

Results-Based Financing Evidence Review and Key Lessons Learned

I. Table of Contents

1. Introduction	2
2. Evidence on Results-based financing (RBF).....	2
1. Impact of RBF	3
2. Impact of programs using RBF.....	4
3. Guides for designing and implementing RBF	8
1. Preparation for RBF.....	9
<i>RBF value-add.....</i>	<i>9</i>
<i>Contextual analysis.....</i>	<i>10</i>
<i>Stakeholder engagement.....</i>	<i>10</i>
2. RBF Strategy.....	10
<i>Instruments.....</i>	<i>10</i>
<i>Design.....</i>	<i>11</i>
<i>Implementation</i>	<i>12</i>
4. Lessons from theory and practice of RBF	13
1. RBF's value-add.....	13
2. Lessons on assessing the conditions for RBF	14
3. Lessons on the RBF Design.....	15
4. Lessons on the implementation of RBF programs.....	17

Authors: Avnish Gungadurdoss, Benjamin Stephens, Lucia Perez-Villar, Sofia Escobar



I. Introduction

The use of Results-based Financing (RBF) has dramatically expanded over recent years, along with the body of literature on effectiveness and lessons for its use. To support bilateral and multilateral funders' RBF initiatives, Instiglio prepared an annotated bibliography of relevant resources on RBF in the development sector. The objective of this annotated bibliography is to (1) provide all funders' staff with an accessible summary of the relevant rigorous evidence on RBF's effectiveness, (2) provide the staff interested in applying RBF an overview of the relevant guides to RBF's effective use and (3) provide the staff interested in understanding experiences with RBF a summary of key insights and lessons from a sample of past programs and reviews.

To ensure the annotated bibliography is useful for funders' work, Instiglio targeted specific examples relevant to what tend to often be priority sectors, including energy, WASH, transportation, land and agriculture, environment, workforce development, education and public sector reform. To identify relevant sources, in addition to drawing on Instiglio's internal library of relevant sources, we examined the publications developed and reviewed by the leading RBF practitioners and researchers globally, including:

- The Government Outcomes Lab (GOLAB)
- The UK's Foreign, Commonwealth & Development Office (previously Department for International Development)
- The Center for Global Development (CGD)
- The Brookings Institution
- The Global Partnership for Results-Based Approaches (GPRBA)

Further papers were identified through cross-referencing relevant papers.

The review identified and drew on papers relating to the following three topics which also provide review's structure.

1. Evidence for RBF's impact

This section provides a review of 1) evidence on the impact of RBF compared to traditional financing approaches (three papers) and 2) evidence on the impact of programs using RBF (12 papers). By summarizing this evidence, this section and the associated references will be a useful resource for funders staff, such as economic analysis teams, interested in understanding the evidence base for RBF.

2. Guides for designing and implementing RBF

This section reviews publications which provide step-by-step guides on how to use RBF to achieve a greater impact. This review covers four guides from a sector agnostic perspective and three sector specific guides to provide guidance on how to apply RBF in the Energy, Agriculture, and Workforce development sectors. All the selected publications provide guidance on two key topics: 1) RBF preparation: RBF value-add and how to assess RBF's feasibility for a specific context, and 2) RBF strategy: design and implementation. By reviewing these papers, this review will provide a useful reference for funders' staff interested in identifying practical guidance on the use of RBF in different contexts.

3. Lessons for effective RBF use

This section explores the main lessons from RBF projects to provide funders' staff with insights relevant to future RBF projects. The papers reviewed include case studies and anecdotal evidence (literature reviews, interviews, surveys, and focus groups discussions, etc.), that provide practitioners lessons learned, success factors, and challenges in the use of RBF in different sectors including health, education, workforce development, energy and environment. The selected publications include lessons on: 1) RBF's value-add and how the four channels through which RBF can drive greater effectiveness in developing spending have been achieved in practice. 2) the three main phases of an RBF project: preparation, design, and implementation, and 3) the provision of incentives (in general, not only focused on RBF) within organizations to have the right incentive structure and improve performance.

2. Evidence on Results-based financing (RBF)

Demand for innovative forms of financing, like RBF, has grown in the past years and evidence about its effectiveness to improve development results compared to traditional input-based approaches is needed. This section first provides a review of evidence on the impact of RBF compared to traditional financing approaches, reviewing four papers providing



rigorous evidence on RBF's effectiveness relative to traditional input-based financing. Second, we review evidence from eleven papers on the impact of programs using RBF but without any comparison to non-RBF programs (i.e. which do not attempt to make causal estimates of RBF's impact). The review draws exclusively on rigorous impact evaluations using randomized control trials (RCTs) or statistical methods including difference-in-difference and quasi-experimental methods.

I. Impact of RBF

While the use of RBF continues to grow, the evidence of its effectiveness remains limited. Only four papers were found that present evidence on RBF's impact compared to traditional funding. All four papers focus on the health sector, specifically covering the use and quality of child and maternal care services in health-care facilities across different countries. The first three papers detailed below use a difference-in-difference methodology and the fourth is an RCT.

In general, the papers reviewed provide some evidence of RBF's potential to drive greater effectiveness but this is not demonstrated across all experiences and further rigorous evidence would be valuable. None of these papers provide evidence on how RBF leads to these improved results, a key knowledge gap (e.g., intensity of the incentive, incentivized agent, sector and beneficiary population). Specific insights from the papers are the following:

1. Basinga, P., et al. (2011) assessed the effect of a Payment for Performance (P4P) program of health-care providers on the use and quality of child and maternal care services in health-care facilities in Rwanda. To do so, 166 facilities were randomly assigned at the district level either to begin P4P funding or to continue with traditional input-based funding. The main outcome measures were prenatal care visits and immunization. The authors use a multivariate regression specification of the difference-in-difference model in which an individual's outcome is regressed against a dummy variable, indicating whether the facility received P4P that year, and a series of individual and household characteristics. The analysis concludes that the use of P4P was associated with improved results, with facilities in the intervention group experiencing a 23% increase in the number of institutional deliveries and a 132% increase in the number of preventive care visits by children between 24 and 59 months of age. The P4P scheme in Rwanda had the greatest effect on those services that had the highest payment rates and needed the least effort from the incentivized agent¹.
2. Bonfrer, I., Van de Poel, E., and Van Doorslaer, E. (2014) assessed a Performance-Based Financing (PBF) program in Burundi which aimed to improve the utilization and quality of maternal and child care. This paper analyses the effect of PBF through a difference-in-difference approach that uses the 2010 Burundi Demographic and Health Survey to assess the impact in variables like quality of care provided during antenatal care (ANC) visits, the probability of an institutional delivery, and the probability of a child being fully vaccinated. The authors identify the effects of PBF by comparing changes in the outcome measures in provinces with PBF (treated) to changes in provinces without PBF (controls). The study concludes that RBF improved the utilization and quality of maternal and childcare in Burundi: the probability of an institutional delivery increased significantly with 4 percentage points as well as a significant increase of 4 percentage points in the probability of a child being fully vaccinated, with effects more pronounced among the poor.
3. Friedman, J., Qamrudding, J., Chansa, C., and Kumar, A. (2016) reviewed a comprehensive impact evaluation of an RBF pilot project in Zambia (Zambia's Health Results-Based Financing or HRBF). To measure the effectiveness of RBF, the study had a three-arm evaluation that tested RBF against an enhanced financing arm (i.e., treatment group) and a pure control group. Quantitative data for the impact evaluation at household and facility level was collected at baseline, implementation stage, and endline from 10 RBF intervention districts; 10 Control 1 (C1) districts; and 10 Control 2 (C2) districts. Districts were selected based on district-matched randomization. The research question "what is the causal effect of the Zambian HRBF on the population health indicators of interest?" was addressed through the quasi-experimental evaluation method of district-level matched difference-in-differences. The RBF districts performed better than control districts in most of the indicators. Institutional deliveries in RBF districts increased by approximately 13 percentage points relative to the pure control districts (C2). Results for deliveries by skilled providers also showed improvements in both the RBF and C1 districts relative to C2. One of the most important gains in the RBF arm was the timing of the first ANC visit which was earlier by two weeks as

¹ Incentivized agent is the agent whose payments are contingent upon results.



compared to both controls. Large part of this success was that the Zambia RBF pilot program benefited from a process evaluation (PE) system which provided regular updates and insights on the implementation to allow for midcourse corrections and evidence-based policy and planning.

4. Walque, D., et al. (2017) present the results of an impact evaluation in Cameroon that seeks to isolate the role of specific components of the Performance-based financing. Four² evaluation groups were established to measure the effects of each component that was studied. These four groups were formed by randomizing Medicalized Health Centers, or primary health centers with a medical doctor on staff, and Integrated Health Centers (primary health care centers without a doctor). In general, the results indicate that PBF was an efficient mechanism to channel payments and funding to the provider level, leading to significant increases in utilization in the performance-based financing for several services, such as child and maternal vaccinations (provision of tetanus toxoid vaccine for children improved by a monthly average of almost 21.5 vaccinations, provision of postnatal care in facilities improved by 4,309 facilities, and modern contraception provision³ improved by 9,240 on average). However, for many outcomes, the differences between the PBF group and the group that received financing based on activities (not linked to performance) are not significant.

Even though the evidence on RBF is limited and focused on the health sector, the results presented above from the four papers reviewed, demonstrate the potential of RBF to support improved results for a variety of health indicators. However, the papers also highlight some limitations of RBF. For instance, the impact evaluation in the health sector in Cameroon (World Bank Group, 2017) found that the difference of antenatal care visits between the group implementing RBF approaches and the group using traditional input-based financing were not significant. This could be because the difference between these two groups was not clear and obvious in the intervention design, making it difficult for staff and management to fully grasp the differences and might have caused errors in the estimates. Also, the report relied on quantitative household and health facility surveys. A companion qualitative study was posteriorly conducted, and its analysis is still pending, which could help further assess program impacts – or their absence – and shed light on possible mechanisms to strengthen impact.

Similarly, in Burundi (Social Science & Medicine, 2014), there was no evidence that PBF affects the likelihood of the mother receiving more than one ANC visit during pregnancy, or of an ANC visit to occur in the critical first trimester. The authors argue this may be because it is harder for PBF to increase utilization of services that depend on patient choices than services that are under the provider's control. They also argue that initiation of care takes more effort than its continuation, suggesting that policy makers might want to increase the PBF unit payments for first time antenatal care use. Last, in Rwanda (The Lancet, 2011), there was no change estimated with the P4P intervention in the probability that women received any prenatal care or in the probability that they had more than four prenatal care visits. This could be caused by several limitations. For instance, the original randomized design was compromised by the decentralization process⁴, which could have inadvertently caused some bias in the estimates. Also, errors by individuals who were interviewed could have affected the accuracy of the estimates, particularly those focused on women health services.

2. Impact of programs using RBF

This section reviews results from eleven programs that used RBF in sectors across education, health, workforce development, agriculture, and water utilities. These studies all used rigorous evidence to assess the program's impact but do not attempt to isolate the effect of RBF as a funding modality. All papers were published in the past eight years and largely relate to programs funded by governments with the support of international multilaterals and donors with the objective of strengthening service delivery through performance-based contracts⁵. Out of the total 12 papers analyzed under this section, there are 11 that are either RCT's or use difference-in-difference. Only one of the papers (Mugisha, S., 2007) uses a stochastic frontier model.

² Specifically, the evaluation compared four arms: (1) the standard PBF package (T1), (2) the same level of financing as T1 but not linked to performance, and with the same levels of supervision, monitoring, and autonomy as PBF (C1), (3) no additional resources or autonomy, but the same levels of supervision and monitoring as PBF (C2), and (4) pure comparison (C3).

³ Modern contraception included women provided with contraceptive implants, injectables, oral contraceptive pills, and the intrauterine device.

⁴ just before implementation of the baseline survey, the district boundaries were redefined by the government in a decentralization process. As a result, some of the districts selected for the assessment were combined with districts that already had the existing P4P schemes

⁵ The only program that implemented an impact bond was the Educate Girls DIB.



All papers found that the assessed RBF programs succeeded in improving results, driven by factors including improved attention to results and motivating accountability and institutional deliveries, though the extent of these improvements were mixed. We now summarize the results of these papers by sector.

In education, two RBF programs were reviewed, where both were linked to higher enrollment, participation rates, and learning gains. First, IDinsight (2018) provides rigorous evidence on the Educate Girls Development Impact Bond (DIB) which operated in the state of Rajasthan 2015-18. The Educate Girls Development Impact Bond is a joint project between the Children's Investment Fund Foundation, Educate Girls, the UBS Optimus Foundation, Instiglio, and IDinsight focused on enrolling out-of-school girls and improve quality education for children in rural, remote and marginalized communities of India. IDinsight found the DIB was associated with significant gains in enrolment and learning, achieving 116% of the enrolment target and 160% of the learning target were achieved by the end of the program. While Educate Girls was consistently on track to meet the enrollment target throughout the DIB, progress against the learning target lagged behind. Two years into the three-year DIB, Educate Girls had reached just half the target. However, large increases in year 3 drove them to exceed the final target by 60%. Learning gains, which accounted for 80% of the final DIB payments, were measured in an RCT. Enrollment of out-of-school girls, which accounted for 20% of the payments, was calculated as the percent of eligible out-of-school girls in treatment villages enrolled by the end of the program.

In contrast to the significant results of the Educate Girls DIB, the Government-Administered Pilot Teacher Performance Pay Program in Punjab, Pakistan (World Bank, 2015) achieved limited results. The program offered yearly cash bonuses to teachers in a sample of public primary schools linked to three school-level indicators: the gain in student exam scores, the gain in school enrollment, and the level of student exam participation. While the program increased school enrollment by 4.1% and student exam participation rates by 3.4 percentage points, both in the third year, the analysis did not find that the program increased student exam scores in any year. The authors present three hypotheses for the absence of positive mean impacts on student exam scores. First, non-school or school factors outside the control of teachers may make raising student exam scores difficult even with greater teacher effort. Targeting schools with the lowest mean student exam scores in the province places the program in circumstances that may be especially handicapping. Second, teachers may lack the knowhow to raise student exam scores. Third, teachers may optimally choose to direct their effort at those incentivized margins that maximize payoffs net of effort costs.

In the health sector, three programs were identified that used RBF programs in targeted national health systems and health institutions with overall positive effects. These were the RBF Intervention in Afghanistan, Argentina's Plan Nacer, and Danish Pay For Performance (i.e., P4P). The RBF Intervention in Afghanistan (The World Bank Group, 2014) was introduced in 2009, where health facility workers were provided additional financial incentives for providing increased quantity and quality of services from baseline targets. Many methodologies were used to calculate the impact of the intervention, including descriptive statistics and difference-in-difference analysis. The program was found to generate large improvements in health indicators since the treatment group had a higher coverage of all indicators. However, the intervention did not have a statistically significant positive effect overall. Possible reasons for the lack of treatment effect may be due to issues with fidelity of intervention, inherent issues in the study design, spillover and/or contamination and competition by the comparison group, or unobserved changes in the comparison group.

Argentina's *Plan Nacer* (The World Bank Group, 2014) provides insurance for maternal and child health care to uninsured families. The program allocates funding to provinces based on enrollment of beneficiaries and adds performance incentives based on indicators of the use and quality of maternal and child health care services and health outcomes. In order to analyze the impact of the program, the authors use data from birth records in seven Argentine provinces for 2004 to 2008 and exploits the geographic phasing in of *Plan Nacer* over time using difference-in-difference models. The paper finds that the program increases the use and quality of prenatal care as measured by the number of visits and the probability of receiving a tetanus vaccine. Specifically, the beneficiaries' probability of low birth-weight is estimated to be reduced by 19% and beneficiaries have a 74% lower chance of in-hospital neonatal mortality in larger facilities. Nevertheless, some of the results were not translated to the population level and were only seen by a focalized sample group.

The Danish Pay for Performance program was introduced in 2009 in the Region of Southern Denmark, where performance payments were distributed from the region to the hospitals based on the mean of each hospital's department-level performance during the first three quarters of each year. To test whether redistributing performance payments to the department level is associated with higher performance rather than keeping performance payments at



the hospital level, the authors apply difference in differences analysis. The coefficient of interest is the department-level incentive which measures the difference in performance between hospital departments with and without a direct department-level incentive for performance after incentive payments were introduced. The coefficient is positive and statistically significant in all models. The authors refer to this difference as the incentive effect. Specifically, results indicate that performance was on average 5 percentage points higher at hospital departments facing a direct financial performance incentive. The results thus reflect that the organizational level does indeed matter for the effectiveness of P4P schemes.

In workforce development, two programs were reviewed. First, the Employment Fund Intervention in Nepal was reviewed, a program operated over 2008-2015 by Helvetas, a Swiss NGO, in partnership with the Government of Nepal drawing on World Bank financing, which aimed to train the youth for a more inclusive labor force, with a special initiative focused on women. It was one of the largest youth training initiatives in the country, serving almost 15,000 youth annually with funding of approximately USD 28 million. A World Bank evaluation (2012) employed a difference in difference technique to estimate the causal effects of the EF training program on various outcomes. The authors found that after three years, the program had positively improved employment outcomes by providing skills training and employment placement services for more than 40,000 Nepalese youth. Participation in the Fund's training program led to an increase in non-farm employment of 15 to 16 percentage points for an overall gain of about 50%.

Second, the World Bank (2021) did a meta-analysis of short-term education and skills training programs for out-of-school youth. This extensive review included youth training programs in Brazil, Colombia, the Dominican Republic, Liberia, Nepal, Uganda, the United Kingdom, and the United States, many of which are among the most widely quoted programs in the literature. Evidence on the Short-Term Education and Skills Training Programs for Out-of-School Youth (World Bank, 2021), shows that the scarcity of appropriate evaluations made it difficult to conclude whether RBF approaches are effective when compared to others for incentivizing the desired behaviors and outcomes. However, results indicate that for RBF programs where impact data exists, most demonstrate a net positive impact on beneficiaries, at least in the shorter-term and in lower-income country contexts. Additionally, the benefits to female enrollees, in the form of post-program employment and earnings, typically exceed those for males.

In agriculture, two performance-based prize competitions were analyzed. First, the Nigeria Aflasafe Project used a Pay-for-Results prize competition, to incentivize the private sector to develop the supply base of Aflasafe-treated or aflatoxin-compliant (AT/AC) maize. Specifically, the prize competition awarded participating maize aggregators a prize of US\$18.75 for each metric tonne of AT maize that they procured from smallholder farmers. To assess the project's impact on smallholders, the evaluation assessed the impact on smallholders using a quasi-experimental design. By the end of project year 3, the supply of AT/AC maize increased, with more than 13,000 farmers producing maize for sale and 57% of smallholders in the treatment group applied Aflasafe on at least one plot. Additionally, the evaluation found that smallholders in the treatment group earned US \$318 more than smallholders in the comparison group. However, results were not sustained after the program ended: participants could not maintain the market for the products and access to financing.

Second, the AgResults Kenya challenge project aimed to incentivize private sector firms to develop and improve on-farm storage (OFS) technologies using a pay-for-results prize competition. This approach included qualitative methods to assess the project's impact, an interrupted time series design to estimate impacts of the project on adoption of improved OFS, and a difference-in-differences design to compare the outcomes of smallholder farmer adopters of the technology with a matched group of non-adopters before and after the project. The study finds that adopters reduced pesticide dust usage by 36 percentage points, and many reported that reduced pesticide use was the key motivation for adopting improved OFS. The competitors' investments in developing distribution networks and promoting improved OFS increased adoption by 23 percentage points in Eastern Kenya (with 28% of farmers reporting that they adopted in total) and 6 percentage points in Rift Valley region (with 10% of farmers reporting they had adopted in total) compared to would have happened in the project's absence.

Finally, two papers in the **water utilities sector** were identified, where both analyzed the effect of incentive applications in various incentive systems of the National Water and Sewerage Corporation (NWSC) in Uganda. The first study (Mugisha, 2013) utilizes cross-sectional data from utilities of the National Water and Sewerage Corporation of Uganda to test application of extrinsic incentive theory on non-revenue water reduction. Applying a multiple regression model, the authors find that utilities with higher levels of promised incentive payments have a higher



likelihood of improving their billing efficiencies, depending on the number of staff employed, level of service coverage and production per connection. Similarly, empirical evidence from Ugandan water utilities (Mugisha, 2007) suggests that reform initiatives that have incorporated significant use of incentives, have had positive impacts. Notably, service coverage increased from 52% to 70% and water network coverage increased by 49% (1,117 km of water mains extensions, primarily from internally generated funds). In addition, new connections increased from 4,317 to 28,312 per year. As a result, total connections were up from 60,826 to 148,312. Metering efficiency (proportion of metered accounts to total accounts) has increased from 70% to 99.6%, while connection efficiency (proportion of active connections to total connections) has improved from 71% to 93.9%. The evidence from Ugandan water utilities suggests that the mechanisms incentivized cost-effective achievements of energy saving by setting performance criteria that could be easily measured and monitored.

The findings described here are summarized in Table I below.

. Summary of the impact of programs using RBF across different sectors

Sector	Program	Description	Insights from evaluations
Education	Educate Girls Development Impact Bond (DIB) in Rajasthan, India (2015-2018)	This DIB focused on enrolling out-of-school girls and improving quality of education for children in rural, remote, and marginalized communities of India. This was led by Educate Girls along with UBS Optimus Foundation (UBSOF) and Children's Investment Fund Foundation (CIFF).	The DIB was associated with significant impact, achieving 116% of the enrolment target and 160% of the learning target by the end of the program. ⁶
Health	Argentina's Plan Nacer (2007-2012)	This government-led program provided insurance for maternal and child healthcare by funding provinces based on the enrolment of beneficiaries. It also added performance incentives based on indicators of maternal and child healthcare services and health outcomes.	Beneficiaries were less prone to have low birth weight and a lower chance of in-hospital neonatal mortality in larger facilities. Nevertheless, some of the results were focalized on the sample group of the analysis. ⁷
	RBF Intervention in Afghanistan (2010-2012)	An intervention where health facility workers were provided additional financial incentives for providing increased quantity and quality of services from baseline targets. This program was implemented by Helvetas Swiss Intercooperation in cooperation with the World Bank.	The program was found to generate large improvements in the coverage of health services. However, the intervention did not have a statistically significant positive effect overall. ⁸
Workforce development	Employment Intervention in Nepal (2018-2021)	A program drawing on World Bank financing, which aimed to train the youth for a more inclusive labor force, with a special initiative focused on women.	The program positively facilitated the provision of skills training and employment placement services for more than 40,000 Nepalese youth. Participation in the training program led to an increase in non-farm employment of 15 to 16 percentage points for an overall gain of about 50%. ⁹

⁶ Kitzmuller., McManus., Shah., Sturla. (2018). Evaluating Learning and Enrollment Gains for the Educate Girls Development Impact Bond. *IDinsight*.
https://static1.squarespace.com/static/5b7cc54eec4eb7d25f7af2be/t/5bce543ee4966befb66c13e5/1540248811132/EG+DIB+Project+Report_Final_High+Res_Web.pdf

⁷ Gertler, P & Giovagnoli, P & Martinez, S. (2014). Rewarding Provider Performance to Enable a Healthy Start to Life: Evidence from Argentina's Plan Nacer. *The World Bank Group*, 6884.
https://www.rbhealth.org/sites/rbf/files/Rewarding%20provider%20performance%20to%20enable%20a%20healthy%20start%20to%20life%20-%20evide..._0.pdf

⁸ The Royal Tropical Institute. (2015). An Impact Evaluation of the Results-based Financing Intervention in Afghanistan. <https://www.kit.nl/wp-content/uploads/2018/10/RBF-IE-2015-Final-Report.pdf>

⁹ Chakravarty & Lundberg & Nikolov & Zenker. (2016). The Role of Training Programs for Youth Employment in Nepal. *The World Bank Group*, 7656.
<https://openknowledge.worldbank.org/bitstream/handle/10986/24232/The0role0of0tr00the0employment0fund.pdf?sequence=5>



	Review of the evidence on short-term education and skills training programs for out-of-school youth (2013)	The World Bank did a meta-analysis that included youth training programs in Brazil, Colombia, the Dominican Republic, Liberia, Nepal, Uganda, the United Kingdom, and the United States, many of which are among the most widely quoted programs in the literature.	This work argued that the scarcity of appropriate evaluations made it difficult to conclude whether RBF approaches are effective. However, results indicate that for RBF programs where impact data exists, most demonstrate a net positive impact on beneficiaries, at least in the shorter-term and in lower-income country contexts. ¹⁰
Agriculture	Nigeria Aflasafe Project (2014-2018)	An AgResults project that aimed to incentivize the private sector to develop the supply base of Aflasafe-treated or aflatoxin-compliant (AT/AC) maize	The supply of AT/AC maize increased, and the smallholders increased their income. However, these results were not sustained after the program ended. ¹¹
	AgResults Kenya challenge project (2014-2018)	This AgResults initiative had the objective of incentivizing private sector firms to develop and improve on-farm storage (OFS) technologies.	This intervention reduced pesticide dust usage by 36 percentage points and increased the adoption of OFS by 23 percentage points in Eastern Kenya and 6 percentage points in Rift Valley. ¹²
WASH	Incentive systems of the National Water and Sewerage Corporation (NWSC) in Uganda (2000–2006)	This initiative consisted of innovative managerial techniques aligning management and corporate performance goals. Performance targets were set for strategic areas of NWSC (e.g., water production and sewerage, customer service, revenue improvement) and actions were outlined pursuant to these goals. ¹³	Utilities with higher levels of promised incentive payments had a higher likelihood of improving their billing efficiencies ¹⁴ . Similarly, reform initiatives that have incorporated significant use of incentives had positive impacts. ¹⁵ Finally, evidence suggests that these mechanisms incentivized cost-effective achievements of energy saving by setting performance criteria that could be easily measured and monitored.

3. Guides for designing and implementing RBF

This section reviews publications providing step by step guidance on whether and how to use RBF. RBF is defined across most guides as an approach to driving greater impact by tying funds of social programs to results rather than to activities and inputs (Instiglio & GPRBA, 2018). The guides selected for this review are written by leading RBF practitioners including the Government Outcome Lab, World Bank, the UK’s Foreign, DFID (currently known as Foreign, Commonwealth & Development Office - FCDO) and the Education Outcome Fund. In particular, this review covers the following five sector agnostic guides:

- 1) The Government Outcomes Lab (2017). *How to Guide Procurement: A technical guide to good procurement practice in outcome-based commissioning*. This guide provides recommendations for local governments and agencies on Outcome Based Commissioning (i.e. RBF) based on advice from practitioners and experts with RBF experience.

¹⁰ Marguerite Clarke, Meghna Sharma, and Pradyumna Bhattacharjee. (2021). Review of the Evidence on Short-Term Education and Skills Training Programs for Out-of-School Youth with a Focus on the Use of Incentives. World Bank Group. <https://openknowledge.worldbank.org/bitstream/handle/10986/35045/Review-of-the-Evidence-on-Short-Term-Education-and-Skills-Training-Programs-for-Out-of-School-Youth-with-a-Focus-on-the-Use-of-Incentives.pdf?sequence=5>

¹¹ Abt Associates. (2020). AgResults Impact Evaluation Report: Nigeria Aflasafe Challenge Project. *AgResults*. <https://agresults.org/learning/37-evaluation-final-report-nigeria-aflasafe-challenge-project/file>

¹² Abt Associates. (2019). AgResults Evaluation: Kenya On-Farm Storage Challenge Project. *AgResults*. <https://agresults.org/learning/48-evaluation-final-report-kenya-on-farm-storage-challenge-project/file>

¹³ Matta, Nadim. 2003. Uganda: Turn-around of the National Water and Sewerage Corporation. Africa Region Findings & Good Practice Infobriefs. *World Bank*. (228) <https://openknowledge.worldbank.org/handle/10986/9729>
License: CC BY 3.0 IGO

¹⁴ Mugisha, S. (2013). Applying incentives to increase revenue water in urban systems. *Journal of Water Supply: Research and Technology*. 62 (5) 268–278. <https://doi.org/10.2166/aqua.2013.059>

¹⁵ Mugisha, S. (2007). Effects of incentive applications on technical efficiencies: Empirical evidence from Ugandan water utilities. *Utilities Policy*. 15 (4) 225-233. <https://doi.org/10.1016/j.jup.2006.11.001>



- 2) Instiglio and World Vision. (2017). *A practitioner's guide to Results-Based Financing: Getting to impact*. Building on Instiglio's experience in the field, this guidebook is tailored to support non-profit implementers use RBF operate.
- 3) Instiglio and GPRBA. (2018). *A Guide for Effective Results-based Financing Strategies*. This guide provides RBF guidance for a funder perspective, outlining the potential of RBF to enhance returns on donor spending and providing detailed guidance and frameworks on how these enhanced returns can be achieved.
- 4) DFID. (2015). *Designing and Delivering Payment by Results Programmes: A DFID Smart Guide*. This note focuses on payment-by-results contracting – including outcome-based, output-based and hybrid approaches. This document gives a guidance overview to people considering, designing and implementing Payment by Results (PbR) interventions. The guidance provided is based on international and domestic lessons learnt.
- 5) Castellnou, M., Jammes, D., and Sienrukos, H. (2021). *Setting Up for Success: Best Practices for the Procurement and Contracting of Results-Based Financing Programs*. This paper provides practical guidance to practitioners, procurement and legal officers on the differences between RBF programs and typical service contracts or grants, for which adaptations of existing practices might be required.

In addition, the following sector specific guides are included to provide guidance on how to apply RBF in the energy, agriculture, and technical and vocational training sectors:

- 1) World Bank. (2015). *Results-Based Aid in the Energy Sector: An Analytical Guide (the Energy Sector Management Assistance Program)*. This guide covers RBF's application to energy based on two practical frameworks presented: (1) the energy results chain tool, which analyzes results and indicators in the energy sector, and (2) the multi-tier framework, which shows how to measure household access to electricity.
- 2) World Bank. (2021). *Impactful Design: AgResults' Pay-for-Results Prize Competition Toolkit*, AgResults. This guide includes a step-by-step process on how to use RBF in prize-based challenges specifically, to stimulate new markets for agricultural technologies.
- 3) HELVETAS and Employment Fund¹⁶. (2015). *Results-based Financing in Technical and Vocational Training A step-by-step implementation guide*. This guide aims at illustrating the interplay and modalities of RBF innovative approaches to ensure a sustainable and well-paid employment. It provides guidelines on how to apply the modalities in a few simple steps and highlights the major lessons learned from six years' experience of working with these approaches.

All the publications detailed above provide guidance on two key topics: i) preparation for RBF and ii) RBF strategy: guidance on design and implementation.

I. Preparation for RBF

The importance of assessing the suitability and feasibility of RBF is widely identified in the literature and is considered an important stage before any project starts. The objective of this phase is to diagnose whether it is feasible to implement an RBF project by conducting contextual analysis and evaluating the stakeholder commitment and engagement.

RBF value-add

To assess how and to what extent RBF can contribute to improved service delivery and results, a useful framework is presented, which analyzes what drives impact in RBF (GPRBA & Instiglio, 2018). The drivers of impact are the distinct channels for how RBF can drive greater development effectiveness and by analyzing them it helps to understand if RBF is the adequate instrument in a given program (GPRBA & Instiglio, 2018, p. 38; DFID, 2015, p. 5). These RBF drivers include: drawing attention to what matters, aligning incentives to the welfare of program beneficiaries, proving flexibility to maximize results, and enhancing the accountability of the incentivized agent to beneficiaries.

¹⁶ The Employment Fund applies innovative approaches to ensure the sustainable and gainful employment of its primary stakeholders.



Similarly, the World Bank's (2021) AgResults toolkit has the following four-part evaluation framework to decide whether to proceed with the design of a prize competition and understand if the project adds value to the organization:

- **Mission Impact:** Does the problem align with mission goals and impact expectations?
- **Prize Design:** Is the prize competition design feasible and measurable?
- **Enabling Environment:** Are the context and enabling conditions neutral-to-supportive?
- **Market Failure:** Is there a market breakdown that a prize competition can fix?

Contextual analysis

All guides highlight the importance of having an in-depth feasibility analysis to assess the viability of using RBF in a given country with particular socio-economic features. Since all the RBF design features are subject to the contextual conditions that can constrain RBF's effectiveness, the guides recommend performing a **contextual analysis** analyzing factors such as (Instiglio & World Vision, 2017; Instiglio & GPRBA, 2018; DFID, 2015):

1. **Technical factors**, which are those related to the availability of suitable interventions, results, and data required to design an RBF project,
2. **Institutional and legal factors** related to the capacity of the stakeholders involved, including the ability to pre-finance and regulations on disbursement, and
3. **Political factors** related to the interests and motivations of the stakeholders.

The World Bank (2021) recommends using a systems lens¹⁷ to identify how **market barriers** impacts value chains to pinpoint where to intervene. Additionally, GPRBA and Instiglio (2018) presents tools to identify these barriers which include a systemic market analysis such as root cause analysis or barrier analysis, or guiding questions such as:

- What is the social issue and what are the results funders seek to achieve?
- Are relevant services being used? If not, why is this the case?
- If relevant services available but not delivering results, why is this the case?
- Which limitations are the most severe and need to be addressed to improve results?

Stakeholder engagement

All guides highlight how **engagement and commitment between RBF stakeholders** is key to facilitate the whole process of designing an RBF model. For instance, the GO Lab (2017) recommends that a steering group is constituted at this early stage, so stakeholders are engaged in key decisions like the focus and ambit of the project. Other recommendations include starting the process with a detailed understanding of the relevant stakeholders; building alignment among stakeholders by creating shared objectives; and effectively planning the sequencing and engagement approach for each stakeholder (GPRBA & Instiglio, 2018).

2. RBF Strategy

Drawing on the contextual analysis, the RBF strategy phase entails making decisions on how RBF should be used in the given context, including the instrument selection and the RBF design. The section provides guidance on the choice of RBF instrument, key design choices, including what outcomes to pay for, how much, and how the achieved results will be verified, and lastly, guidance on issues related to strengthening and scaling the RBF strategy over time.

Instruments

Most guides start with instrument selection as the first step in designing an RBF strategy. The guides provide a short description of instruments that are commonly used. An easy-to-understand framework to decide which instrument to use is outlined in by GPRBA and Instiglio (2018, pp. 55), in which the instruments are organized in terms of the first-order question: 'who is the incentivized agent?'. This framework allows practitioners to have a clearer understanding of not only the different types of RBF instruments, but when should each one be used. Another useful framework included in this guide is the maturity model which provides a complementary categorization of instruments based on the

¹⁷Focusing on nodes along the value chain where target populations experience challenges, such as value loss, or market exclusion.



program's maturity in terms of the level of evidence available about the intervention's impact and implications for RBF's potential to drive enhanced effectiveness.

Design

All guides provide advice on defining key RBF design features, such as how much funding should be tied to results, which results should be paid for, how much should be paid for each result, how they should be measured and who should deliver them. We summarize common advice on these topics below.

Portion of funds tied to results

The portion tied to results sets the level of funding tied to results versus the level attached to inputs and activities. GPRBA and Instiglio (2018) highlights that the practitioners deciding the optimal portion of funding tied to results, should aim to balance the need to provide the incentivized agent with strong incentives and flexibility against the need to limit risk by accounting for the agent's capacity to absorb risks and the ability to get upfront financing. Likewise, the DFID (2015), presents some considerations to guide the portion tied to results, noting:

- If it is a program with outcomes-level metrics and targets, it is highly unlikely that the incentivized agent would bear 100% of the outcome risk. In addition, it is important to consider the incentivized agent's control in achieving outcome results. It may be more realistic to consider a lower proportion of funding tied to results to decrease non-performance risk.

Payment metrics

Payment metrics (i.e., outcomes and outputs to be measured and verified as the basis of payment) are the foundation of every RBF agreement, as mentioned across all guides. It may be that an appropriate measure does not exist, in which case RBF will not be feasible (DFID, 2015). Key considerations to assess when selecting payment metrics are provided (Instiglio & World Vision, 2017; Instiglio & GPRBA, 2018; Government Outcomes Lab, 2017; DFID, 2015):

- Having a well-defined theory of change
- Aligning the metrics with the desired impact and objective
- Selecting metrics that are within the manageable control of the implementers.
- Payment metrics that are objective and easy to measure
- Assess the costs of monitoring and data collection of the selected metrics

Pricing structure

The pricing structure determines the prices attached to different metrics and across populations. Most of the advice in this field highlighted in the publications is high-level and limited on aligning RBF pricing with existing approaches in order to set funding levels. A GPRBA and Instiglio (2018) presents five different pricing methods that could be used to determine how much should be paid for the achievement of the defined results:

1. **Cost-plus:** this approach is based on the known costs of delivering the results and adding a percentage.
2. **Comparative cost effectiveness:** with this method the price is derived by comparing the average cost for delivering the same results with similar interventions.
3. **Market determined:** when incentivizing providers, a competitive procurement process can be used to determine the most appropriate price.
4. **Cashable benefits:** this approach is based on an estimate of government future savings or revenue that could be made by achieving the defined results. It is most appropriate where government is the funder.
5. **Social value:** prices can also be set based on an estimate of the social value attaining the specified results.

In low data environments or where organizations are used to setting budgets based on costs, it is common for organizations to use cost driven approaches consistent with the first option above (African Development Bank, 2017). These approaches use the costs of delivering results as the base and adds on a predetermined percentage to the costs to account for the risk of not achieving results. Consistent with this, [GoLab's guidance](#) on pricing outcomes proposes to use cost estimates to set minimum prices, while using value estimates to estimate maximum prices.

From a process perspective, the GoLab (2017) identifies two broad methods for commissioners to determine the price to be paid for outcomes: 1) the market offers a price based on specified outcomes (aligned with the market determined method mentioned above), and 2) the commissioner determines the value of the outcomes by defining a "rate card"



and the market responds with an offer based on the number of outcomes that can be delivered for that value (potentially based on historical data or other analysis aligned with the points on comparative cost-effectiveness, cashable benefits or social value methodologies outlined above).

Verification method

Deciding on the approach for the verification and reporting on payment metrics is another key design feature when designing RBF mentioned across the guides. A GPRBA and Instiglio (2018), states that two key decisions are involved to determine the verification approach, both involve balancing rigor against cost and capacity demands.

1. Funders should decide on their preferred measurement method, defining who collects data, when and how. Data collection by independent third parties and larger samples, typically increase the confidence in the results but are more expensive.
2. Funders need to choose the verification method, deciding whether they want to pay for observed results or results attributable to the program.

World Bank (2021) provides considerations from best practices to guide the design of verification and evaluation of the results:

- Verifiers should be independent from the project manager
- A thorough verification plan may include multiple methods (e.g., ‘checks and balances’ approach or audits from external verifier)
- Verification should be robust but also financially and logistically feasible in the project context
- Verification should try to align with existing stakeholder activities or opportunities

In order to create a comprehensive yet pragmatic verification plan, designers must carefully consider the available data, payment indicators, and incentives for fraud or gaming by competitors. The DFID (2015) states that without robust and preferably independent verification there is a higher risk of dispute over payments, doubts over what has been achieved, fraud risks, and gaming¹⁸.

Selection of service providers

The selection of the service provider is an important design feature as it is the one responsible for implementing the intervention and achieving the predefined results. Castellnou, M., et al., (2021) give guidance on when a competitive process or a direct award is appropriate when selecting the service providers:

- A direct award (or a sole source) is appropriate when it is determined that there is only one implementer that can do the work and/or the implementer has unique skills or services.
- An open and competitive process is appropriate when there are multiple implementers able to provide the service. It is considered as the best way for commissioners to ensure that they receive value for money when there are multiple implementers.

Implementation

Strengthening and Scaling RBF

GPRBA and Instiglio (2018) recommends that while implementing an RBF strategy, funders plan for the RBF instrument’s evolution over time. This involves two dimensions: 1) how to strengthen RBF, and 2) how to scale it up. Additionally, practitioners should start RBF implementation with an explicit focus on testing and calibrating key design features can help improve the intervention and add to its sustainability. This can be achieved by piloting an RBF instrument at small scale and/or with a moderate level of incentives. An RBF pilot serves as a proof of concept for its future scale up and to understand which areas need improvement (Helvetas, 2015). These pilots can help answer questions such as:

1. Is the intervention likely to produce unanticipated (positive or negative) effects?
2. What benefits and challenges will the intervention produce for the incentivized agent?

¹⁸ A ‘gamable’ result is a result that can be manipulated by the recipient often in an unfair or unscrupulous manner and, as a consequence, the objective is not achieved.



3. Has the cost of delivering results been correctly estimated?

Results and performance management

Several of the guides highlight that practitioners should have appropriate performance management systems in place to help improve the delivery of results such as capacities in data collection, management, and analysis and the ability to devise and implement course corrective measures based on data-driven insights (GPRBA & Instiglio, 2018). Commissioners should expect and encourage the provider to adapt their way of working in response to the evidence coming from the operation (GO Lab, 2017).

4. Lessons from theory and practice of RBF

This section reviews key lessons on when and how to use RBF based on practical experiences drawing on publications including case studies and anecdotal evidence from literature reviews, interviews, surveys, and focus groups discussions. Based on this review, this section provides practitioners with the lessons learned from designing and implementing RBF, covering success factors, and challenges in the use of RBF in different sectors including health, education, workforce development, energy and environment.

This section is structured as follows: first it reviews lessons regarding the value-add of RBF and how RBF can drive impact. Secondly, it covers relevant lessons from the RBF project cycle from design to implementation. Last, we review lessons on the importance and potential of providing incentives within organizations to improve performance, including experiences beyond RBF.

I. RBF's value-add

The publications reviewed highlight the potential of RBF to drive change in how development interventions are delivered and achieve greater results. To achieve this change, publications emphasize the importance of assessing how RBF can add value before starting the design process and the challenges faced when designing innovative financing approaches compared to traditional input-based financing. For instance, Perakis and Savedoff (2015)¹⁹ highlight that RBF programs are more likely to be approved and be distinctive compared to traditional input-based approaches based on substantial upfront effort to demonstrate the value-add.

The value-add analysis rests on an understanding of RBF's drivers of impact commonly known as the channels through which RBF can drive greater effectiveness in developing spending. The benefits of using RBF are strongest when there is a value proposition based on activating certain drivers of impact which address specific delivery challenges. For instance, Perakis and Savedoff (2015) propose a set of four RBF drivers based on empirical observations. While the publications reviewed provide little evidence on how these theoretical channels drive impact, below we provide insights related to each of these drivers:

- 1. RBF schemes can make results visible for all actors by linking them to the payments of the program.** In traditional activity-based funding, service providers are incentivized to focus on implementing prescribed activities on-time and within budget, which may decrease their ability to concurrently pay sufficient attention to achieving the desired results. In RBF programs, by tying funding to results, the focus is shifted from activities to outcomes, contributing to greater results being achieved. For instance, the GAVI Immunization Services Support (ISS)²⁰ linked payments made to governments to defined outcome indicators, specifically for increases in routine immunization, measured as coverage rates of the diphtheria-tetanus-pertussis vaccine (DTP3), which drew the attention of government officials to the achievement of these outcomes through a change in management decisions, and focused the attention of implementers on the purpose of their activities (Center for Global Development, 2015).
- 2. RBF can facilitate alignment of actors and actions around results.** In complex interventions where stakeholders have diverse, and sometimes opposed objectives, incentives are often misaligned, creating confusion and posing challenges in achieving concrete results. This was the case in aspects of the Dominican Republic's

¹⁹ Does Results-Based Aid Change Anything? Pecuniary Interests, Attention, Accountability and Discretion in Four Case Studies, 2015.

²⁰ 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.



education system prior to the National Pact for Education Reform (2014),²¹ where RBF was able to address misalignments and haphazard divisions of labour by ensuring a clearer division of responsibilities that contributed to driving greater education results (World Bank Group, 2017).

3. **RBF can enhance accountability to beneficiaries for results they care about.** RBF's pre-defined metrics and their respective verification systems place incentivized agents' performance center stage and give funders and other stakeholders reliable information against which to hold a provider's management accountable. In the Kenya Electricity Expansion Project, at the time the contract was tendered, poor accountability was considered a key challenge within the utility. However, as the contract came into force, a committee with representatives of KPLC and MHI was created to oversee the execution of the contract and enforce accountability. Moreover, the management contract required submitting monthly, quarterly, and bi-annual reports regarding financial performance. These reports were then used by the division of Corporate Performance Monitoring and Strategic Planning to perform audits of KPLC's accounts (GPRBA, 2016).
4. **RBF can provide flexibility for incentivized agents to innovate in pursuit of more effective solutions.** The lack of flexibility in activity-based funding models can have important implications on the likelihood of attaining funders' ultimate goals. As part of the Kenya Electricity Expansion Project²², the Kenya Power and Lightening Company (KPLC) received funds to support grid expansion into slum areas. Tying funds to outputs, rather than activities was critical for KPLC to have sufficient flexibility to adapt and achieve agreed targets. As program implementation began, some slum residents "were reluctant to switch to legal connections due to issues of trust, payment barriers, and fear of reprisals from local cartels." To address this challenge, KPLC tailored its approach depending on whether the household was in an area with a strong cartel presence or not. Given the flexibility of the program, KPLC was able to make changes in the implementation and strengthen their outreach and increase connections in slum areas (GPRBA, 2016).

2. Lessons on assessing the conditions for RBF

Assessing the suitability and feasibility of RBF is widely recognized as an important phase before project design. The objective of this phase is to diagnose whether it is feasible to implement an RBF project in a given context and relevant publications highlight the importance of committing sufficient time and resources to this phase (DFID, 2017). Relevant lessons and challenges regarding this preparation phase are presented below:

1. **Performing a contextual analysis is essential given that the design features and suitability of an RBF instrument are subject to the contextual conditions which can constrain RBF's effectiveness.** Assessing the political will and performing a participatory stakeholder analysis helps ensure that the diagnosis considers local knowledge and builds ownership in the process. RBF schemes are more likely to succeed when there is strong leadership and political commitment. For example, the Nepal Performance Based Maintenance Contract in Nepal, under the Road Maintenance and Development Program (RMDP)²³ supported by the International Development Association, allowed bidders to compete for a fixed monthly lump sum fee per km of road for the maintenance of the road (remuneration paid to the contractor covered all physical and nonphysical maintenance services provided by the contractor). The success of this program was impinged by weak stakeholder commitment, a constraint which was not identified until implementation issues emerged as adequate political analysis had not been undertaken at the outset. This lack of support resulted in political interference in the project, with money being taken from the Road Board and allocated to new roads rather than improved maintenance, limiting the project's success (DFID, 2003).

²¹ The reform aimed to recruit and train primary and secondary school teachers; assess student learning; evaluated early childhood development services and help decentralize public school management.

²² This program was financed by the World Bank in 2012, where GPRBA provided the Kenya Power and Lightening Company (KPLC) a USD 5.15 million grant to support grid expansion into slum areas. Payments were disbursed upon verified working connections with pre-paid meters.

²³ The first pilot-based Performance Based Maintenance Contract, focused on strengthening road maintenance in Nepal, started in June 2003 under the Road Maintenance and Development Program (RMDP) supported by IDA (International Development Association). Five additional contracts under RNDP supported by Asian Development Bank (ADB) were implemented in Nepal as Performance Based Maintenance Contract from 2005-2010.



2. **Availability of high-quality data is essential for a well-designed RBF model.** Data gaps like periodicity, quality and timeliness of administrative data poses a challenge in the use of RBF as high-quality data is key in designing targets and a pricing approach. From a review of seven RBF programs done by DFID (2010), lessons highlight that there are important gaps in the data available of many countries, especially for fragile states. The gaps relate to the periodicity of household census and survey data and the quality and timeliness of administrative data, in particular. RBF contracts encourage and support developing countries to develop the capacity they need to produce the data required for M&E. Hence, there is a need for further support if these RBF schemes are to be effectively monitored.
3. **Assessing the capabilities of implementers is essential to evaluate their ability to deliver results and if needed, invest in their capacity building.** Service providers or other incentivized agents with weak capacity, weak performance management systems and limited evidence on the effectiveness of their interventions pose a high program risk as they may be less likely to achieve agreed outcomes. Therefore, ensuring adequate organizational structures and institutional capacity is a key step before starting a project. In the Helvetas Skills and Knowledge for Youth (SKY) – Vocational Training for Young People program in Ethiopia²⁴ program, technical support was needed to ensure capacity constrained providers could benefit from RBF. Trainers mainly received “soft” inputs such as capacity building to offset the risks they needed to bear. Soft inputs also included sharing or strengthening partnerships with government bodies and employers (NORRAG, 2021).
4. **Using existing institutions or partnerships can help enhance a project’s implementation quality and sustainability by saving time and resources when designing an RBF model.** Pre-existing conditions can enhance the achievement of results, particularly when a short RBF duration limits the opportunities for capacity development from the RBF mechanism itself. Based on the research conducted by the Center for Global Development (2015) on Scaling Up Performance-based Transfers for Reduced Tropical Deforestation, joining existing partnerships was greatly beneficial. Combining existing bilateral/trilateral agreements may be preferable because they may have been partially “liberated” from conventional aid practices and are more easily able to channel funding. Nevertheless, it is important to ensure that an existing institution’s mission and procedures do not constrain innovative or systemic uses of funds (Center for Global Development, 2014).

3. Lessons on the RBF Design

The effectiveness of an RBF program is critically dependent on design parameters and decisions, including what outcomes to pay for, how much, and how the achieved results will be verified. Effective RBF design must align the incentives of incentivized agents with the desired impact. Where incentives are not aligned with impact, incentivized agents may face perverse incentives to ‘game the system’ by maximizing their payments rather than employment outcomes. For instance, this can manifest in ‘cream-skimming’, where incentivized agents focus on beneficiaries who are most likely to achieve the defined results even without support, representing a deadweight loss to the program, while ignoring (‘parking’) those most in need of services.

We have structured design lessons into three categories relating to 1) incentive sizes; 2) Setting payment metrics, and 3) Determining the pricing structure. These lessons are explored below:

- 1) **Promoting a healthy level of competition between implementers will motivate them to achieve results, particularly in contexts where multiple implementers are being incentivized for the same results.** Where possible and if multiple service providers have the necessary capacity and experience for RBF, working with multiple implementers can create a space for them to share experiences and ignite a positive competition between them. Lessons from the SKY program (PLoS ONE, 2017) highlight how a degree of competition among training providers across a range of public and private training providers and an existing demand for labor in the market was key in the implementation process. If a labor market is completely saturated, the implementation of RBF may also be difficult. SKY was more easily implemented in urban areas with stronger

²⁴ The project applies results-based financing, and more precisely performance-based contracts, where training providers are paid for predefined results achieved (skills improvement and employment of graduates) following independent evaluation. Helvetas acts both as funder and implementer of the project. It is implemented by the Helvetas project office in Ethiopia and owned and financed by Helvetas Switzerland/headquarters.



economic activities and better training providers available than in rural areas with limited employment opportunities.

2) **Setting payment metrics that are measurable and verifiable at a reasonable cost, ambitious but realistic, and in the manageable control of implementers**, is a widely identified challenge. Well-designed payment metrics are critical as they define what success means and the incentives provided. Key challenges and related lessons on metric selection from the literature review include:

- **Indicators selected must be easy to measure and objective at a reasonable cost and time frames.** This will contribute to ensuring that all stakeholders are clear on what the desired impact is and drive alignment to focus on those results. Lessons from a review by the World Bank of 74 RBF programs highlight the need to have an accepted definition of the desired outcome and corresponding tools to measure it. For example, from the RBF programs reviewed in this study, the number of Solar Home Systems installed was the RBF output, which contributed to “delivering” a measurable volume of CO2 emission reductions, defined as the RBF outcome (World Bank, 2017). Using indicators that are easy to measure at reasonable costs supports cost-effective performance management systems and enhances and accountability of an RBF program.
- **Metrics (along with other design features like target populations) need to be tightly defined to focus on impact and avoid perverse incentives.** The literature highlights that poor metric selection can cause of RBF to create incentives that risk undermining the program goals by shifting the attention to outcomes with payments attached to them. These so-called perverse incentives might cause undesirable effects, including (1) ignoring unrewarded activities, (2) inducing unnecessary demand for remunerated services, (3) hampering sustainability, (4) deteriorating service quality, (5) gaming and fraud, (6) so-called cherry-picking or cream-skimming, or (7) reducing intrinsic motivation (Grittner, 2013). These issues were identified in the Tennessee Department of Children’s Services (DCS) Performance Based Contracting²⁵, where there was a concern that by paying quick and effective placement of children in permanent homes, providers could be encourage to only serve the easiest to place children, and not those most in need. To resolve this risk, DCS requires providers to complete exception reports each time they refuse to take a child, enabling analysis to identify trends in the reasons for refusal (Beeck Center, 2019).
- **To create effective incentives, results should be within the manageable control of the provider and not be overly influenced by exogenous factors.** The provider’s existing capabilities, financial resources, and decision-making autonomy should be considered. For example, in the Indonesia Power Distribution Development Program-for-Results, residential energy sales were selected as a payment metric based on the premise that increased household connections and improved quality of service (also incentivized through the RBF) would translate into higher electricity consumption. However, energy sales are impacted by numerous exogeneous factors, such as economic growth (or lack thereof) and individual households’ income and employment status. In Indonesia, despite considerable progress in expanding access, reducing technical losses, and improving supply reliability (all of which contribute to more energy available for sale), energy sales remained significantly below expectations and the restructuring paper published in 2019 noted that “this phenomenon is widespread in all Islands and not under the control of the Project” (World Bank, 2019).

3) **Determining an adequate pricing structure is a key RBF design feature as it determines by how much outcomes will be paid.** RBF builds on the existing incentive environment by adding financial incentives to further align interests and reward better results. To achieve this, RBF requires a price per outcome, while ensuring that this price is sufficient to affect the incentivized agent’s behavior. There are different methodologies that can be used to price outcomes, which depend on the data environment, the confidence in the pricing and availability of benchmarks, and the regulatory framework of the organization determining the price per outcome. Two challenges when determining the pricing structure are highlighted in the publications: 1) how to choose a pricing

²⁵ In the wake of a child welfare crisis in 2006, Tennessee adopted an outcomes-based agreement model to reduce the amount of time taken to place children in permanent homes. Providers that improve on baseline performance receive a share of the state’s savings, and those that perform below the baseline reimburse the state for cost overages.



method given a specific context and, 2) how to ensure the pricing structure incentivizes the desired results and impact. Lessons from these two challenges are presented below, respectively:

- **The data environment influences the decision of what pricing method to use.** Stakeholders interested in promoting RBF in low data environments may face challenges in determining pricing structures based on reliable information. For “Empleando Futuro” in Colombia, a significant amount of data was collected to determine the adequate price per outcome, increasing the availability of quality data for the sector. In this case, the pricing method was cost-driven, where the unit price per result achieved was based on the costs of the program. Additionally, both the budget assigned and the unit price per metric were agreed by all actors involved in the design process to guarantee the pricing structure created incentives and met the objectives of the program (Instiglio, 2018).
- **The pricing structure influences the behavior of incentivized agents and should ensure all prioritized outcomes are sufficiently incentivized.** Incentivized agents are not only incentivized by the price per outcome or payment metrics, but also by how different payment metrics interact. There are different pricing structure design choices that can be used to drive different behaviors. For example, differential pricing can be used when an intervention wants to target different groups and create an incentive to target the harder to reach. The SKY program used a differential pricing system where the total payment for training and gainful employment differed per trade, and the incentives were based on the target groups. Some of the lessons drawn from this program include how the differential pricing system benefited both the providers and the beneficiaries. Specifically, the service providers were incentivized to serve all populations, by setting differential prices equal to differential costs of generating results for different populations. Additionally, this pricing method reduced the risks the training providers faced since they were paid not only for training delivery but also for job placement services (NORRAG, 2021).

4. Lessons on the implementation of RBF programs

Key lessons from the literature in relation to the RBF implementation phase can be categorized into three key topics: 1) adaptive management practices, 2) monitoring and evaluation systems in place, and 3) incentivizing increased performance within organizations. The learnings that stood out from the literature in these three categories are:

- 1) **Adaptive management approaches can improve the results achieved.** The publications reviewed highlight the importance of using adaptive management to enable course correction to enhance results. The Salud Mesoamericana Initiative 2015 (SMI) was a five-year initiative that aimed to reduce equity gaps in maternal and child health across Central America²⁶. SMI provided measurement of the results and enabled SMI's partners to identify implementation bottlenecks and devise solutions and corrective actions on the spot, as well as terminate a country's participation if it did not meet targets. One of these corrective measures was the knowledge that targets had not been met at the end of the first operation in Chiapas, which opened the door for an action plan and allowed a correction of the course that the initiative was taking. Hence, constant monitoring of implementation and progress fostered an experience-based learning environment across countries and localities. Freely available data from stakeholders, lessons-sharing meetings between participating countries, and knowledge-sharing as a core component of the Theory of Change of the project made this possible (PLoS ONE, 2017).
- 2) **A well-managed evaluation and verification system is need to rigorously assess results as the basis of payment.** Robust results verification systems give outcome funders confidence that the results that are being paid for are accurate. The SMI demonstrates that credible measurement is essential for programs that pay for results. Independent verification, both in terms of the institutional independence of the organization responsible for the surveys (IHME) and the independence of the information collected, was an important part of the program. It was important not only because it assured 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 (Center for Global Development, 2015). In the absence of strong existing systems, independent verification can come in the

²⁶ SMI2015 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 with a portion of funding to governments contingent on progress on health-related development goals



form of third-party verification, through contracting with private actors. This can be a good short-term solution but will not resolve the capacity constraints in the long-term (The World Bank Group, 2017).

- 3) **Incentives can improve performance within organizations, contributing to greater results.** The right incentive structure can enhance an institution's performance through 1) behavioral change within the organization, and 2) refining the organization's operations. The literature shows that with the right incentives, individuals and teams within organizations are encouraged to increase their effort and work more efficiently. In addition, they are more focused on achieving the institution's goals because incentives can enlighten additional interests besides their set responsibilities (Overseas Development Institute, 2016., DFID, 2003).²⁷ This can be achieved by, for example, promoting incentives in Human Resources for management and staff. Lessons from a pilot to review organizational incentives in the World Bank's Nigeria country team (Overseas Development Institute, 2016) show that focusing on Human Resource incentives for management and staff is key. The World Bank applied new competencies for staff and management, created space and incentives for management to lead by example and actively rewarded individuals and teams for behaviors that were likely to lead to greater organizational effectiveness such as problem solving, adaptation, and learning.

²⁷ The publications reviewed (Overseas Development Institute, 2016., DFID, 2003) focus on the benefits of providing incentives to generate performance improvements within organizations in general, not restricted to RBF approaches.