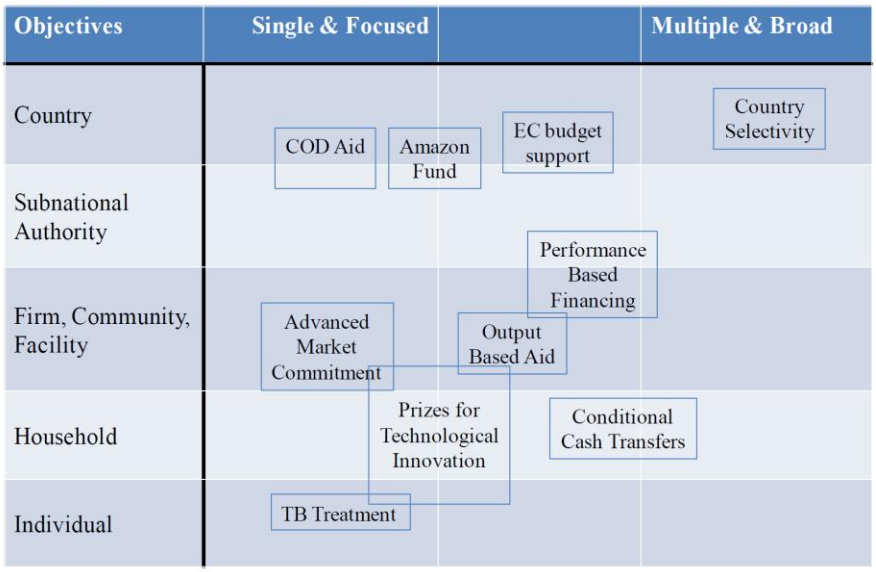
For a number of reasons, most results-based programs diverge from the ideal model described above. First, the assumption that lack of funds is a key obstacle to making any progress leads most results-based programs to provide at least some funds "up front" and independent of results achieved. Second, the concern that funds may be stolen or diverted has led many results-based programs to preserve the architecture of traditional input-tracking programs – which is considered to be the conservative approach to fiduciary risk – and overlay results as either an additional condition or a non-binding condition for disbursement. Third, many results-based aid programs assume that recipients do not have

[6] A detailed discussion of this model can be found in Birdsall and Savedoff (2010). For a survey of results-based approaches and how they compare, see Savedoff (2011) and Pearson et al (2010).

the capacity to design their own programs and consequently impose detailed designs – in terms of planning procurements, activities, and tasks – for achieving results. This may involve multiple additional contracts for consulting services and ancillary projects and end up distracting managers from the measurement of results. Fourth, results-based programs can end up with so many indicators and measure them so poorly that disbursements become contingent on negotiating skills rather than the outcome measurements.

The range of programs that seek to pay for results is large. They vary at least along two important dimensions – the nature of the recipient and the number of indicators (see Figure 3). Programs focusing on individuals sometimes pay for results, as in the case of rewarding people infected with tuberculosis who adhere to a prescribed regimen and complete their treatment. Conditional cash transfer programs are generally aimed at households, for example, paying women whose children remain in school. Sometimes service providers are the focus of results programs. For example, performance based financing (PBF) programs reward health facilities for meeting quantity and quality targets while output based aid (OBA) programs withhold some payments until final delivery and operation of water or electricity utilities is verified. Subnational and national authorities are the recipients in the case of budget support and COD Aid programs. The design of a results-based program should reflect the differences across recipients, both in terms of the risks they can reasonably bear, the scope of action they have to respond to the incentive, and the likely responses to different ways of structuring the payment.

Figure 3: Incentive Programs by Agent and Objectives



Source: Savedoff 2011, p. 12.

The other relevant dimension is the number of objectives. Some programs have very few indicators. In the case of a tuberculosis treatment program, adherence (and cure) is the single

indicator for reward. In many conditional cash transfer programs, a number of results are measured – such as school attendance, growth monitoring, and vaccinations. Many results-based programs for facilities incorporate a range of indicators – one program in Haiti focused on 7 service indicators (Eichler et al 2007) while another in Rwanda included 14 (Basinga et al 2011). At the national level, it is more common to find a large number of indicators included in results-based programs. Budget Support programs may contain 30 or more indicators. By contrast, COD Aid programs typically contain only a few; for example, DFID's secondary school program in Ethiopia pays for each child who completes the secondary school exam, only differentiating the unit payments with respect to location (higher for marginalized regions) and gender (higher for girls). Amongst other outputs, OBA-based payment programs have been based around sustainable connections to utilities, verified provision of basic health services, and the extension of irrigation.

When a result-based program pays for a single or a few indicators, political and managerial attention is directed more clearly to those goals. The agreement is also more credible. When programs include a large number of indicators, it is easier to convince funders to disburse funds based on successful measures even when progress is poor in other areas. Only if disbursement schedules are clearly linked to each individual indicator can this problem be avoided. The inclusion of many indicators tends to increase the discretion of both funder and recipient in a way that payment for a single result – or explicit link to a single indicator – would not permit.

At the same time, results-based programs must either have good direct measures of the desired outcomes or indicators for outputs which are closely linked to those outcomes. For example, a program could pay directly for reducing the prevalence of measles (the outcome) but the number of children vaccinated for measles is an output that is so closely linked to the spread of measles that it is a good proxy for a results-based program. Paying directly for the number of students who demonstrate they have mastered basic skills in primary school would also be a good outcome measure, but the number of teachers trained or schools built would not be particularly good as measures for development impact. Reductions in transportation costs might be a good outcome indicator but it is difficult to measure. On the other hand, miles of road paved would be a particularly poor indicator for improvements in transportation since the road could be poorly constructed and quickly fall apart or lead to nowhere and be unused. At the least, a results-based approach to road construction would want to include an indication of the quality of roads built and an economic rationale for route choice to have some assurance that the output being purchased will, indeed, lead to improved transportation outcomes.

Even when outcomes can be measured well, results-based programs need to pay attention to minimizing unintended consequences. As it might be, an output-based project for the construction of a 147 meter tall stone pyramid under a Pharaonic dynasty would certainly be a bad idea unless it monitored working conditions. But in more typical programs, the unintended consequences that can result from the recipient's discretion over how to achieve results depends on a range of contextual factors such as social accountability mechanisms,

the characteristics of the outcome, and the visibility of unethical or destructive activities. In implementing its Payment for Results modality, for example, a recent World Bank program in Nepal excluded bridges in environmentally sensitive areas precisely for this reason, but applied a results-based modality to bridges in other areas. Thus, unintended consequences are not an objection to paying for outcomes, only an issue to be assessed and mitigated.

The experience of results-based programs with regard to corruption, diversion of funds or the impact on outcomes has not been fully studied. Some programs may be overpaying for results because they have relied on self-reported performance, as appears to have happened with the Global Alliance for Vaccines' (GAVI) Immunization Support Services program (Lim et al 2008). While criminal intent has been demonstrated in some cases of fraudulent reporting (e.g. US Medicare payments to hospitals), we are unaware of proven fraud in the case of results-based payments in foreign assistance programs. In a World Bank–financed OBA project in Southeast Asia, allegations of corruption resulted in the project being put on hold and eventually canceled. The project was halted before outputs were fully delivered and verified, so the World Bank had not disbursed all related funding (Mumssen and Kenny, 2007). In GAVI's case, 39 countries over reported but another 8 countries underreported, suggesting that some of the variation might have been due to real errors in reporting rather than systematic fraud.

In the case of well-designed programs with a single simple measured outcome or output, it is difficult to imagine how funds could be diverted. For example, the Amazon Fund pays Brazil on the basis of reductions in deforestation. The reductions are measured by satellite imaging and certified by an independent international committee of experts. No payment occurs without real performance. This is one case but it demonstrates quite clearly the power of a pure results-based approach. The only possible objections are that (1) Brazil would have made those policy changes anyway (i.e. no additionality) or (2) the Amazon Fund is paying more than the true cost of the results. Both of these objections rely on a counterfactual that is difficult to establish. Furthermore, these objections apply equally to traditional input-tracking approaches.

While traditional input-tracking programs seek to control corruption by predefining appropriate uses of funds and then monitoring the application of money, results-based programs explicitly manage corruption by only paying when results are achieved. In other words, the kinds of corruption that obstruct programs from achieving results cannot, by definition, extract money from funders. The only kinds of corruption that can succeed in stealing funds from a results-based program are either forms that have limited impact on results or forms that are able to misrepresent results (i.e., making it seem that they were achieved when, in fact, they were not).

As a result of this fundamental difference in design, results-based programs do not have direct costs associated with monitoring and controlling corruption. The costs of administering the results-based program – costs associated with measuring and verifying the results – are essentially the same with and without corruption, although more exacting approaches to verification may be required in contexts where corruption is of greater concern.

In reality, results-based programs are not perfectly implemented. Errors in choosing indicators, measuring progress or verifying results can lead programs to over- or underpay. When conditions are favorable to achieving results unrelated to the recipient government's actions, opportunities will open for individuals to divert funds to other purposes. Because payments are ex-post there is always the possibility that a recipient government will achieve results in one period and then use the subsequent payments in unethical or criminal ways. But importantly, the results-based model has a built-in mechanism to detect and respond when outcomes fail to materialize. The costs of corruption in terms of foregone benefits cannot be concealed indefinitely without interrupting the flow of funds and therefore ultimately limiting the amount of harm done by that corruption.

5. Formal comparison of input-tracking and results-based programs

The contrast between input and results-based programs can be formalized to show how input-based payments allow funds to be diverted at the expense of outcomes. Furthermore, it shows how difficult it is to divert funds from results-based modalities unless outcomes are achieved. It further demonstrates why results-based payment models are likely to be preferable under most circumstances and particularly when corruption risks are high.

The Basic Model:

Consider a program where the true production function is

$$\mathbf{O} = \alpha^{*} X \qquad \text{(eq. 1)}$$

for outcome O, inputs X, and transformational (or efficiency) parameter α.

Outcomes are not directly observed. Instead an indicator (**I**) is measured as:

$$\mathbf{I} = \mathbf{O} + \delta \qquad \text{(eq. 2)}$$

where δ represents the difference – positive or negative – between the outcome and the indicator. This parameter (δ) can be further decomposed into three parts – measurement error (e.g., when estimated from a sample); an indication of the "distance" between the proxy and the true outcome (e.g., students taking a test only approximate the actual learning achieved which is the aim of the project); and changes in outcomes not related to the recipient's efforts (i.e. "windfalls" and "bad luck").

The input modality:

The input-based program is designed to apply X units of inputs with α^* efficiency at a price of P_x per unit of X. In reality, only x inputs are applied with efficiency α and at a price that could be lower than P_x by the share of funds that are skimmed off or paid as bribes (ϱ).

The program receives statements of expenditures from the implementation unit for X units of inputs at price P_x. Therefore, the project pays out P_x X.

$$\text{Payout} = P_x \mathbf{X}$$

However, the statement of expenditures may overstate the price and the number of inputs applied. The actual amount of money spent on inputs by the recipient is $(1-\varrho)P_x x$ with the remaining $\varrho P_x x$ being stolen or paid in bribes. The program *infers* that the total outcome produced is equal to α^* X.

From this information we can derive the "diverted funds" and the "failure costs."

Diverted funds are measured in monetary units (e.g. US$) and can be written as the difference between the reported price times inputs and the actual price times inputs:

$$\text{Diverted funds} = P_x X - (1-\varrho)P_x x$$

This can be rewritten as:

$$\text{Diverted funds} = P_x X [-(1-\varrho)P_x X + (1-\varrho)P_x X] - (1-\varrho)P_x x \text{ and decomposed}^7 \text{ as}$$

$$\text{Diverted funds} = \varrho P_x X + (1-\varrho)P_x (X - x) \qquad \text{(eq. 3)}$$

The first term on the right hand side are the funds diverted through graft (ϱ) while the second term represents fund diverted due to charging for more inputs than were actually applied ($X - x$). At one extreme, if there is no overcharging or graft ($\varrho = 0$) then all diversion takes place in the form of undersupplying inputs. At the other extreme, as graft approaches 100%, the difference between programmed and actual inputs becomes insignificant.

Failure costs are measured in terms of foregone benefits (e.g., children who could have been educated for the same amount of money). For the input modality, the foregone benefits are the difference between the outcome expected from the inputs that were paid for (α^* X) and the outcome which occurred as a function of the actual inputs (x) and efficiency with which those inputs were applied (α).

---

[7] The decomposition can be done two different ways, in terms of X or x. For expositional purposes, the two are equivalent. However, if the decomposition were quantified, the choice would alter the share of diversion attributed to the two components. This is similar to the difference between base period and end period normalization in Laspeyres and Paasches price indices and reflects similar issues in the decomposition of wage discrimination gaps (Oaxaca and Ransom 1994). This holds for the next decomposition in this section, as well.

$$\text{Failure costs} = \alpha^{*} X - \alpha x$$

This can be rewritten and decomposed as:

$$\text{Failure costs} = \alpha^{*} X[- \alpha X + \alpha X] - \alpha x$$

$$\text{Failure costs} = (\alpha^{*} X - \alpha X) + (\alpha X - \alpha x)$$

$$\text{Failure costs} = (\alpha^{*} X - \alpha X) + (\alpha X - \alpha x)$$

$$\text{Failure costs} = (\alpha^{*} - \alpha) X + \alpha (X - x)$$

(eq. 4)

The first term on the right hand side represents the failure costs that result from applying inputs less efficiently than expected ($\alpha^{*} - \alpha$) while the second term represents failures costs that result from applying fewer inputs than expected ($X - x$).

These results are summarized in Table 1.

The results-based modality:

The results-based modality uses the same model for inputs (X), outcomes (O) and efficiency (α^{*}). However, this program receives no statements of expenditures. Instead, it receives a report of how much the indicator has changed (I) and pays $P_{\mathbf{I}}$ for each unit of progress. This can be written as:

$$\text{Outcome payment} = P_{\mathbf{I}} \mathbf{I} = P_{\mathbf{I}} (\mathbf{O} + \delta) = P_{\mathbf{I}} (\alpha x + \delta)$$

Taking into account price skimming and undersupply of inputs, the true unit cost of **O** is the *actual* amount of money spent per unit of outcome, that is, $(1-\varrho)P_x x/\mathbf{O}$. So, the price per unit ($P_{\mathbf{I}}$) is likely to differ from the true unit cost by a factor φ as:

$$P_{\mathbf{I}} = [(1-\varrho)P_x x/ \mathbf{O}] + \varphi. \qquad \text{(eq. 5)}$$

If $\varphi > 0$, then the program "overpays" for the outcome. If however $\varphi < 0$ then the program is paying less than the true unit cost.

As before, we can now derive the "diverted funds" and the "failure costs."

In the case of the results-based modality, diverted funds (measured in monetary units) can be written as the difference between the amount paid by the funder and the amount spent by the recipient on actual inputs:

$$\text{Diverted funds} = P_{\mathbf{I}} \mathbf{I} - (1-\varrho)P_x x$$

The outcome payment is made against the indicator which measures the true outcomes with some error as discussed above.

$$\text{Diverted funds} = P_{\mathbf{I}}(\mathbf{O} + \delta) - (1-\varrho)P_x x = P_{\mathbf{I}}\mathbf{O} + P_{\mathbf{I}}\delta - (1-\varrho)P_x x$$

Using equation 5 to substitute for $P_{\mathbf{I}}$ in the first term yields:

$$\text{Diverted funds} = [(1-\varrho)P_x x/\ \mathbf{O} + \varphi]\ \mathbf{O} + P_{\mathbf{I}}\delta - (1-\varrho)P_x x$$

$$\text{Diverted funds} = [(1-\varrho)P_x x/\ \mathbf{O}]\mathbf{O} + \varphi\mathbf{O} + P_{\mathbf{I}}\delta - (1-\varrho)P_x x$$

$$\text{Diverted funds} = (1-\varrho)P_x x + \varphi\mathbf{O} + P_{\mathbf{I}}\delta - (1-\varrho)P_x x$$

$$\text{Diverted funds} = \varphi\mathbf{O} + P_{\mathbf{I}}\delta$$

(eq. 6)

Thus diverted funds in the results-based program are affected by the degree to which the price per unit of outcome is over- or underestimated (first term on the right hand side) and the gap between the indicator and the outcome (δ). This final result has a series of important implications.

First, diverted funds will be greater (1) when the funder overpays ($\varphi>0$ and $P_{\mathbf{I}}$ are larger) and (2) when the indicator (**I**) overestimates outcomes (i.e. $\delta>0$). However, the converse is also true: diverted funds will have to be *smaller* or zero in those cases (1) when the funder underpays ($\varphi<0$ and $P_{\mathbf{I}}$ are smaller) and (2) when the indicator (**I**) *under*estimates **O** (i.e. $\delta<0$). In these cases, there is no "room" to extract funds. These key parameters can be designed or tested to minimize the diversion of funds. For example, choosing a good indicator and designing a system to reduce measurement error can make δ small. As another example, risks of overpayment can be reduced by initially making low offers and iterating upward until the desired impacts are achieved.

Second, when $\varphi>0$ the recipient faces positive incentives to produce more outcomes; but when $\varphi<0$, only honest recipients have an interest in maximizing outcomes (and the payments they receive offset their costs). Dishonest recipients who are only interested in diverting funds, however, will see no profit in generating outcomes when $\varphi<0$ and will either refuse to enter such a contract or default on it.

Third, with input modalities, diverted funds are directly affected by the amount paid in bribes (ϱ), skimmed off by using fewer inputs (X - x), or applied inefficiently (α). By contrast, in the results-based approach, these parameters only influence the potential for diverting funds through their impact on outcomes. With results-based programs, corruption that reduces outcomes will reduce the potential for diverting funds; only corruption that increases outcomes can make space for diverting funds.

For the results-based modality, failure costs (measured as foregone benefits) are simply the difference between the outcomes that were paid for (**I**) and the outcomes actually achieved (**O**). Thus,

$$\text{Failure costs} = \mathbf{I} - \mathbf{O} = \mathbf{O} + \delta - \mathbf{O}$$

$$\text{Failure costs} = \delta \qquad \text{(eq. 7)}$$

In other words, the only failure costs in an outcome program are generated by the measurement errors, distance between the actual outcome and the proxy indicator, or windfalls and bad luck. These results are summarized in Table 1.

Table 1: Cost comparison of input and results-based modalities

	Input modality	Results-based modality
Diverted funds ($)	$\varrho\ P_x X + (1-\varrho)P_x (X - x)$	$\varphi\mathbf{O} + P_{\mathbf{I}}\,\delta$
Failure costs (foregone benefits)	$(\alpha^* - \alpha)\ X + \alpha\ (X - x)$	δ

Comparing the two modalities:

There are five key parameters to consider when comparing the two modalities:

- the share of the price that is stolen or paid in bribes (ϱ)

- the actual efficiency with which the recipient applies inputs (α)

- the difference between reported and actual inputs ($X - x$), and

- the divergence between the indicator and the true outcomes (δ)

- the relationship between the outcome payment and the true cost of the outcome (φ)

The main factor affecting *diverted funds for the input modality* is the difference between reported and actual inputs. The price difference (ϱ) mainly affects the share of the diverted funds attributable to diversion through pricing and diversion through input quantities.

Diverted funds in the results-based modality are neither influenced by the price difference nor by the difference between reported and actual inputs. The potential for diverting funds in this modality is primarily influenced by the level set for the indicator payment and the distance between the indicator and true outcomes. Note also that the choice of payment level has no impact on the failure costs.

The two modalities differ most significantly in terms of failure costs. The *foregone benefits in an input program* increase whenever inefficiencies are high (larger α) or inputs are over reported (X - *x* is big). Thus, input programs spend a large amount of time and energy trying to measure and increase efficiency while also trying to verify the actual amount of inputs that are used. By contrast, the only factor affecting *foregone benefits in an outcome program* is the difference between the indicator and the outcome it is trying to measure. Careful choice of an indicator in terms of accurately reflecting the desired outcome and minimizing measurement error are the ways to reduce the failure costs of an outcome program.

Deciding whether one modality is preferable depends on several parameters, particularly on uncertainty regarding the production function, the accuracy of the indicator, and the payment level. It also depends on whether the funder and recipient are more concerned about diverted funds or foregone benefits. And finally, it depends on the relative costs for monitoring and implementing the programs.

Under the input model funders pay P_x X. They pay the same when there is graft, and when x is smaller than X, but don't need to worry about δ. Essentially, funders get what they pay for minus reductions due to graft, under-delivery, and inefficiency.

With the results-based modality, funders pay for the efficient delivery of the actual outcomes (**O**) and for δ. They pay *less* when there is graft, when x is smaller than X, and when efficiency is low but more when δ is big. Essentially, funders get what they pay for minus (or plus) δ.

Therefore, results-based modalities are almost always preferable in terms both of reducing foregone benefits and reducing diverted funds. Only if the impact of δ is greater than the impact of ϱ, (X- x) and (α*- α) combined should we choose an input-modality. Furthermore, this demonstrates that outcome modalities are preferred especially in circumstances where risks of corruption are high; that is where you think the values of ϱ, (X- x) and (α*- α) are likely to be large (i.e., high corruption low capacity).

These results indicate that vulnerability to corruption by either measure – diverted funds or failure costs – will depend critically on (1) whether a good indicator for the outcome can be found and (2) whether the production function is known and well-specified. This leads to four possible scenarios with direct implications for the preferability of one modality over the other:

1. If a good indicator can be found and the production function is known, then δ and φ can be made small and there is no reason to use an input modality.

2. If a good indicator can be found and the production function is not known, δ can be small but φ might be high. In such cases, the input modality is still no better than the results-based modality because there is no way to accurately estimate or assess α and ϱ. The foregone benefits of the input-modality are affected by α and ϱ but the

results-based modality is *not* affected by the size of φ. Therefore, in this case, the results-based modality is likely to be preferable.

3. If a good indicator cannot be found and the production function is also not known, then we're in a situation where relative costs cannot be judged well. Part of assessing the difference requires considering whether the risks of the input-modality – in which outcomes are inversely related to diverted funds – make sense compared to the risks in the results-based modality for which outcomes are compatible in incentive terms for both honest and dishonest recipients.

4. Only when a good indicator cannot be found and the production function is well known does the input-modality seem to be preferable. However, it is difficult to imagine how the production function can be well-specified without being able to measure the results. So even in this case, paying for changes in a noisy outcome proxy might still, in expectation, be better than trying to control a noisy and costly to monitor production process.

6. When will a results-based program probably do less harm?

We have shown that corruption has many different costs that vary across different modalities and we have explicitly analyzed costs in terms of *diverted funds* and *failure costs* (e.g. not getting the outcomes you paid for). Funds can be diverted from both input-based aid and results-based aid but the diversion is inversely related to outcomes in the former case and positively associated in the second. Failure costs with input-based aid can be significant and are directly related to the extent of corruption in pricing and delivery; whereas failure costs in results-based aid are linked only to difficulties in measuring outcomes and are unrelated to corruption.

Table 2 applies our analysis to sectoral examples, distinguishing those cases for which inputs and outputs are hard or easy to measure. Technical assistance represents an example for which it is hard to measure both inputs and outputs. Advanced education may be a case where it is comparatively easy to measure inputs (number of teachers, buildings, etc.) but harder to measure outputs (e.g., PhD quality). The inputs to service delivery reforms are very difficult to measure (e.g. legal changes, training, reorganizations) but the outputs are reasonably countable and measurable (more children vaccinated, more students taught). Roads, food aid, and basic education are examples where both inputs and outputs are fairly measurable.

Where each example belongs within this table is arguable but it is provided to be illustrative. While road construction is labeled 'easy to measure' on both dimensions – in that it is fairly straightforward to count bags of concrete and see if there is a road there – transport provides many examples of the wrong quality or quantity of concrete or reinforcement being used, resulting in low quality infrastructure. At best these are comparative positionings.

Where inputs are hard to measure, there is considerable danger of diverting resources by undersupplying goods (i.e., X-x takes a large value). Even if inputs are easy to monitor, inputs may still be supplied at a price that allows for graft (large ϱ) or used with low efficiency ($\alpha^* - \alpha$ is large). Such problems are only straightforward to observe ex-post in cases where 'true' values for P_x and α^* are well known, but in turn this implies being able to measure outcomes in order to generate values for the production function $\mathbf{O} = \alpha^* X$. Again, such problems will only be observed, if at all, considerably after payments are made.

Where outputs are hard to measure, there is a risk that δ is large, and payments based on an imperfect proxy indicator **I** may well deliver little in the way of the desired outcome **O**. Even in cases where φ is well calibrated to pay only for an efficient delivery of **O** (i.e. where (X-x), ϱ and $(\alpha^* - \alpha)$ are necessarily small), a large δ creates space for diverting funds. Note also that calibrating φ so that P_I is close to the actual cost of delivery requires 'true' values for P_x and α^*. This implies being able to measure inputs in order to generate values for the production function $\mathbf{O} = \alpha^* X$

Table 3 suggests which modality might be best suited to limit the impact of corruption in relation to the difficulty of measuring inputs and outputs. When outcomes are easy to measure, results-based approaches are preferable. They are more effective at limiting the effects of corruption and delivering results. Only in cases where inputs are easy to measure and outcomes are difficult to measure do input-based approaches possibly have an edge. But even if inputs are easy to measure the input-based approach is flawed because easily measured inputs may still be supplied at a price that allows for graft and be applied with low efficiency. And in cases where outcomes are hard to measure such problems will be difficult to detect *ex-post*. In short, if outcomes are hard to measure, neither results-based nor traditional approaches work very well.

When outcomes are difficult to measure, good programs have to find other mechanisms, such as independent professional reviews and expert standards, to try to control abuses and assure performance. When neither inputs nor outcomes are easily measured, results-based approaches are likely to be preferable. Calibrating payments (P_I) to the actual cost of delivery by minimizing φ will be difficult in such cases but the implications of *underestimating* P_I is to have fewer outcomes rather than more corruption. This suggests that in poor information contexts, the best strategy may be to create a results-based program with a payment that is deliberately underestimated so as to avoid corruption risk. Complementing this with professional reviews of performance would allow funders to adjust the results payment up or down depending on the level of response.

Table 2: Illustration of projects by ease of measuring inputs and outputs

		Inputs	
		Easy to Measure	*Hard to Measure*
Output	*Easy to Measure*	Road construction Food aid Basic education	Service delivery reforms
	Hard to Measure	Advanced education	Technical assistance Regulatory reforms

Table 3: Preferred modalities by ease of measuring inputs and outputs

		Inputs	
		Easy to Measure	*Hard to Measure*
Output	*Easy to Measure*	Results-based COD Aid Output Based Aid	Results-based COD Aid Output Based Aid
	Hard to Measure	Input-based with professional review of performance?	Results-based with price iteration and professional review of performance?

6. Conclusion

A common criticism of result-based aid modalities is that they are somehow more vulnerable to corruption than input-based modalities which monitor inputs and impose specific procedures for procurement and financial accounting. In fact, corruption control strategies applied to input-based aid modalities are often ineffective and - despite their widespread use - largely unproven. Ironically, the input-tracking approach leads to strategies that undermine rather than improve the effectiveness of foreign aid. When scandals erupt over improper use of foreign aid, agencies adopt a standard set of actions that tighten the control of funding in the inputs chain –involving greater financial controls, rigid procurement rules, and enhanced

forensic investigation. While we have little evidence to show that these measures reduce the diversion of funds, we do know that these mechanisms are introduced without reference to their impact on the ability of programs to achieve results.

By contrast, the mechanisms used to measure outcomes (or outputs) in results-based modalities necessarily reduce the scope for corruption. Particularly when outcomes are underpriced, individuals who wish to defraud the program have no incentive to participate. Even when outcomes or outputs are overpriced, those who wish to divert funds from a program can only do so by generating results. Any reduction in outcomes or outputs interrupts the flow of aid funds from which individuals can extract resources for personal enrichment.

This comparison demonstrates a few key differences between the modalities. First, the information requirements of input modalities are much higher than outcome modalities. The input modality requires information about prices, quantities of inputs and efficiency of application while the results-based modality requires only information about the outcomes. Second, people who want to defraud an input-based program will thrive by reducing the program's efficiency – forcing prices up and reducing deliveries. By contrast, someone who wants to defraud a results-based modality program can only do so by improving efficiency – generating more outcomes so that more aid money is disbursed and available for diversion. Finally, honest agents in an input-based program face the same or higher transaction costs than dishonest agents. After all, honest agents have to do their jobs (achieve results) as well as honestly comply with all reporting and auditing requirements. Dishonest agents don't have to do their jobs and can make up reports and invoices. By contrast, in an outcome-based program, honest agents face lower costs than dishonest agents. Honest agents can focus on doing their jobs – and generate results which trigger disbursements. It is the dishonest agents who have to find a way to achieve the outcomes and then, in addition, find ways to divert funds.

Whenever it is possible to measure an outcome directly or an output that is closely linked to that outcome, the results-based modality will be preferable. Only in cases where it is possible to measure inputs but difficult to measure outputs do the advantages of an input-tracking approach begin to materialize. However, when inputs are also difficult to measure, the results-based approaches are again the preferred option.

The critical factor in choosing among aid modalities is to recognize that failure costs – the foregone benefits of a program that has been defrauded – are the true costs of corruption. Once attention is focused on whether or not the program is achieving results it is not only possible to improve the program's effectiveness but also possible to limit the impact of corruption on development.

Bibliography

Africon (2008) Unit Costs of Infrastructure Programs in Sub-Saharan Africa AICD Background Paper 11, Washington, DC: World Bank.

Ahmedov, H. and C. Kenny (2010) World Bank Infrastructure Programs in Developing Countries: Analysis of Market Concentration, mimeo, Washington, DC: World Bank.

Alexeeva, V., G. Padam and C. Queiroz (2008) Monitoring Road Works Contracts and Unit Costs in Sub-Saharan Africa for Enhanced Governance, mimeo, Washington, DC: World Bank.

Auriol, E, (2006) Corruption in procurement and public purchase. *International Journal of Industrial Organization*, 24(5):867-885.

Bardhan, P. (1997) Corruption and Development: A Review of Issues. *Journal of Economic Literature,* 35:1320-46.

Basinga, P., Gertler, P. J., Binagwaho, A., Soucat, A. L., Sturdy, J., & Vermeersch, C. M. (2011). Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: an impact evaluation. *The Lancet*, 377(9775):1421-1428.

Berkman, Steve. 2008. *The World Bank and the gods of lending*. Kumarian Press.

Birdsall, N. and W. Savedoff (2010) *Cash on Delivery: a New Approach to Foreign Aid*. Washington, DC: Center for Global Development.

Belli, P. and L. Pritchett (1995) Does Good Economic Analysis Improve Program Success mimeo, Operations Policy Division, Washington, DC: World Bank.

Casartelli, G. and E. Wolfstetter (2007) World Bank Policy on the Selection and Employment of Consultants: Study of its Effectiveness, mimeo, Washington, DC: World Bank.

Davis, J. (2004) Corruption in Public Service Delivery: Experience from South Asia's Water and Sanitation Sector. *World Development* 32(1).

Devarajan, S., L. Squire and S. Suthiwart Narueput (1997) Beyond the Rate of Return, World *Bank Research Observer* 12(1).

Dollar, D. and V. Levin (2005) Sowing and Reaping: Institutional Quality and Program Outcomes in Developing Countries World Bank Policy Research Working Paper 3524. Washington, DC: World Bank.

Eichler, R., Auxila, P., Antoine, U., & Desmangles, B. (2007). Performance-based incentives for health: six years of results from supply-side programs in Haiti. CGD Working Paper 121, Washington, DC: Center for Global Development.

Gulati, M. and M. Rao (2006) Checking Corruption in the Electricity Sector, mimeo, Washington, DC: World Bank.

Hines, J. (1995) Forbidden Payment: Foreign Bribery and American Business After 1977, NBER Working Paper 5266. Cambridge, MA: National Bureau of Economic Research.

Isham, J., D. Kaufmann and L. Pritchett (1997) Civil Liberties, Democracy and the Performance of Government Programs World Bank Economic Review 11, 2.

Kenny, C. (2007) Corruption, Construction and Developing Countries, World Bank Policy Research Working Paper 4271. Washington, DC: World Bank.

Kenny, C. (2008) What is Effective Aid? How Would Funders Allocate it? *European Journal of Development Research* 20, 2

Kenny, C. (2009) Measuring and Reducing the Impact of Corruption in Infrastructure *Journal of Development Studies* forthcoming.

Kenny, C. and D. Williams (2001) What Do Economists Know About Economic Growth – Or Why Don't They Know Very Much? *World Development* Vol. 29 No.1.

Kenny, C. and M. Musatova (2009) What Drives Contracting Outcomes in Infrastructure Procurement? Mimeo, Washington, DC: World Bank.

Kremer, M., N. Chaudhury, F. Rogers, K. Muralidharan and J. Hammer (2005) Teacher Absence in India: A Snapshot, *Journal of the European Economic Association* 3:2-3

Lim, S. S., Stein, D. B., Charrow, A., & Murray, C. J. (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.

Moss, T. (2005) Ten Myths of the International Finance Facility, CGD Working Paper 60, Washington, DC: Center for Global Development.

Mumsen, Y. and Kenny, C. (2007) Output-based aid in infrastructure: a tool for reducing the impact of corruption, World Bank OBApproaches note 16. Washington, DC: World Bank.

Olken, B. (2004) Monitoring Corruption: Evidence from a Field Experiment in Indonesia, NBER Working Paper 11753.

OPCS (2009) Investment Lending Reform: Concept Note, mimeo, Washington, DC: World Bank.

Pearson, M., M. Johnson, and R. Ellison (2010) Review of major Results Based Aid (RBA) and Results Based Financing (RBF) schemes. London: DFID Human Development Resource Centre.

Pohl, G. and D. Mihaljek (1992) Program Evaluation and Uncertainty in Practice: A Statistical Analysis of Rate-of-Return Divergences of 1,015 World Bank Programs**,** *World Bank Economic Review* 6: 255-277.

Pritchett, L. (2000) The Tyranny of Concepts. World Bank Policy Research Working Paper 2341. Washington, DC: World Bank.

Rioja, F, (2003) Filling Potholes: Macroeconomic Effects of Maintenance Versus New Investments in Public Infrastructure. *Journal of Public Economics*, 87:2281-2304

Rogers, F. and E. Vegas, (2009) No More Cutting Class? Reducing Teacher Absence and Providing Incentives for Performance (February 1, 2009). World Bank Policy Research Working Paper 4847. Washington, DC: World Bank.

Savedoff, W. (2011) Incentive Proliferation? Making Sense of a New Wave of Development Programs. CGD Essay, Washington, DC: Center for Global Development.

Soraya, G. (2009) Indonesian Experience Governance and Anticorruption Bali Urban Infrastructure Program (BUIP) presentation to the World Bank GAC in Programs Group, 4/21/09

Soreide T. (2006) Procurement Procedures and the Size of Firms in Infrastructure Contracts Paper Prepared for the World Bank Annual Conference on Development Economics, Tokyo 29-30 May.

Strassman P. and J. Wells (1988) *The Global Construction Industry: Strategies for Entry Growth and Survival* London: Unwin.

Transparency International (2001) *Global Corruption Report 2001*. Berlin: Transparency International.

Vawda, A., P. Moock, J. Gittinger and H. Patrinos (2001) Economic Analysis of World Bank Programs and Program Outcomes, World Bank Policy Research Working Paper 2564, Washington, DC: World Bank.

Wong, S. and S. Guggenheim (2005) Community-Driven Development: Decentralization's Accountability Challenge in World Bank (ed) *East Asia Decentralizes: Making Local Government Work* World Bank: Washington DC.

World Bank (2006) *Infrastructure at the Crossroads: Lessons from 20 Years of World Bank Experience*, Washington DC: World Bank

World Bank (2007) Indian Road Construction Industry: Ready for Growth? mimeo, South Asia Regional Transport Team.