



A practitioner's guide to  
**RESULTS-BASED  
FINANCING**

---

**Getting to impact**

A practitioner's guide to  
**RESULTS-BASED  
FINANCING**

---

**Getting to impact**

---

Authored by Instiglio.

Copyright © 2017 Instiglio and World Vision.

Some rights reserved. The text in this document may be reproduced free of charge provided it is reproduced accurately and not used in a misleading context or for profit purposes. The material must be acknowledged as Instiglio and World Vision copyright and the title of the document specified.



**Instiglio, Inc.**  
1875 Connecticut Ave NW  
10<sup>th</sup> Floor  
Washington, DC  
20009  
USA

[www.instiglio.org](http://www.instiglio.org)

Photo credits:  
Instiglio  
World Vision

Designed by:  
SocialThink  
[www.socialthink.co](http://www.socialthink.co)



**World Vision Canada**  
1 World Drive  
Mississauga, Ontario  
L5T 2Y4  
Canada

[www.worldvision.ca](http://www.worldvision.ca)

# ACKNOWLEDGEMENTS

We would like to thank numerous people and organizations for their contributions to this guidebook. First of all, we offer our profound gratitude to World Vision for sponsoring the publication and providing invaluable feedback during the whole process.

We are also grateful to all of the individuals who participated in the interviews conducted during the preparation phase, and whose valuable input guided the content and allowed us to refine the knowledge gap we seek to address through this publication. Each of these conversations added to the final product.

Particular appreciation also goes to all the organizations who agreed to make public the references to their respective programs and experiences in the field of international development. These contributions have allowed us to enrich the theoretical concepts presented with relevant and enlightening examples, and thus we would like to extend our gratitude to Educate Girls, Escalera, Fundación Casa del Cerro, Fundación San Carlos de Maipo, Village Enterprise, and World Vision.

Finally, our experience and work would not have been possible without the multiple actors – governments, donor agencies, foundations, implementers, opinion leaders, and investors – who believed in the promise of Results-Based Financing and embarked on a tireless pursuit of greater social outcomes.

# FOREWORD

## BY WORLD VISION

At World Vision, we continually strive to improve development outcomes. Thus, we have begun a process of experimenting with new tools such as Results-Based Financing (RBF); and in this process, we identified the need to learn more, and get targeted guidance, about RBF. The uniqueness of RBF is not only the focus on desired results, but also the learning and organizational changes which come with this commitment to the delivery of measurable results.

When implementing RBF, often the issue arises of how to best implement an adaptive change process. For many organizations, RBF will challenge assumptions and require them to embrace change. Although there are core parts of an organization which should be preserved, ultimately the resulting DNA of an outcomes-focused organization should allow for real-time project adaptation and iteration.

To leverage this acquired freedom for creativity, and innovation as to what combination of activities and outputs will be most effective in achieving project objectives in a given context, World Vision has found value in fostering a culture of experimentation and evidence-based decision making – beyond our organization's experience or the opinion of experts. Moreover, this experience has required us to build performance management systems and the accompanying internal capacity to understand what adjustments can generate greater impact.

As funders are becoming increasingly interested in RBF, they should not only be keen on financing the development of these performance management systems, but also on encouraging and providing the space for innovation through a thoughtful RBF design. World Vision's experience has shown that the careful selection of payment metrics and the presence of investors' who are willing to take on an additional risk can lead to effective experimentation and improved outcomes.

As development practitioners, we all want to generate greater outcomes, and this RBF Guidebook is a contribution toward sharing World Vision's experience and ongoing journey toward that end. Our hope is that you will find this guidebook useful and that it will positively influence how your organization views the development process.

**Carlos Piedrasanta**

*Global Lead Grants Acquisition & Management  
World Vision International*

**David O'Leary**

*Director of Impact Investing  
World Vision Canada*

## BY INSTIGLIO

At Instiglio, we have worked with a diverse range of innovators and practitioners over the last 5 years to improve the performance of development programs. Our quest for greater impact has led us to support the development and implementation of Results-Based Financing (RBF) models that focus programs on tangible results. We have observed first-hand how RBF can rapidly accelerate development outcomes by fostering a problem-solving mindset, enabling implementers to develop and iterate more relevant and effective solutions with a greater focus on quality.

Driven by this potential, we dedicated our first years to carving out a more prominent space for RBF in development discourse and practice. We supported the demonstration of new models, such as the Educate Girls Development Impact Bond in India, a Social Impact Bond to improve labor outcomes of vulnerable youth in Colombia, and the first outcomes fund and Development Impact Bond to finance livelihood outcomes in Sub-Saharan Africa. We helped develop a stronger outcomes-orientation in existing RBF models, supporting the design of government-led RBF systems in Morocco, Peru, and Egypt. We also worked together with implementers like World Vision to help them upgrade their performance management systems and prepare them for RBF contracts.

Through this work, we have witnessed the challenges of getting RBF right; we have had to evolve, revise, and even abandon certain design features. We have encountered situations where RBF made things worse, which has underscored the lesson that RBF is simply one tool, and not a silver bullet. We have appreciated the leap that RBF requires from implementers and governments, in terms of developing an entire set of new capabilities. In the words of one of our long-term partners, “RBF sounds so simple. Yet, when you start managing to outcomes, the upgrade that you need to succeed feels like a reinvention.”

Today, as the development sector is poised to rapidly expand the use of RBF, guiding the thoughtful application of these models has become central to our work. Through this guidebook, we are sharing the expertise, practices, design frameworks, and the other insights and lessons we have acquired and developed thus far. We hope it will impart a deeper understanding of where to apply RBF and how to design high-quality and successful RBF projects, by providing concrete technical guidance to practitioners. The guidebook lays a foundation that we hope will ultimately contribute to turning RBF from an art, at times a hype, to a well-understood service delivery mechanism.

**Avnish Gungadurdoss**

*Managing Partner and co-Founder  
Instiglio*





# CONTENTS

<b>INTRODUCTION TO THE GUIDEBOOK</b>	<b>12</b>
Why the interest in Results-Based Financing?	12
About this Guidebook	14
Who is this Guidebook for?	14
How to use this Guidebook	15
<b>UNDERSTANDING RBF AND ITS POTENTIAL BENEFITS</b>	<b>18</b>
Focusing programs on the desired results, instead of the activities delivered	19
The three key steps of RBF	20
The potential value-add of RBF for implementers	22
RBF Instruments	27
Key takeaways	34
<b>ASSESSING THE IMPLICATIONS OF THE RBF DESIGN</b>	<b>38</b>
Introduction to the RBF design components and process	38
Guiding questions for the assessment	41
Assess the boundaries of the project: financial size, duration, and target population	43
Assess the payment metrics	44
Assess the evaluation and measurement method	51
Assess the payment structure	61
Assess the RBF contract features	73
Prepare for negotiations	76
Key takeaways	78
<b>BUILDING THE CAPACITY TO LEVERAGE THE BENEFITS OF RBF</b>	<b>84</b>
Why it matters	84
Stakeholder management skills	84
Capacity for performance management	87
Key takeaways	97
<b>REFERENCES</b>	<b>99</b>
<b>APPENDIX</b>	
Appendix 1: Results-Based Financing (RBF) Cheat Sheet	104
Appendix 2: Results-Based Financing Instruments	106
Appendix 3: Overview of evaluation methods	108
Appendix 4: RBF design negotiations checklist	112
Appendix 5: IT systems for performance management	113





# ABBREVIATIONS

---

<b>CIFF</b>	Children's Investment Fund Foundation
<b>COD</b>	Cash on Delivery
<b>DFID</b>	Department for International Development
<b>DIB</b>	Development Impact Bond
<b>EG</b>	Educate Girls
<b>FY</b>	Fiscal Year
<b>GPOBA</b>	Global Partnership for Output Based Aid
<b>HIV/AIDS</b>	Human Immunodeficiency Virus infection and Acquired Immune Deficiency Syndrome
<b>HSDP</b>	Health Sector Development Program
<b>IDRB</b>	International Bank for Reconstruction and Development
<b>IT</b>	Information Technology
<b>M&amp;E</b>	Monitoring and Evaluation
<b>MDE</b>	Minimum Detectable Effect
<b>MDG</b>	Millennium Development Goals
<b>NGO</b>	Non-Profit Organization
<b>NORAD</b>	Norwegian Agency for Development Cooperation
<b>PBC</b>	Performance-Based Contract
<b>PBG</b>	Performance-Based Grant
<b>PBL</b>	Performance-Based Loan
<b>PDBD</b>	Performance Debt Buy-Down
<b>PMS</b>	Performance Management System
<b>RBA</b>	Results-Based Aid
<b>RBF</b>	Results-Based Financing
<b>RCT</b>	Randomized Controlled Trial
<b>SDG</b>	Sustainable Development Goals
<b>SIB</b>	Social Impact Bond
<b>UK</b>	United Kingdom
<b>US</b>	United States
<b>USAID</b>	United States Agency for International Development
<b>VE</b>	Village Enterprise



# 1

## INTRODUCTION TO THE GUIDEBOOK



# INTRODUCTION TO THE GUIDEBOOK

## WHY THE INTEREST IN RESULTS-BASED FINANCING?

The failings of business-as-usual development practices when tackling complex social problems are widely recognized. A litany of studies has highlighted problems, such as the rigidities of traditional programming, too little focus on results, disincentives for using evidence, and an overreliance on cookie-cutter responses rather than searching for context-specific solutions.<sup>1</sup> Given these constraints and the growing pressure to achieve greater cost-effectiveness with social spending, the development sector has begun shifting away from traditional, activity-based funding towards results-based approaches.

Results-Based Financing (RBF) is one approach to driving greater impact from social spending.<sup>2,3</sup>

By tying the funding of social programs to results rather than to activities and inputs, well-designed RBF introduces performance incentives. Further, it provides implementers with greater flexibility to adjust their programs, empowering them to innovate, learn, and adapt their programs in pursuit of impact.

By embracing RBF, implementers are also better placed to respond to funders' growing focus on measurable outcomes. Beyond demanding more rigorous impact evaluations, funders are increasingly demanding options which allow them to directly pay for outcomes. Responding to these demands is critical for implementers, especially in the face of an increasingly competitive funding landscape.

“RBF has become a valuable tool for World Vision in not only measuring outcomes, but in managing outcomes much more effectively. In so doing, World Vision has the potential of achieving much greater impact.”

Carlos Piedrasanta  
World Vision Global Leader  
Grant Acquisition and Management

<sup>1</sup> Andrews, Pritchett, & Woolcock (2017); Bain (2016); Ramalingam (2013); Rondinelli (1993); Webster (2016).

<sup>2</sup> Birdsall & Savedoff (2010); Kenny (2017); Perakis & Savedoff (2015); Ramalingam (2013).

<sup>3</sup> RBF has also been endorsed as a measure for aid effectiveness by Angela Merkel, Jim Yong Kim and Ban Ki Moon, see: “New Way to Finance Health in World’s Less Developed Nations” (2013).

Furthermore, since funders only pay if results are achieved, RBF enables implementers to offer greater value for money. This may help implementers engage funders and secure support for innovative programs - even if funders are relatively risk averse and if programs are based on fairly nascent evidence.

The burgeoning significance of RBF is reflected in the recent growth of RBF spending. We estimate that in the decade to 2016, over US\$ 20 billion of development spending was funneled through RBF, up from just a few billion in the preceding decade.<sup>4</sup> Illustrations of this shift include:

- The World Bank's Program-for-Results has been steadily increasing in use since its creation in 2012, totaling almost US\$ 15 billion worth of RBF funding which has supported over US\$ 70 billion of government programs.<sup>5</sup>
- The United Kingdom's Girls' Education Challenge Fund has used RBF in 25 of its 37 education programs implemented since 2012, worth £300 million.<sup>6</sup>
- In 2016, the Department for International Development (DFID) announced that up to 30% of its core funding to United Nations humanitarian and development agencies will be disbursed to agencies that meet pre-agreed performance targets.<sup>7</sup>

With these trends set to continue, learning to use RBF effectively is an opportunity and a necessity for implementers committed to operating on the cutting edge of development effectiveness.

<sup>4</sup> The Instiglio RBF Database, April 2017. This database compiles information on RBF programs in low- and middle-income countries, utilizing sources like the World Bank, DFID, USAID, Asian Development Bank, NORAD.

<sup>5</sup> World Bank (2017).

<sup>6</sup> DFID (2014); DFID (2017).

<sup>7</sup> DFID (2016).

## ABOUT THIS GUIDEBOOK

Like any tool, RBF's effectiveness depends on how it is used. Designed and used poorly, RBF can distort implementers' priorities and end up being ineffective.<sup>8</sup> The RBF Practitioner Guidebook offers practical guidance to non-profit implementers who want to embrace RBF as a way to improve development outcomes. It responds to a scarcity of resources available for implementers, especially regarding the detailed technical guidance necessary to ensure that RBF designs align with their mission, capacities, and context.

There is no fixed model or formula to determine whether and how an implementer should enter into an RBF agreement. As such, this guidebook is not a compliance manual, but rather provides information and frameworks that allow implementers to make informed decisions. The objectives of the guidebook are threefold:

1. Introduce implementers to RBF and help them analyze the potential value added for their organization and programs.
2. Equip implementers with practical frameworks to assess the implications of the design of an RBF agreement on the incentives they face and the payment risk they bear.
3. Outline the capabilities needed to successfully implement RBF and leverage the benefits it offers.

Importantly, the guidebook does not equip an implementer with the necessary skills and expertise to design RBF agreements. The process of designing an RBF agreement is very technical, context-specific, multi-dimensional, and iterative; it requires experience and expertise that is beyond the scope of the guidebook. However, it does prepare organizations to ask the right questions and to better assess the key risks and benefits of an RBF agreement, ultimately serving as a point of departure to successfully negotiate the RBF elements being proposed.

## WHO IS THIS GUIDEBOOK FOR?

The guidebook is intended for implementers of development programs that are looking to make their operations increasingly results-oriented, and who are interested in RBF as a tool to potentially increase their outcomes. Each chapter is intended for a different audience within these organizations, as noted in the descriptions of the chapters below.

While the guidebook responds to implementers' need for guidance and capacity building in this new practice, we hope it will be relevant to funders and investors alike - as well as useful in academic settings as a starting point for future development practitioners.

<sup>8</sup> For example, randomized evaluations measuring the impact of RBF compared to traditional financing for maternal and child health services in Zambia and Afghanistan found no impact of RBF. See World Bank (2016a) and Engineer et al. (2016).

## HOW TO USE THIS GUIDEBOOK

The guidebook is divided into the following three chapters:

2

### UNDERSTANDING RBF AND ITS POTENTIAL BENEFITS

**This chapter introduces RBF and explains how it can help implementers enhance the impact of their programs and improve their funding prospects. It is geared towards senior leadership teams interested in learning more about RBF and determining their strategy with regard to RBF.**

- Learn about the key steps involved in an RBF agreement.
- Understand how RBF helps improve the impact of your program and expand funding prospects.
- See the range of contracting structures for RBF.

3

### ASSESSING THE IMPLICATIONS OF THE RBF DESIGN

**This chapter equips implementers to assess the alignment of the RBF design with their mission and the financial risk they bear. It is intended to be used by business development staff and program directors responsible for determining whether the organization should pursue a funding opportunity.**

- Understand the components of the RBF design and the main phases of work involved.
- Explore each design component in turn and assess the implications on the incentives your organization faces and the financial risks transferred.
- Learn about important contract features that protect your organization against risk outside your manageable control and protocols for addressing risks and disputes.
- See recommendations on how to negotiate favorable design components with outcome payers.

4

### BUILDING THE CAPACITY TO LEVERAGE THE BENEFITS OF RBF

**This chapter explains important capabilities to successfully implement RBF. It is intended to help program directors and M&E leadership consider the relationships with external stakeholders thoughtfully, and guide plans to build performance management capabilities.**

- Understand the importance of building trust-based relationships with funders and translating RBF incentives to internal teams and implementing partners.
- Explore how performance management can support your organization in the necessary shift in focus from activities to results. Learn about organizational processes and culture that empowers staff to actively identify strategies that maximize impact and minimize payment risks.

The guidebook has been designed to maximize its utility to readers—providing actionable lists wherever possible and allowing users to navigate to the information they need quickly. Each chapter begins with a short overview of the content and concludes with a box summarizing key takeaways.

Deep dives into specific topics and case studies of real-world RBF agreements are included in boxes, and can easily be skipped when not relevant.

The appendix of the guidebook includes a 2-page Cheat Sheet on RBF concepts, which can be used for reference by senior leadership or printed for staff trainings. It also includes an RBF design check-list that summarizes the key design considerations provided in Chapter 3 and which can be used before entering negotiations.





# 2

## UNDERSTANDING RBF AND ITS POTENTIAL BENEFITS

*This chapter introduces Results-Based Financing (RBF), and how it can add value for implementers by: 1) enhancing the impact of their programs, and 2) improving their funding prospects.*

*Following a brief introduction to the concept, the chapter begins by presenting RBF as a three-step process. It introduces the three drivers of impact, critical for fulfilling the first value proposition: improving impact. Next, it explains the second value proposition, expanding funding prospects for implementers, which may come from two sources: improving program competitiveness and facilitating the engagement of risk-averse funders. The chapter concludes by introducing common RBF instruments, providing context on when each instrument is most appropriate.*

# UNDERSTANDING RBF AND ITS POTENTIAL BENEFITS

## Box 1. Key concepts of the chapter

1. **Results-Based Financing (RBF):** A program financing arrangement in which payments are contingent upon the achievement of predefined results, which are usually verified by an independent evaluator. In the field of international development, this can mean, for example, that a government (an outcome payer) makes payments to a nonprofit organization (an implementer) only after it delivers increased child literacy rates (a predefined result).
2. **RBF instrument:** A specific RBF modality. The nomenclature, which is not particularly well structured around any specific dimension of RBF, includes Performance-Based Contracts, Impact Bonds, and Prize-Based Challenges.
3. **RBF design:** The specific parameters of the RBF instrument, including the selection of the results to be paid for, how results are verified, and the total payment made for the achievement of results.
4. **Funder:** An organization that provides money for development programs. Examples include philanthropic donors, multilateral and bilateral donor organizations, and governments.
5. **Outcome Payer:** A funder who makes payments conditional on the achievement of predefined results within an RBF instrument. Note that outcome payers are a subcategory of funders.
6. **Investor:** An organization that provides upfront working capital to implementers within an RBF instrument. Upon the achievement of verified results, investors expect to receive this capital back – often with a return. Not all RBF instruments utilize an investor. Examples of investors include impact investors, foundations, private organizations, and nonprofit organizations.
7. **Implementer:** An organization that executes a program to deliver the predefined results within an RBF instrument. Implementers are typically nonprofit organizations, or private service providers.
8. **Activity:** An implementer's action to generate outputs and achieve defined outcomes. For example, providing books to children (an activity) to ultimately improve child literacy test scores (a desired outcome).
9. **Results:** A generic term for outputs, outcomes, and impact.
10. **Output:** A tangible product directly produced by the implementer's activities. For example, providing books (an activity) directly results in 350 children receiving books (an output).
11. **Outcome:** A change in the beneficiaries' knowledge, skills, or behavior. For example, an implementer supplies books to children (an activity) and 350 children receive books (an output), which ultimately improves literacy test scores for these children (an outcome). Outcomes can be immediate, intermediate, and long-term.
12. **Impact:** The desired long-term and sustained effect of an implementer's program on the lives of its beneficiaries.



## 2.1 FOCUSING PROGRAMS ON THE DESIRED RESULTS, INSTEAD OF THE ACTIVITIES DELIVERED

In 2016, an evaluation of the United Kingdom government's £1.3 billion Troubled Families Programme concluded there was no discernible impact on unemployment, truancy, or criminality, the main outcomes targeted by the government.<sup>9</sup> Unfortunately, these stories have also become all too common in the field of international development, where complex social problems require persistent engagement and context-specific problem solving, rather than cookie-cutter solutions.

Results-Based Financing (RBF) offers a solution to the risk of funding activities that do not achieve the intended results for beneficiaries, by tying at least part of the funding for social programs directly to measurable results. This is a significant departure from traditional funding, which tends to be disbursed based on receipts or the execution of activities. RBF rewards implementers for achieving a predefined set of results, which, compared to traditional funding, can improve a program's impact and offer funders an additional guarantee of value-for-money.<sup>10</sup>

For instance, consider a maternal and child health program implemented in Rwanda: Researchers found that financing the program using RBF resulted in 23% more institutional deliveries and 132% more preventive visits for children aged two to five than when the program's funding was disbursed based on activities, despite receiving the same amount of money.

Imagine if, with the same funding, you could alleviate poverty 23% faster or provide life-saving health services, education, and livelihood opportunities to 23% more children. This guidebook is an invitation to consider, learn about, and thoughtfully apply RBF to significantly accelerate the pace at which implementers address pressing social problems.

The examples found in Figure 1 illustrate this central challenge that RBF seeks to solve: activities do not always lead to outcomes. For instance, building a school does not guarantee an improvement in literacy. The more uncertain and complex the pathway from activities to outcomes, the more valuable RBF can be in closing the gap between good intentions and results.

<sup>9</sup> O'Carroll (2016).

<sup>10</sup> Basinga et al. (2010).



**FIGURE 1.** Receiving payments for activities vs. payments for results

Program goal	Tying payments to activities	<b>Vs.</b>	Tying payments to results
<b>Improved education for children</b>	<b>Implementer receives payments for:</b> <ul style="list-style-type: none"> <li>·Building a school.</li> <li>·Giving a teacher training.</li> <li>·Providing books and supplies to students.</li> </ul>		<b>Implementer receives payments for:</b> <ul style="list-style-type: none"> <li>·Improvement in enrollment and attendance rates.</li> <li>·Improvement in test scores for reading comprehension and math.</li> </ul>
<b>Improved health for rural populations</b>	<b>Implementer receives payments for:</b> <ul style="list-style-type: none"> <li>·Building a clinic.</li> <li>·Building access roads to clinics.</li> <li>·Delivering medicines and supplies to clinics.</li> </ul>		<b>Implementer receives payments for:</b> <ul style="list-style-type: none"> <li>·Number of children fully immunized.</li> <li>·Reduction in the incidence of an illness.</li> </ul>
<b>Improved livelihoods for low-income populations</b>	<b>Implementer receives payments for:</b> <ul style="list-style-type: none"> <li>·Conducting an entrepreneurial skills training course.</li> <li>·Constructing a community garden.</li> </ul>		<b>Implementer receives payments for:</b> <ul style="list-style-type: none"> <li>·Increase in income.</li> <li>·Decrease in beneficiaries reporting poor food security.</li> </ul>

## 2.2 THE THREE KEY STEPS OF RBF

Results-Based Financing follows a three-step process. Below we expand on these steps in turn.



### I. Signing the RBF agreement

An outcome payer signs an RBF agreement with implementers, where it conditions all or part of its funding on the achievement of measurable, predefined results.<sup>11</sup> When signing the RBF agreement, outcome payers and implementers must decide on the following defining features:

1. Which results will trigger payment? For example, in the case of an education program, the relevant result could be the percentage improvement in literacy rates of students in grade 3.
2. How will the payment be linked to those results? In this same example, the outcome payer could agree to pay \$50,000 for every percentage point improvement in literacy rates for the target population.
3. How will results be verified? In this case, the outcome payer and the implementer could agree to compare the current year's literacy rate with the previous year to calculate the improvement, which would be verified by an independent evaluator.

These three features define the core incentives of the RBF agreement: which results it rewards, what these rewards are, and how to measure these results.

<sup>11</sup> Remember that “results” can refer to both outputs and outcomes. Further guidance on how to select outputs or outcomes is provided in Box 2.



## 2. Verifying results

An independent evaluator – who has no financial or other stakes in the results – measures and reports the results achieved. Continuing with the education example, at the end of the year an independent evaluator would measure literacy rates as compared to the previous year, and report on the progress made to the outcome payer.



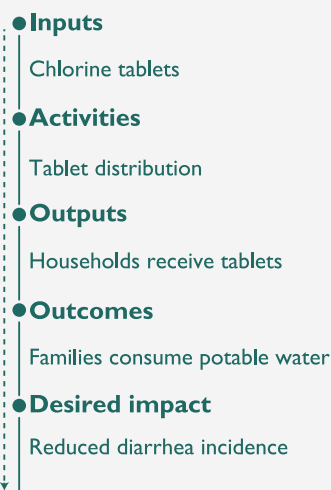
## 3. Paying for achieved results

The third step of RBF is the disbursement upon verification of achieved results. Following our simple example, if the independent evaluator reported that literacy improved by four percentage points, the outcome payer would disburse \$200,000 to the implementer.

### Box 2. Deep dive: Paying for results: outputs or outcomes?

In a theory of change (see Figure 2), outcomes are “closer” than outputs to the implementer’s desired impact. Paying for outcomes gives implementers incentives that more strongly link to their desired impact. From an incentive perspective, paying for outcomes is often the best choice for RBF.

**FIGURE 2.**  
Sample theory of change



Outputs, on the other hand, are “further away” from the desired impact and, unless there is a strong causal link between the output and impact, paying for outputs does not guarantee the achievement of the desired impact. For example, receiving chlorine tablets (an output) does not itself reduce the incidence of diarrhea (the desired impact). Households may not use the tablets or poor water quality may not be the cause of high incidences of diarrhea in the first place. Thus, paying for the delivery of chlorine tablets may not result in the achievement of the desired impact.

However, where there is a strong, evidence-based link between outputs and the desired impact, paying for outputs may be an effective option as they tend to be easier to measure, are observable within shorter timeframes, and are more firmly under the control of implementers. For example, consider a vaccine that is proven to effectively reduce the incidence of smallpox: measuring and paying for the number of vaccines administered (an output) may be less costly compared to measuring a reduction in the incidence of smallpox (the desired impact). Paying for outputs does not compromise the desired impact in this case.

For more on how to choose the results that are paid for, see Chapter 3.

## 2.3 THE POTENTIAL VALUE-ADD OF RBF FOR IMPLEMENTERS

When deciding whether to engage with RBF, one of the first key questions for implementers to consider is if RBF will add value to the organization and its programs. When designed well, RBF can offer at least two relevant value-adds to implementers: improving impact and expanding funding prospects.

### 2.3.1 IMPROVING IMPACT

When designed well, RBF provides implementers with the right mix of incentives and flexibility to deliver socially meaningful results. This helps implementers better serve their beneficiaries and achieve more impact with the same amount of funding.

Illustrating RBF's faculty to improve efficiency, consider improvements in health services following an RBF project in the Democratic Republic of the Congo. An impact evaluation showed that health facilities receiving payments for predetermined health services and quality indicators provided comparable or better services and quality of care than health facilities in control districts, despite lower levels of external financing.<sup>12</sup> Further, the study found that participating health facilities increased their revenues from user fees, reflecting patients' views that the quality of care was better in these facilities than in the control ones. Moreover, revenues grew without increasing out-of-pocket health spending of the poorest 25% households.<sup>13</sup>

The experience of RBF points to three principal drivers of impact:<sup>14</sup>

**a. Drawing attention to what matters:** You cannot manage what you do not measure.

Paying for results requires investing in data systems and regularly tracking results. Thus, RBF gives implementers continuous feedback on their impact trajectory. This new level of data-driven insight is an essential input for dynamic performance management (i.e., program adaptation to achieve greater results).

An example of the importance of a focus on results and a strong data system in supporting implementers to improve performance comes from the first Development Impact Bond (DIB) in India. RBF provided Educate Girls (EG), an education service provider, with rigorous impact evaluation reports on a yearly basis. In addition, the organization built a performance management system to collect real-time data on the impact of their program on girls' enrolment and children's literacy. The feedback loops prompted important improvements in the design and delivery of the program. For instance, when mid-year data showed some students lagging behind in English literacy,

<sup>12</sup> Approximately US\$ 2 per capita per year in participating districts and approximately US\$ 9 to 12 in control districts.

<sup>13</sup> Soeters et al. (2011).

<sup>14</sup> Adapted from Perakis & Savedoff (2015).

EG started to invest more in improving the teaching skills of the EG volunteers and restructured classroom time to include greater exposure to English. Consequently, EG not only closed the learning gap, but generated greater learning gains in English than in the other subjects.

**b. Aligning incentives to the beneficiaries' welfare:** What gets rewarded, gets done.

By tying funding to agreed-upon results, RBF aligns the objectives of outcome payers and implementers with the beneficiaries' welfare. Thus, RBF financially rewards the implementer for achieving its mission, rather than for meeting any other stakeholders' interests. Under RBF, implementer field teams are explicitly guided by the welfare of the beneficiaries, actively monitoring and managing external risks. Many implementers experience a mindset shift where the beneficiaries become clients.

For example, while conducting an education RBF program in India, EG faced a unique challenge when they discovered that some villages did not have upper primary schools where girls aged 11 to 14 could attend. In order to attend school, girls in those villages had to travel long distances to the closest school. Capacity concerns arose and it became apparent that it would not be possible to enroll many of these girls. Driven by the incentive to improve beneficiary welfare, EG quickly devised and implemented a strategy to provide continuity through distance education or open schools. Under a more typical program (i.e., not RBF), many implementers might have instead written this challenge off as an unexpected external factor.

**c. Providing flexibility to maximize results:** Improving performance is difficult under constraints.

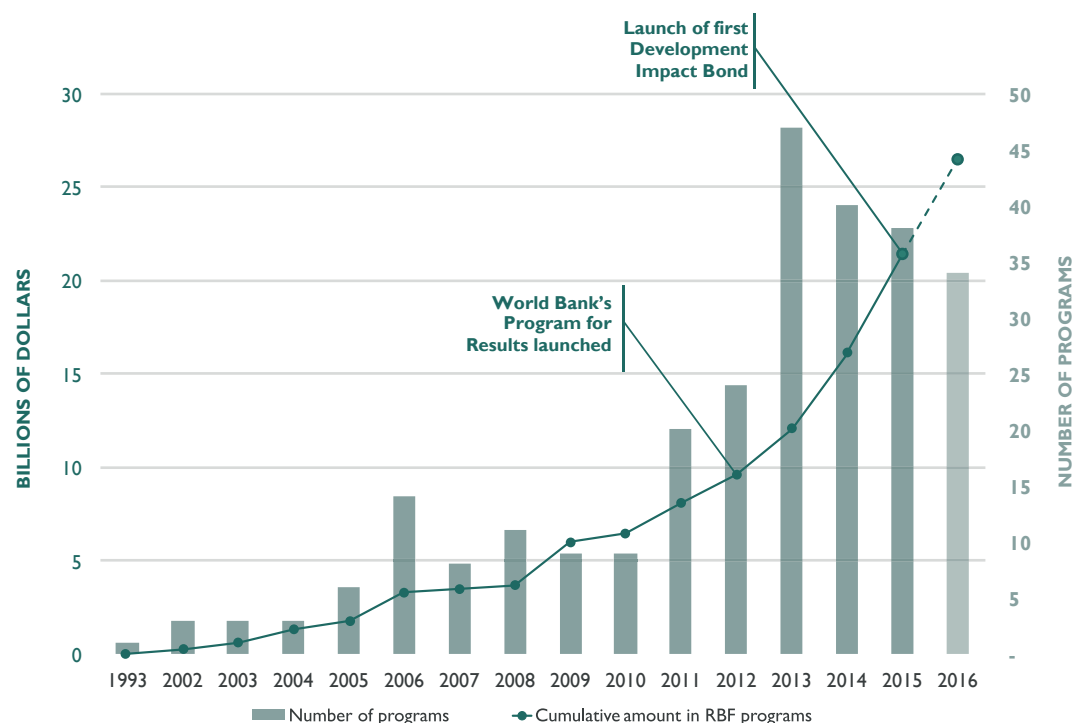
RBF offers the space for adapting and iterating program design and delivery practices. RBF tightens the control over achieving results; allowing outcome payers to relax their control over activities and grant implementers the flexibility and freedom to pursue a range of strategies to achieve greater impact. Implementers are free to invest in necessary overheads, to monitor and adapt their program to changing contexts, and to experiment with new delivery strategies without lengthy funder consultations or approvals.

For instance, consider an RBF agreement in a poverty graduation program in Uganda and Kenya that aims to increase beneficiary income. In this agreement, the implementer provides small seed grants to beneficiaries, together with business-related services and trainings, to help them launch a new business as a means to improve income. Outcome payers pay US\$ 1 for every US\$ 1 increase in beneficiaries' income. The design of the RBF agreement allows the implementer to adapt its program without the outcome payer's approval. Taking advantage of this flexibility, and propelled by data-driven insights, the implementer is currently tripling the size of its seed grants to further increase its impact on income.



## 2.3.2 EXPANDING FUNDING PROSPECTS

**FIGURE 3.** Financing tied to results in low- and middle-income countries



SOURCE: THE INSTIGLIO RBF DATABASE, APRIL 2017. THIS DATABASE COMPILES INFORMATION ON RBF PROGRAMS IN LOW- AND MIDDLE-INCOME COUNTRIES, UTILIZING SOURCES LIKE THE WORLD BANK, DFID, USAID, ASIAN DEVELOPMENT BANK, AND NORAD.

Funders, including governments, are facing pressure to achieve more impact with fewer resources and to justify the value of their international development strategies to their various stakeholders.<sup>15</sup> At the same time, they are emphasizing the use of evidence in funding decisions.<sup>16</sup> These two trends highlight the importance of delivering value for money and utilizing evidence as a means to allocate funding.

To achieve more with less and encourage more evidence-based programming, funders are rapidly expanding the use of RBF (see Figure 3), with more than US\$ 26.9 billion committed across 78 low- and middle-income countries.

<sup>15</sup> Insight deduced from discussions with government sector leaders surrounding the ability to create more program impact with less funding (see, for example, Killefer (2011)), as well as the growing trend of seeking alternative funding models for development programs amidst a shifting landscape and often declining aid budgets (see, for example, Saldinger (2017)). Further, regarding shifting landscapes and declining aid budgets, consider the impact of the 2007-2008 financial crisis and the UK's vote to withdraw from the European Union (a.k.a. "Brexit") (see, for example, Chonghaile (2016)).

<sup>16</sup> Stephens (2017).

The following three initiatives exemplify this trend:

- Since 2003, the Global Partnership on Output-Based Aid has tested over 40 projects using results-based approaches with more than US\$ 220 million in funding.<sup>17</sup>
- The World Bank's Program-for-Results has been steadily increasing in use since its creation in 2012, totaling almost US\$ 15 billion worth of RBF funding which has supported over US\$ 70 billion of government programs.<sup>18</sup>
- The United Kingdom's Girls' Education Challenge Fund has used RBF in 25 of its 37 education programs implemented since 2012, worth £300 million.<sup>19</sup>

In this shifting landscape, implementers who are willing to use RBF can improve their funding prospects in at least two ways:

**a. They are more competitive in securing funds**

RBF helps high-performing implementers make their programs stand out when competing for limited resources. Since RBF requires implementers to bet on their performance, only high-performing implementers tend to respond. Therefore, funders develop greater confidence in implementers who are ready to engage with RBF.

For instance, Village Enterprise, a non-profit delivering a poverty graduation program in east Africa, explored RBF after an impact evaluation showed its program to be effective at increasing income levels. Before this, the organization, with a total income of US\$ 2.9 million (FY2017), mostly received funding from smaller philanthropies and high-net worth individuals, which limited its funding base. By proposing to be paid entirely based on results in a forthcoming RBF agreement,<sup>20</sup> Village Enterprise secured US\$ 4.4 million in funding for outcome payments from two leading bilateral donors and a major donor-advised fund.

In the last three years, more than ten high-performing implementers have proactively explored RBF programs together with Instiglio. In an evidence-scarce environment, those who enjoy or can rapidly generate evidence are likely to be the first movers to catch the wave of RBF funding.

**b. They can engage new risk-averse funders**

Given funders only pay (at least the totality) if the program achieves predefined results, risk-averse funders may be more willing to fund RBF programs – particularly relatively innovative ones. This can help attract more funding and create opportunities to expand programs and generate more impact.

<sup>17</sup> GPOBA (n.d.).

<sup>18</sup> World Bank (2017).

<sup>19</sup> DFID (2014); DFID (2017).

<sup>20</sup> The Village Enterprise Development Impact Bond (VE DIB) is currently being finalized with agreement signatures across all parties expected in September 2017 and program implementation set to begin in November 2017. For reference, see the VE DIB Design Memo, which will be publicly available in Fall 2017, after program implementation commences.

For example, consider an organization working to reduce urban violence using behavior change methods, with promising evidence of their program's impact on homicide and shooting rates. Nevertheless, funders have no assurance that the program is effective across new geographies or target populations and are unwilling to commit funding. Not to mention that funders facing multiple restrictions, such as city governments, may be reluctant to fund an unconventional program with an innovative approach.

By using RBF, however, the program becomes a more attractive funding opportunity; the organization bears the risk of assuming the program's cost if results are not achieved, which in turn lowers funders' financial risk.

### Box 3. RBF is not a silver bullet

RBF should not be viewed as a panacea that guarantees all, or any, of the potential value-adds described in this section. Instead, the ability of RBF to deliver value depends on many factors. Implementers must identify which drivers of impact are likely to contribute the most to their organization and assess whether the RBF design activates those specific drivers. As Owen Barder from the Center for Global Development stated simply, “for payment by results to work, you have to get a lot of things right.”<sup>21</sup>

Poor design can not only fail to unlock RBF's potential, but can even hurt the impact of the program. Consider the following:

- If the results are not thoughtfully selected, RBF could draw attention to results that are not meaningful for the long-term welfare of the beneficiaries.
- Similarly, incentives could insufficiently affect the implementer's behavior; they could crowd out intrinsic incentives among implementing teams or, worse, they could provoke perverse behavior.
- If not properly bounded, flexibility could result in field teams exploring unproductive or superficial strategies to achieve short-term results at the expense of sustainable results.

In supporting the design of effective RBF, we detail key considerations to help implementers leverage its benefits in Chapter 4 and address central design considerations in Chapter 3.

<sup>21</sup> Barder (2014).

## 2.4 RBF INSTRUMENTS

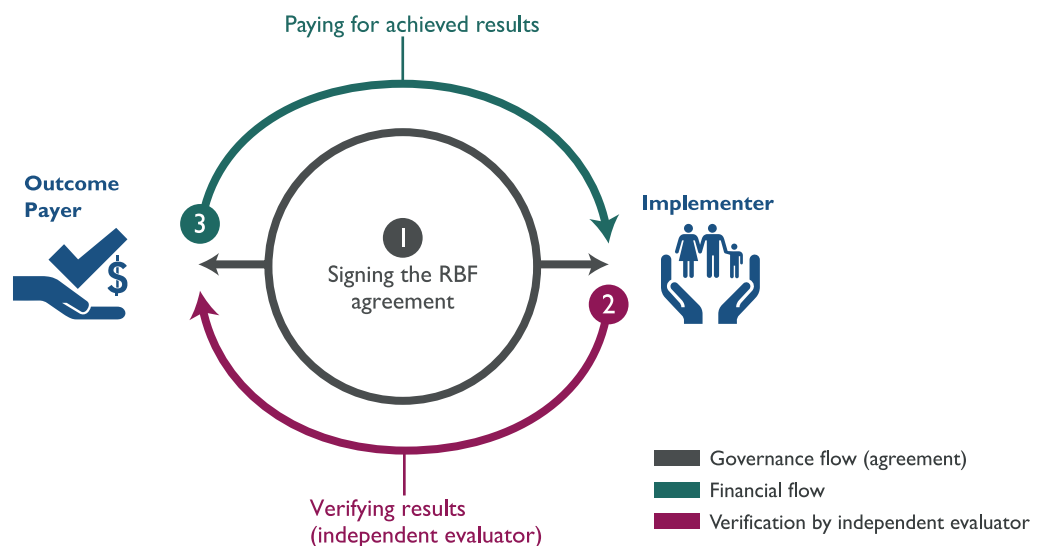
There are a variety of ways to structure RBF agreements and, over time, the terminology has emerged to categorize various types of RBF instruments. This section presents the most relevant instruments for implementers, including Performance-Based Contracts, Impact Bonds, and Prize-Based Challenges. A complete list of RBF instruments can be found in Appendix 2.

Despite this common terminology, RBF is less a set of rigid instruments and more a practice of thinking strategically about the power of financial incentives and how they can motivate better program performance. There are endless ways of structuring, tweaking, and adapting RBF to fit a given context and the desired performance. The RBF instrument that works best for a young organization in a learning phase is different from the instrument needed by an evidence-based organization in a scale-up phase. This section concludes with a framework of the RBF instruments that are typically best suited to programs depending on their maturity.

### » Performance-Based Contract (PBC)

In a PBC, the outcome payer conditions part of its payment to an implementer on the achievement of predefined results. Upon the verification of the achievement of results by an independent evaluator, the outcome payer disburses a payment to the implementer. Figure 4 depicts the structure of a PBC.

**FIGURE 4.** Performance-Based Contract structure



The proportion of funding that is conditioned upon the achievement of results can vary, but often constitutes a small portion of the total contract size. PBCs therefore transfer some risk to implementers, as they do not receive full payments if results are not achieved.

The greater the portion of funding tied to results, the greater the risk for implementers. The option with the least risk for implementers utilizes a bonus payment for the achievement of predefined results. In this case, 0% of the funding needed to implement the program is conditioned upon the achievement of results, and the implementer receives additional bonus payments as a reward for extraordinary performance.

#### Box 4. Spotlight on a PBC - Escalera Foundation

In Chiapas, Mexico, an average of 40% of middle school students do not continue to high school despite the proven benefits of every additional year of education and the widespread and longstanding government conditional cash transfer program Oportunidades/Prospera that pays mothers for sending their children to school. The amounts transferred increase for higher grades and represent a significant proportion of average family incomes in the region.

In light of this, Escalera, a local NGO in Chiapas, wanted to test whether under-enrollment was best addressed by targeting an information gap (e.g., students lack education role models, have poor access to information on scholarships and the enrollment process, and do not understand the real economic returns of additional schooling), or by targeting economic barriers.

Escalera delivered high school preparation curriculum kits to some public schools one year before high school enrollment. These kits contained practical and motivational information to help students make an informed decision regarding high school enrollment. Other schools received vouchers that would cover students' high school entrance exam and enrollment fees. Some schools received both the vouchers and the kits.

In 2014, the Government of Chiapas signed a PBC to enable Escalera to scale up their program to 2,700 students who otherwise would not have participated in the program. Under this contract, the Government of Chiapas, the outcome payer, paid Escalera for each additional student that enrolled in high school because of the program. After the first year, a randomized evaluation found that high school enrollment increased by 6 percentage points due to the program, and that vouchers were more impactful and cost-effective than the kits. In 2015, Escalera and the Chiapas government signed a three-year PBC to continue to scale the program. The PBC allowed Escalera to double the program's budget, which will benefit 42,600 students over four years of operation.<sup>22</sup>

#### » Social Impact Bond and Development Impact Bond (SIB and DIB)

Impact Bonds utilize an investor that provides the upfront working capital to the implementer to deliver their program. The outcome payer repays the investor(s), often with a return, only if results are achieved and have been verified by an independent evaluator. Impact bonds differ from PBCs in that they shift the financial

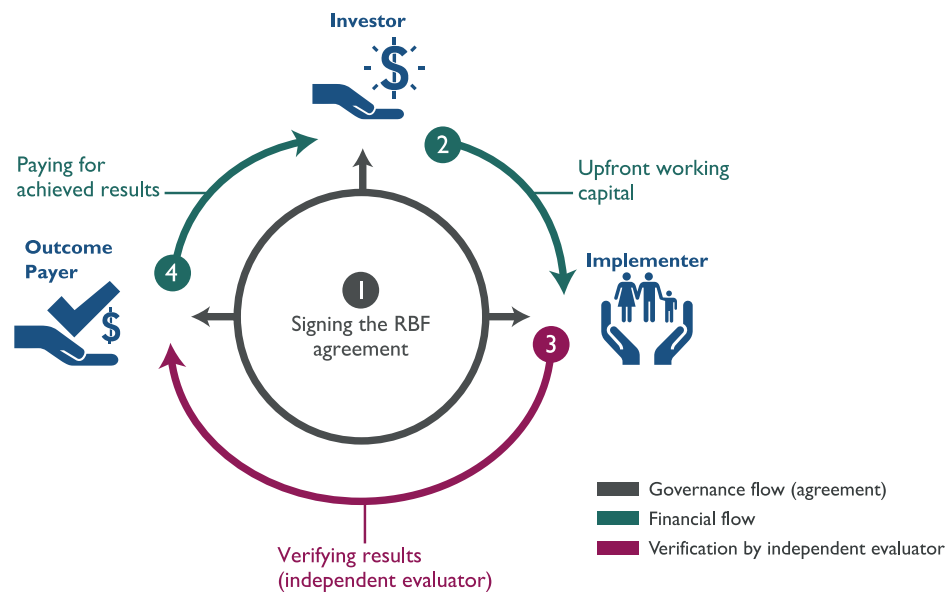
<sup>22</sup> Unfortunately, there is currently a risk that the government of Chiapas will not fulfill its payment obligation. The financial consequences of this would be particularly damaging because Escalera has already finished the implementation of the program cohorts in the contract and are awaiting results which are a few months out.



risk of not achieving results from the implementer to the investors. As such, Impact Bonds respond to the constraint of needing upfront capital for program implementation.

When the outcome payer is a government, we use the term Social Impact Bond (SIB), and when the outcome payer is a funder, such as a foundation or a bilateral or multilateral development agency, we use the term Development Impact Bond (DIB). Figure 5 depicts the structure of an Impact Bond.

**FIGURE 5.** Impact Bond structure



#### Box 5. Spotlight on a DIB - Educate Girls

In India, an estimated 3.7 million girls are out of school. In Rajasthan, 40% of girls drop out before reaching fifth grade and, for those who remain, learning quality is low. Uneducated girls in India are three times more likely to contract HIV, earn 10% less income, and marry three years earlier than educated ones. To address this enrollment issue, Educate Girls (EG), a local implementer, identifies champions for girls' education at the village level who can identify out-of-school girls and work with the community to enroll and teach them in the classroom.

In 2015, the Children's Investment Fund Foundation (CIFF) (outcome payer), EG (implementer), and the UBS Optimus Foundation (investor), created a Development Impact Bond (DIB) to help EG improve its impact on enrollment, retention, and learning for marginalized girls and boys in Rajasthan, India. Under the DIB, all parties agreed to tie payments to results.

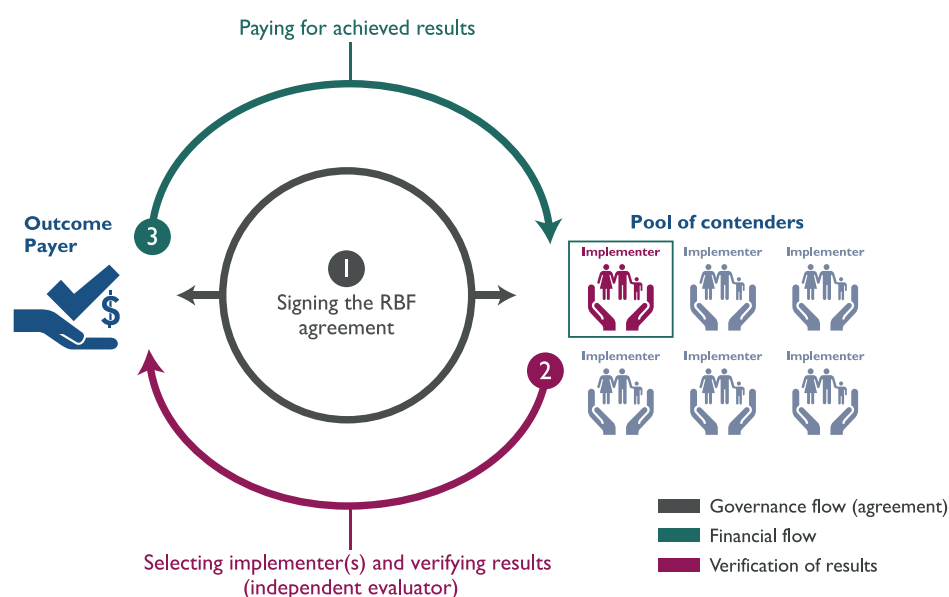
The investor, UBS Optimus Foundation, provides upfront capital to EG and EG delivers their program in Rajasthan. Upon verification of the achievement of predefined results – conducted by IDinsight, an independent evaluator – CIFF will pay the UBS Optimus Foundation their initial investment plus a return, depending on performance. Results after two years show that EG has achieved 87.7% of the 3-year enrolment target and 50.3% of the 3-year learning target. The investor remains on track to recoup its initial funding of US\$ 267,000.

## » Prize-Based Challenge

A Prize-Based Challenge is an open bid competition that awards a financial prize for the best innovation developed within a predefined timeframe. Although similar to other forms of competition, Prize-Based Challenges reward competitors for delivering the best solution rather than producing the best proposal, with competing solutions often assessed based on the results they deliver. Figure 6 depicts the structure of a Prize-Based Challenge.

Note that a Prize-Based Challenge differs from Impact Bonds and Performance-Based Contracts in that it engages many competing implementers who usually develop a *new* solution, rather than focusing on improving the impact and funding prospects of implementers' existing programs.

**FIGURE 6.** Prize-Based Challenge structure



### Box 6. Spotlight on a Prize-Based Challenge - XPRIZE

Global Learning XPRIZE<sup>23</sup> is a program that challenges 135 global teams to develop software that enables children in developing countries to teach themselves basic reading, writing, and arithmetic within 18 months. While the use of multiple stages is not required, Prize-Based Challenges may often utilize stages to further incentivize participation and assist the selection process. During the first stage of XPRIZE, five finalists receive US\$ 1 million each to field-test their solutions with 3,000 to 4,000 children. Finalists are selected based on the likelihood of the technology producing significant learning gains and on various design features, such as usability and customizability. In the second stage, a winning team is selected and awarded US\$ 10 million, which they can use to scale up their solution and enter the market. The winner is the team that generates the greatest improvement in literacy and numeracy among 7- to 10-year-olds. Improvement will be assessed relative to a control group, which receives no learning technology.



### » Applicability of RBF instruments

Programs tend to benefit from different characteristics of RBF instruments as they develop stronger evidence of the contributors to their impact and mature from an innovative idea, to a promising solution, to an evidence-based program. As mentioned, RBF is an evolving practice rather than a collection of rigid instruments. Nevertheless, the best-suited RBF instrument is typically a function of the program's maturity.

### Classifying programs according to their maturity

One way of mapping RBF instruments to your context is to think about the maturity of your program in terms of its evidence base. Let's first define three broad categories of maturity:

1. **Outcomes-oriented innovation.** A low program maturity is a scenario of no or little evidence, where the solution to a social problem is unknown or there are many potential competing ideas to address the problem. First order questions, such as the best beneficiary recruitment strategy or the right theory of change, must be addressed. For example, effective responses to mitigate the spread and impact of the Zika virus would fall in this category.




<sup>23</sup> Global Learning XPRIZE (n.d.).

2. **Catalyzing promising interventions.** Most development programs have some promising evidence of impact, yet questions remain regarding how impactful and cost-effective they can be across different contexts. For example, consider the poverty graduation approach; a promising multifaceted intervention aimed at sustainably transitioning the ultra-poor out of poverty, which has shown high impact in some countries, but less and even negative impact in others.<sup>24</sup>
3. **Scaling with impact.** The other end of the maturity model (high maturity) is a scenario where evidence-based programs generate the desired outcomes with high predictability in a variety of different contexts. Examples include school-based deworming or immunization. These solutions are known to be impactful and the goal is simply to scale them up with quality. This phase is characterized by a management and delivery challenge, as opposed to an innovation or discovery challenge.

### How can RBF help to deliver greater impact in each stage of maturity?

RBF can add different forms of value to programs across this maturity spectrum (Figure 7).

**FIGURE 7.** Typically best-suited RBF instruments by level of program maturity

Level	Why RBF?	Common instruments
<div style="writing-mode: vertical-rl; transform: rotate(180deg);">PROGRAM MATURITY</div> <div style="display: flex; align-items: center;"> <div style="width: 10px; height: 100px; background: linear-gradient(to bottom, red, grey);"></div> <div style="margin-left: 5px;"> <div style="width: 10px; height: 10px; background-color: red; margin-bottom: 5px;"></div> <div style="width: 10px; height: 10px; background-color: grey; margin-bottom: 5px;"></div> <div style="width: 10px; height: 10px; background-color: grey;"></div> </div> </div> <div style="display: flex; justify-content: space-between; margin-top: 10px;"> <span>LOW</span> <span>HIGH</span> </div>	<p><b>Outcomes-oriented innovation:</b> By specifying desired outcomes, and providing full discretion and incentives, RBF can incentivize rapid learning toward the development of a credible and effective solution.</p> 	<ul style="list-style-type: none"> <li>• Prize-Based Challenges</li> <li>• PBCs</li> </ul>
	<p><b>Catalyzing promising interventions:</b> By providing incentives to learn and adapt your program and the flexibility required to stretch the impact of promising interventions, RBF can help to improve results and facilitate their adoption.</p> 	<ul style="list-style-type: none"> <li>• Impact Bonds</li> </ul>
	<p><b>Scaling with impact:</b></p> <ul style="list-style-type: none"> <li>• By incentivizing results, RBF encourages implementers to keep the quality of their program at the center while implementing it at scale.</li> <li>• SIBs have the unique property of attracting private investors to scale up evidence-based preventative services that can reduce treatment costs down the line.</li> </ul> 	<ul style="list-style-type: none"> <li>• PBCs</li> <li>• Impact Bonds (only for prevention programs)</li> </ul>

During the **outcomes-oriented innovation phase**, the goal is to incentivize rapid learning towards the development of a credible and effective solution for a social program. Both Prize-Based Challenges and Performance-Based Contracts can be used in this phase under certain conditions:

- **Results-oriented Prize-Based Challenges** are most suitable to boost the discovery of new solutions that deliver the outcomes of interest. In essence, by specifying and drawing attention to the

<sup>24</sup> Banerjee et al. (2015).

desired outcomes, providing full discretion to a number of competing organizations and a large enough prize for the winner, Prize-Based Challenges set the right incentives for a group of problem-solvers to accelerate the development of a solution.

- **Low-stakes Performance-Based Contracts** can also be used in this phase and are particularly suitable for programs that have a good sense of the key activities needed to deliver results. In this case, providing a bonus payment for the achievement of results draws the implementer's attention to what really matters and provides the incentives to invest in continuous learning. Given their lack of evidence, transferring a higher payment risk by conditioning some portion of the funding to results is not advisable. However, given the low maturity of the program, the funding provided to implement the program should not limit the implementer's flexibility to adapt it. With tightened accountability, measurement of results, and alignment of incentives provided by the PBC, a funder may be willing to provide this funding with greater flexibility; for example, by defining a menu of potential activities and allowing the implementer to vary the activities it uses.

During the **catalyzing promising interventions phase**, the aim is to provide incentives to learn and adapt the program in order to become more cost-effective.

- **Social and Development Impact Bonds** are interesting mechanisms in this phase because they create a highly incentivized environment by tying all of the program funding to results. They also give full flexibility to iterate the design and delivery practices. Impact Bonds are thus a type of incubator that stretches an organization's impact; but, given their relatively high costs (e.g., transaction costs, return for investors), they are usually not a suitable instrument for programs whose impact has already been stretched and which are ready to be implemented at scale.

In the **scaling with impact phase**, the goal is to provide the right incentives to the management and frontline workers to keep the quality of their program at the center, while implementing it at scale. The first order question is about how to deliver a proven cost-effective program well at scale, and flexibility to adapt the program is not the key driver of impact.

- **Performance-Based Contracts** are often the right instrument in this phase. Tying that portion of funding to results strikes a balance between the potential benefits – including creating effective incentives – and the associated costs – including the upfront capital the implementer needs to raise and the risk premium associated with transferring payment risk to the implementer.
- **Impact Bonds** have a role in this phase in particular contexts. Sometimes, governments do not have sufficient resources to contract preventative services given the already high treatment costs associated with certain diseases (such as diabetes). In this case, choosing an evidence-based preventative program and scaling it up using the upfront capital provided by investors will reduce the treatment costs down the line, freeing up the resources the government is currently spending on treating the problem to pay back investors. Thus, Impact Bonds help to solve a credit problem through private money, and reverse the vicious spending cycles while fundamentally addressing a social issue.



## 2.5 KEY TAKEAWAYS

- Results-Based Financing is a program financing arrangement in which payments are contingent upon the achievement of predefined results, which are usually verified by an independent evaluator.

### The three key steps of RBF

- The three principal steps of RBF are signing the RBF agreement, verifying results, and paying for achieved results.

### The potential value-add of RBF

- The potential value-add of RBF to implementers is a) improving impact and b) expanding funding prospects.
  - RBF may improve impact through the following drivers: by drawing attention to what matters, by aligning incentives to beneficiaries' welfare, or by increasing flexibility in implementation.
  - RBF may expand a program's funding prospects because it makes the implementer more competitive for funding or helps the implementer access funding from risk-averse funders.
- RBF is not a silver bullet. It is important for implementers to understand that value is not automatically generated by engaging in RBF. Therefore, implementers can benefit from an assessment of the drivers of impact which could add the most value to their organization and of whether the RBF design activates those desired drivers appropriately.

### RBF instruments

- RBF instruments relevant for implementers include Performance-Based Contracts, Impact Bonds, and Prize-Based Challenges.
  - In a Performance-Based Contract, the outcome payer conditions part of its payment to an implementer on the achievement of predefined results. Upon the verification of the achievement of results, the outcome payer disburses payment to the implementer.
  - In an Impact Bond, an investor provides the upfront working capital to the implementer to deliver their program. The outcome payer repays investors, often with a return, only if results are achieved.
  - In a Prize-Based Challenge, the outcome payer finances an open bid competition that awards a financial prize for the best innovation developed within a predefined timeframe. Although similar to other forms of competition, Prize-Based Challenges reward competitors for delivering the best solution rather than producing the best proposal.
- RBF is an evolving practice rather than a collection of rigid instruments. Nevertheless, the best-suited RBF instrument is typically a function of the program's maturity.









# 3

## ASSESSING THE IMPLICATIONS OF THE RBF DESIGN

*The powerful potential of Results-Based Financing will only be realized if it is tailored to programmatic and organizational needs. Especially where funders initiate the RBF design process, implementers need to be equipped to assess how the RBF design may enable or hinder their success, and negotiate the right RBF.*

*This chapter begins with a brief overview of the various RBF design components and the RBF design and contracting process. Next, it introduces the key criteria an implementer should use to evaluate each design component. The core of the chapter explores the design components in turn and describes how they affect these criteria. Nuanced considerations are provided, including how the effects of design choices may differ across sectors and vary with the implementer's capacity and evidence base. Furthermore, important contract features of the RBF agreement are introduced. Finally, the chapter concludes by providing recommendations on how implementers might negotiate favorable design components with outcome payers.*

# ASSESSING THE IMPLICATIONS OF THE RBF DESIGN

## 3.1 INTRODUCTION TO THE RBF DESIGN COMPONENTS AND PROCESS

### » RBF design and contract components

The RBF design is critical for your success because it determines the focus, incentives, and risks for the implementation of the program. The design centers on the components that are outlined in the table below and addressed in sections 3.3 to 3.6.

**TABLE 1.** RBF design components

RBF design components	Description	Example: Educate Girls Development Impact Bond <sup>25</sup>
<b>Boundaries of the project</b>	The boundaries of the project include the financial size, the duration, and the target population under the RBF agreement. Note that these are not specific to RBF.	<ul style="list-style-type: none"> <li>• <b>Financial Size:</b> US\$ 0.27 million (investment) + 10% (internal return rate).</li> <li>• <b>Duration:</b> 3 years.</li> <li>• <b>Target population:</b> Approximately 15,000 girls and boys in 166 schools in Bhilwara district of Rajasthan.</li> <li>• <b>Learning:</b> Girls and boys in grade 3 to 5.</li> <li>• <b>Enrollment:</b> Out-of-school girls aged 7-14 eligible for enrollment into grade 2-8.</li> </ul>
<b>Payment metrics</b>	The payment metrics define the results that are paid for and apply to the defined target population.	<ul style="list-style-type: none"> <li>• <b>Learning</b> in Hindi, English and math measured by the ASER test.</li> <li>• <b>Enrollment</b> into government schools.</li> </ul>
<b>Evaluation and measurement method</b>	The evaluation method determines the degree to which results are attributable to the implementer. The measurement method defines how, when, and by whom data is collected and results measured.	<p><b>Learning</b></p> <ul style="list-style-type: none"> <li>• Evaluation method: Results are measured against a control group in a randomized evaluation, with impact defined as the difference in students' learning gains between the treatment and control groups.</li> <li>• Measurement method: ASER test administered yearly by the independent evaluator.</li> </ul>

<sup>25</sup> Instiglio (2015).



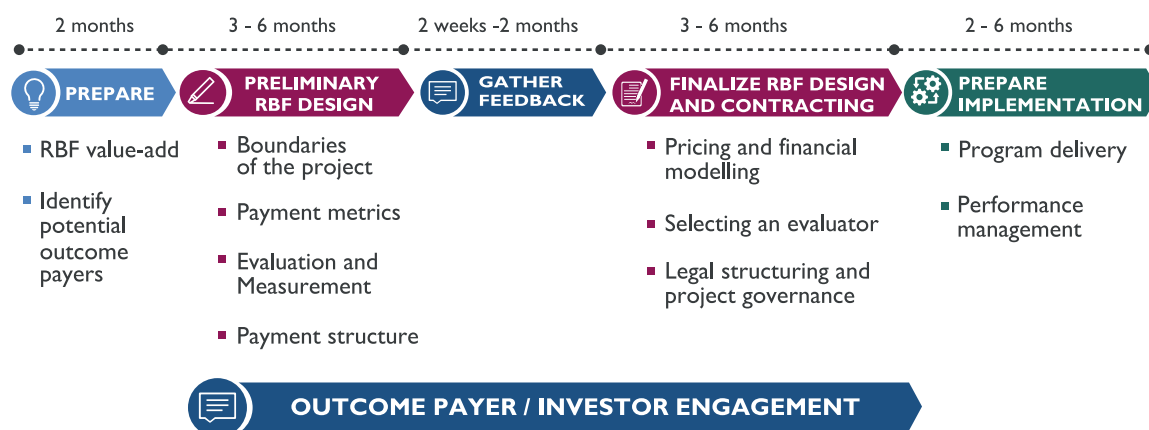
RBF design components	Description	Example: Educate Girls Development Impact Bond
		<b>Enrollment</b> <ul style="list-style-type: none"> <li>Evaluation method: Pre-post analysis. Enrollment status is measured for out-of-school girls at the start of the program.</li> <li>Measurement method: EG measures the enrollment status of out-of-school girls using school registers (yearly). Enrollment is verified on a sample basis by the independent evaluator.</li> </ul>
<b>Payment structure</b>	The payment structure includes a) the payment function which specifies how payment varies with the results achieved and b) the payment schedule, which determines the timing of payment to the implementer (or investor(s)).	<b>Payment function</b> <ul style="list-style-type: none"> <li>Maximum outcome payment: US\$ 0.42 million.</li> <li>80% of the total outcome payment is allocated to the achievement of learning outcomes, 20% is allocated to the achievement of enrollment outcomes.</li> <li>Price per unit of results achieved: US\$ 49 per unit of learning gained, US\$ 910 per percentage point enrollment.</li> <li>Payments are linear (i.e., same payment for each unit of results achieved).</li> <li>Payment schedule: All payments will be made at the end of the contract.</li> </ul>

In addition to the RBF design, good contracting is critical to adequately reflect the terms of the agreement, protect parties against risks that are outside their manageable control, and to specify protocols for addressing a variety of scenarios. Section 3.7 introduces important contract features of the RBF agreement.

### » RBF design and contracting process overview

The RBF design process is critical in achieving alignment between various stakeholders. The outcome of this process, the RBF design, may vary significantly depending on whether it is led by the implementer or developed in response to a proposition from a funder. If you are analyzing an existing opportunity, it is less likely that you will be able to iterate and tailor the design to your organization's needs as much as if your organization is leading the process. See section 3.8 of this chapter for tips and techniques to negotiate favorable design components with funders.

For implementer-led design processes, Figure 8 provides an overview of the typical design process and indicative timelines. Understanding the design process and timeline will help implementers assess whether they can afford this process and plan better. Knowing when to get feedback from stakeholders helps to create buy-in and alignment in an efficient manner.

**FIGURE 8.** Implementer-led design process

Before starting the design process, implementers should think about how RBF adds value to their organization and programs (see Chapter 2). We also recommend identifying and engaging potential funders early on to ensure their preferences and constraints can inform the design choices.

Moving to the design process, we recommend developing a “rapid prototype” or “preliminary RBF design.” The prototype is essentially a rough draft of the full RBF design and is intended to make key design components tangible enough for stakeholders to understand and react to. Prototyping allows for rapid feedback from stakeholders on key design choices such as payment metrics, before investing time in creating detailed budgets and pricing results. Successful prototyping helps to drastically improve the efficiency of the design process compared to traditional stop-and-go design processes, which seek consensus on one design feature before moving to the next. Design features in RBF are all inter-connected and one can only iterate a feature in connection to everything else. Rapid prototyping has helped drive down design costs and timelines by at least a factor of two. For instance, this technique enabled the completion of a Social Impact Bond design in 4 months in Colombia, compared to our Development Impact Bond in India, which took 8 months.

Following the prototyping, more detailed design work is undertaken based on iterative feedback from stakeholders, and once the specifics of the contracting are finalized between the outcome payer and the implementer (along with any other parties, such as the investor or independent evaluator). The efforts and timelines associated with this phase will vary depending on the number of stakeholders and how well the prototype responds to their initial preferences.

Finally, planning sufficient time to set up program implementation, hire key staff, build capacity, finalize agreements with sub-contractors – and if relevant, investors – and possibly establish performance management systems (see Chapter 4) ensures your organization can focus on reaching the desired impact from the start of the RBF agreement. Planning for a healthy mobilization period is critical. In the words of the Educate Girls program manager, for lack of a sufficient mobilization period the first few months “felt like flying a plane while building it.” Preparatory work and capacity building took efforts and resources away from the delivery of the intervention. The mobilization period can easily take between 2 to 6 months depending on the complexity of the implementation set up.

## 3.2 GUIDING QUESTIONS FOR THE ASSESSMENT

There are two primary considerations when assessing the design of an RBF agreement:

- (1) **Alignment to your organization's desired social impact** and;
- (2) **Payment risks** or the likelihood that you will not receive payment conditioned on results.

### » Alignment with your organization's desired social impact

The RBF agreement should support your organization in achieving its mission. To assess alignment, the desired social impact of your organization must first be clear. This includes various dimensions, such as: a) the socio-demographic and geographic characteristics of the beneficiaries you want to impact; b) the intended impact on the beneficiaries; and c) the time period over which you want to achieve the desired impact.

A mission-aligned RBF agreement creates the incentives and provides the space for your organization to prioritize what matters to achieve your mission. A good RBF agreement should also minimize any undesirable effects, such as crowding out intrinsic motivation or incentivizing activities that do not contribute to, or even detract from, the achievement of the desired impact. If incentives are not well aligned with your organization's desired impact, this might result in:

- (1) Prioritization of individuals in the highest income quintile to reach the predefined results, when your organization's mission is to improve the lives of the lowest income quintile;
- (2) Prioritization of a result that does not contribute to the intended impact on the beneficiary; or
- (3) Prioritization of improving results in the short-term, such as savings, when your mission is to sustainably lift beneficiaries out of poverty through building their income-generating capacity.

Perverse incentives are not unique to RBF agreements. Traditional financing that pays for activities may also create perverse incentives; for example, by incentivizing implementers to adhere to a rigid implementation plan, rather than adapt when they know that the program can be improved. In fact, RBF agreements are designed to mitigate perverse incentives created by activity-based financing. Nevertheless, one must be careful not to introduce new perverse incentives through the RBF agreement.

### » Appropriate payment risk

Both activity-based financing and RBF involve certain financial risks, such as unpredicted increases in the costs of program delivery, changes in the exchange rate, or the default of the funder. Given that they tie payments to the achievement of pre-defined results, RBF agreements carry the additional risk that your organization will not earn the payments because it does not deliver expected results. We will refer to this as the “payment risk.” There are two main sources of payment risk:

- (1) The expected results are not (fully) achieved. This could be because your organization did not have the capacity or flexibility to implement RBF well or because the occurrence of external factors negatively affected your program’s effectiveness. It could also be that the targets for the expected results were set too high at the outset. Taking more ambitious bets on results may make your organization more competitive for funding. However, that also increases the payment risk.
- (2) The selected evaluation and measurement method did not accurately measure the results delivered by your program and led to under-payment.

You should carefully assess RBF agreements to understand the sources of payment risks, and ensure that you either effectively mitigate these risks through a negotiated RBF design or that you seek fair compensation for taking on these risks.

### » Other criteria

Throughout the chapter, we also emphasize a variety of other benefits and costs your organization may consider important in assessing an RBF agreement, including:

- a. **Evidence and learning.** Because RBF implies the generation of evidence, RBF projects can add to your evidence base and allow your organization to learn more about the aspects of your program that contribute most to your organization’s impact.
- b. **Capacity-building.** Because funders are starting to recognize that RBF requires new systems and competencies, RBF projects often come with capacity-building resources, enabling investments in data systems and staff training, for instance.
- c. **Replicability or long-term suitability of the specific RBF design.** Because RBF can be an attractive long-term funding mechanism for funders (particularly governments), the RBF project can be designed for future use in different contexts and at greater scale.
- d. **Costs.** Because RBF entails administrative costs and financing costs, the benefits associated with the RBF project should justify its costs. Measurement costs are usually assumed by the outcome payer, but become particularly relevant for implementers where the total funding for the RBF project is fixed and higher measurement cost results in less funding for program implementation.

### 3.3 ASSESS THE BOUNDARIES OF THE PROJECT: FINANCIAL SIZE, DURATION AND TARGET POPULATION

The boundaries of RBF projects are similar to other projects, including the financial size, duration of the agreement, and target population. We introduce these terms below and outline key considerations for implementers specific to RBF.

#### » Financial size

The financial size is the total funding provided under the RBF agreement, including funding not tied to results and any payments made if pre-defined results are achieved. As with other projects, the size of the RBF agreement is largely determined by the funding required to implement the program at the desired size and the funder's willingness to pay.

Implementers should ensure funding levels are sufficient to cover quality program delivery and any additional capability building, which is needed in most RBF projects. This includes the costs of organizational capacity building linked to RBF, including accessing the right skills and processes to capture and act on lessons learned (see Chapter 4).

As with other projects, implementers should pay attention to the trade-offs involved between:

- a) Smaller scale implementation which places less strain on organizational leadership and capacity and allows for more rapid iterations and adaptations of the program based on lessons learned; and
- b) Larger scale implementation which benefits from economies of scale, allowing for more cost-effective programs.

Especially when implemented for the first time, RBF projects typically place greater demands on the leadership team of an organization. To increase the chances of achieving the agreed-upon results, we recommend implementing RBF projects that are small enough in their geographical and programmatic scope that they can be staffed with high-quality teams and the organization is not diluting the quality. Smaller scale RBF projects also pose lower absolute financial risk if results are not achieved.

#### » Duration

In assessing the proposed duration of an RBF agreement, implementers need to consider the optimal duration of their program to achieve and measure the desired results and impact. Implementers and other project stakeholders may seek contract periods longer than the program duration to allow them to work with multiple sequential cohorts and integrate learnings from one cohort into the program design and implementation for the next cohort.



The optimal contract duration may, however, be constrained by other factors, such as the outcome payers' contracting rules or government administrative cycles. For example, governments are often restricted in committing funding beyond budgetary cycles. Likewise, implementers or investors may seek to limit the contract duration to reduce the risk associated with external factors impacting results.

### » Target population

RBF agreements are based on achieving results with a clear target population, often defined based on socio-economic and geographic characteristics. In defining the target population, implementers should focus on the population with which their mission is aligned and for whom their program has been designed. Further, implementers should consider their prior experience and track record of achieving results with that target population and ensure that there are no barriers to working with the target population. This is especially important given the payment risk that implementers face.

## 3.4 ASSESS THE PAYMENT METRICS

Payment metrics (the results that are paid for) are the foundation of every RBF agreement. They define what success means and have a strong influence on the incentives provided and the risk transferred. The payment metrics apply to a given target population. Table 2. provides examples of payment metrics and target populations from existing RBF agreements.

**TABLE 2.** Examples of payment metrics

RBF agreement	Payment metrics	Target population
<b>Colombia Workforce Social Impact Bond (SIB)</b> <sup>26</sup>	<ol style="list-style-type: none"> <li>1. # of participants in formal work at the end of the intervention.</li> <li>2. # of participants in formal work 3 months after the intervention.</li> <li>3. # of participants in formal work 6 months after the intervention.</li> </ol>	<p>514 unemployed vulnerable individuals in Bogota, Cali, Pereira who meet the following criteria used by Prosperidad Social (the sponsoring government entity):</p> <ul style="list-style-type: none"> <li>• Have a SISBEN score (poverty measure) of 0 to 41.74, are registered in Red Unidos (ultra-vulnerable group) or are victims of displacement due to the armed conflict;</li> <li>• Are between 18 and 40 years old;</li> <li>• Are high-school graduates; and</li> <li>• Have not participated in Prosperidad Social's employment programs in the last two years.</li> </ul>
<b>Vietnam Rural Water Project (PBC)</b> <sup>27</sup>	<ol style="list-style-type: none"> <li>1. # of households connected to water supply.</li> <li>2. # of households with 6 months of satisfactory water service delivery.</li> </ol>	30,000 households in low-income communities in five provinces of central Vietnam.

<sup>26</sup> Departamento Administrativo para la Prosperidad Social (2017).

<sup>27</sup> World Bank (2016b); World Bank (2007).

RBF agreement	Payment metrics	Target population
<b>Village Enterprise Development Impact Bond (DIB)</b> <sup>28</sup>	<p>Increases in consumption and assets, which are defined as follows:</p> <ol style="list-style-type: none"> <li>1. Consumption: The sum of household food and beverage consumption, household recurring consumption, and household infrequent consumption. Consumption will be measured based on the widely-utilized Consumption and Expenditure survey.</li> <li>2. Assets: The net household assets (household savings and tangible household assets, net of household liabilities) plus net business assets (business and Village Enterprise business savings groups savings and tangible business assets, net of business liabilities), accounting for business ownership by the household. No distinction between productive and non-productive assets is made.</li> </ol>	<p>Approximately 18,000 ultra-poor households living in rural Kenya and Uganda. Beneficiary households will be selected through a targeting exercise which is based on the <i>Progress Out Of Poverty Index</i>. Each beneficiary household will select one family member to participate in the graduation program.</p>

Choosing the right payment metrics is critical for the success of an RBF program. The payment metrics are the source of many of the incentives that your organization will experience and will strongly shape what your teams do on the ground. This section provides the key considerations when selecting or negotiating payment metrics, beginning with an explanation of how your organization's theory of change is helpful for the assessment of payment metrics.

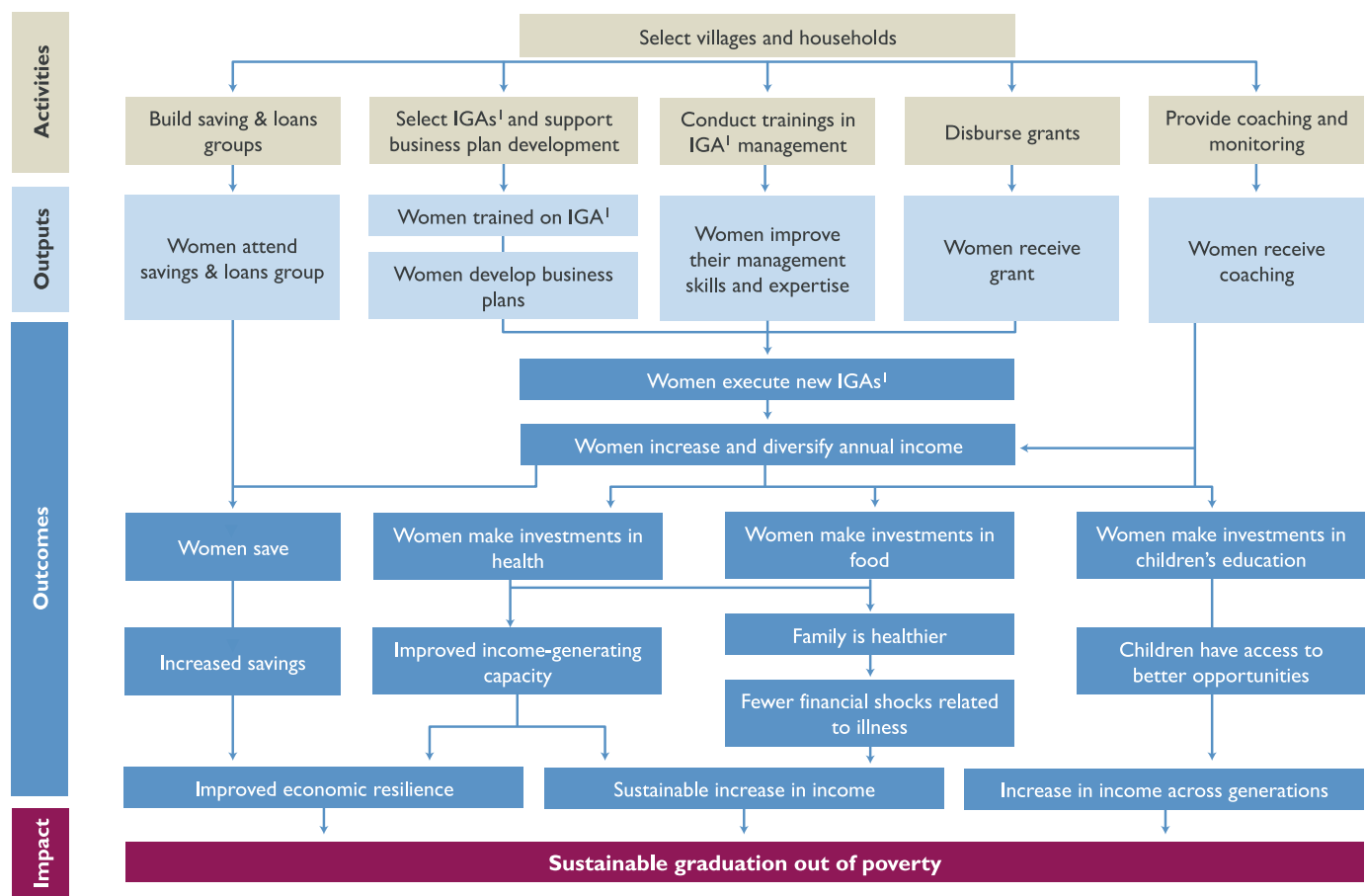
### 3.4.1 THE ROLE OF THE THEORY OF CHANGE IN ASSESSING PAYMENT METRICS

Before getting into a discussion of payment metrics, an implementer would benefit from a well-articulated theory of change. A program's theory of change illustrates the expected causal pathway from activities to the desired impact. It defines the activities, outputs, outcomes, and desired impact, and the expected causal links between them (see definitions of each of these components in Chapter 2). The assumptions underlying each of the linkages, as well as the influence of external factors, are considered when constructing the theory of change.

<sup>28</sup> The Village Enterprise Development Impact Bond (VE DIB) is currently being finalized with agreement signatures across all parties expected in September 2017 and program implementation set to begin in November 2017. For reference, see the VE DIB Design Memo, which will be publicly available Fall 2017 after program implementation commences.

An example of a theory of change for a poverty graduation program, which intends to support beneficiaries to permanently lift themselves out of poverty, is shown in Figure 9.

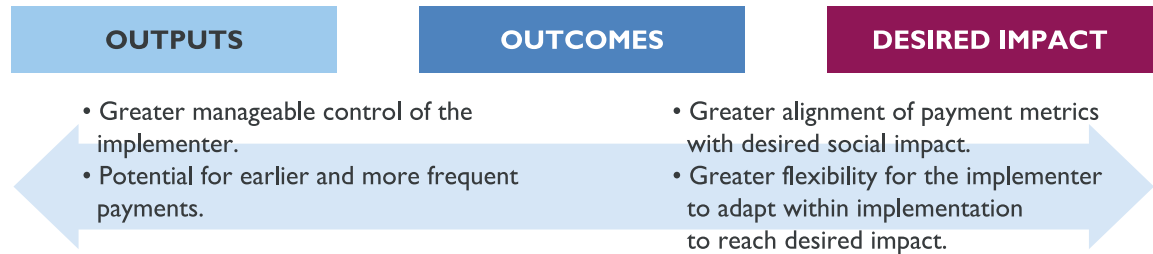
**FIGURE 9.** A stylized theory of change from a Poverty Graduation Program



Developing the theory of change helps implementers navigate discussions around which outputs and outcomes their program delivers. It helps to identify the risk associated with delivering each metric by making explicit the assumptions the metric is dependent upon and the external factors it is vulnerable to. The theory of change also helps assess how well aligned the proposed payment metrics are to the core channels of impact and whether the payment metrics allow flexibility to adapt the program's activities based on experiential learning during implementation.

The typical trade-offs between payment metrics closer and further along the results chain are presented in Figure 10.

**FIGURE 10.** Common trade-offs between payment metrics across your theory of change



### 3.4.2 SELECTING THE RIGHT PAYMENT METRICS

Payment metrics should create the right incentives and allow you the flexibility to achieve your mission and reach the desired levels of impact. To ensure the incentives are effective and the level of risk is acceptable, your organization should be confident it can influence the metrics. Below, we outline key considerations when assessing proposed payment metrics.

#### » Alignment of payment metrics with desired impact

To provide the ideal incentives, or avoid the “folly of rewarding A, while hoping for B,”<sup>29</sup> the RBF agreement should pay for the desired impact (B) directly. However, in many cases it might not be feasible to pay for the desired impact directly for one or more of the following reasons:

- **Measurability:** It may be difficult or costly to measure the desired impact objectively. For example, child mortality requires a large sample size to measure due to its relatively low frequency.
- **Time horizon:** The time required to achieve a detectable impact might be too long for the preference of the parties. For example, the impact of an early childhood development intervention on increased lifetime earnings may need to be measured decades after the program is implemented.
- **Attribution:** If the desired impact is highly sensitive to external factors (e.g. climate conditions) it may be hard to attribute observed results to the implementer’s program. Attribution is critical from the perspective of both the outcome payer and the implementer. Implementers may not be able to assume risks of underperformance arising from external factors over which they have no control. Similarly, outcome payers do not want to pay for results that would have occurred in the absence of the program, and thus want to ensure results are attributable to the implementer.

<sup>29</sup> Steven Keer, 1975 as cited by Sturgess & Cumming (2011), p.22.

Where it is not possible for payments to be tied to your desired impact, payments in an RBF agreement will usually be based on a combination of other outputs or outcomes. The payment metrics should be certain and strong contributors to your organization's desired impact individually and as a group, and should minimize potential perverse incentives. Your theory of change will help you assess the extent to which the selected metrics create incentives that support or distract from your desired impact.

### **Assess individual payment metrics**

Payment metrics are a more certain contributor to the desired impact when there is a strong causal link with the desired impact. To assess the causal link, implementers should identify the assumptions the link is dependent upon and how vulnerable it is to changes in external factors.

Generally, outcomes have a stronger causal link with the desired impact than outputs since the causal link relies on fewer assumptions and there are fewer external factors that could jeopardize outcomes translating to impact. Some outcomes might also have a stronger causal link with the desired impact than others. For example, improved student learning might have a stronger causal link to increased future income than school enrollment given that enrollment's impact on future income is dependent on the quality of education.

### **Assess the set of payment metrics**

When assessing the quality of a set of payment metrics, consider how well the metrics collectively capture the key contributors to the desired impact and the measurement costs involved. Fewer payment metrics can reduce measurement costs and help to draw attention to the most impactful elements of your intervention. However, in multi-dimensional programs with many contributors to your desired impact, or programs where there is uncertainty regarding the most impactful contributors, fewer payment metrics increase the risk of de-prioritizing important contributors (see below).

### **Assess potential perverse incentives**

Finally, assess whether the payment metrics introduce any incentives for undesirable actions. That is, do they create perverse incentives such as the following?

#### ***De-prioritization of key results in multi-dimensional programs***

Well-designed incentives ensure that implementers prioritize what matters most. Strong incentives to prioritize one aspect of a program inevitably lead to a de-prioritization of others. In multi-dimensional programs with many results of interest, there is a risk of introducing perverse incentives if the set of payment metrics only partially captures the key results. For example, in a poverty graduation program, the metrics that capture the breadth of outcomes required to sustainably lift an individual from poverty might include savings, food security, health outcomes, and education outcomes. If savings were the only payment metric, it might diminish attention to the other necessary outcomes and reduce the program's overall effectiveness.

***De-prioritization of subsets of the target population: Differences in effort required for impact***

Certain payment metrics might also create an incentive to work with a subset of your target population that may not be aligned with your organization's mission. Consider an education project in which payments are tied to percentage point increases in enrollment rates. If it takes more effort to enroll children from distant communities, the incentive might induce your organization to neglect these children in favor of children living in nearby communities. A payment function with higher prices for the subset of your target population that is hardest to impact would mitigate this potentially perverse incentive (see section 3.6).

***De-prioritization of subsets of the target population: Binary vs. continuous measures***

There are important trade-offs between using binary ("yes" or "no") and continuous metrics. For example, where there is significant heterogeneity in the target population, a binary payment metric – such as the number of participants who experienced at least a 10% increase in income – creates an incentive to target a broader range of individuals than a similar continuous metric, such as the increase in income across all participants. The continuous metric may create an incentive to increase the income of a few individuals by a great deal while the binary metric creates an incentive to work with more participants. However, a binary metric may also create incentives to neglect participants that would be able to increase their income but not reach the 10% threshold.

***Conflict with best interest of the beneficiaries***

Payment metrics can create incentives that lead to an oversupply of services or harm the target population. For example, consider an implementer that offers family therapy and support to improve parenting practices. Paying the implementer for each child that remains living with their parents may create an incentive for the implementer to discourage out-of-home care, even if it is in the best interest of the child. Often perverse incentives can be mitigated if the actor responsible for determining what is in the best interest of the beneficiary is not part of the RBF agreement; for example, by having the appropriate government agency assess whether a child should be placed in out-of-home care.

**» Payment metrics and flexibility**

As noted in Chapter 2, flexibility to adapt activities to reach the desired impact is a key benefit of RBF. This is especially important if your organization is still testing components of a program's theory of change or needs to adapt the program to changing contexts.

In such situations, some payment metrics may adversely affect your flexibility to adapt your program if they tie payment to outputs that are no longer relevant or impactful in changing contexts. This may increase your payment risk if one metric prescribes a certain path towards improving another metric. For example, consider an RBF agreement which ties payment to the number of condoms delivered and the increase in the contraceptive prevalence rate. Contraceptive preferences are highly context sensitive. During implementation, it becomes evident that this population is not using the condoms and thus the



contraceptive prevalence rate is not increasing. The implementer is placed in a challenging position where it is unable to increase both condom delivery and the contraceptive prevalence rate, and thus is unlikely to receive full payment.

Alternatively, using outputs as payment metrics can limit flexibility and discourage innovation or experimentation. For example, consider a program that aims to reduce child deaths due to diarrhea and ties payment to the provision of diarrhea treatment. Such an approach assumes that the delivery of diarrhea treatments is the most effective intervention to reduce child deaths and inhibits your organization's ability to experiment with other methods of reducing child deaths. For instance, a water, hygiene, and sanitation intervention might prove to be more impactful given the local context. For a real-world example of this type of issue see Box 7.

#### **Box 7. Payment metrics for a school voucher program PBC**

Escalera, an NGO operating in Chiapas, Mexico, delivered “mentor-in-a-box” high school prep kits to public schools in the year before high school enrollment. The kits informed a student's decision of whether to enroll in high school through videos of role models, workbooks, and classroom lesson plans. Additionally, to test whether money was also a barrier for high school enrollment, Escalera provided students from some schools with vouchers to cover enrollment fees.

During the payment metric selection exercise, the number of students that used vouchers to enroll was considered for inclusion as a payment metric because it is easy to verify and closely related to actual enrollment. However, it became clear that including vouchers as a payment metric would prescribe a specific path towards increasing high school enrollment rates and might therefore limit Escalera's ability to adapt and innovate its approach to increase high school enrollment rates and respond to changing needs of the target population.

Given Escalera's aim to adapt the intervention to the local context while scaling up the PBC, vouchers were not included as a payment metric.

Note that flexibility is most useful if accompanied by performance management that provides data on program activities and results in real-time. This can inform your organization's decisions regarding how to iterate and adapt the program to reach its desired level of impact. This will be explained in greater detail in Chapter 4.

### » Manageable control, risk, and incentives

Payment metrics should be largely within your organization's manageable control. Payment metrics over which your organization has relatively little manageable control carry greater payment risk and do not create incentives to improve impact.

Outputs, which are direct products of the activities implemented, are usually more within your manageable control than outcomes, which rely on assumptions, depend on the beneficiary's response, and are more sensitive to changes in external factors. For example:

- (1) For a community health organization, the number of trainings conducted on the importance of antenatal care (output) is more within their control than the number of women who attend antenatal visits (outcome). The number of women attending health facilities is dependent on the effectiveness of education, on the willingness of women to go to facilities, but also on the quality of facilities and distance to facilities, which are completely outside the community health workers' control.
- (2) For an agriculture extension services provider, the amount of fertilizer used (output) is more within its manageable control than improved agricultural production (outcome). Increases in agricultural production are subject to external factors outside the control of the provider, such as rainfall, and thus less within its manageable control. The influence of changes in external factors can be mitigated using an evaluation method that estimates the causal impact of your program, effectively<sup>30</sup> isolating the effect of these external factors (see section 3.5).

Evidence from your program or other programs can help you assess the degree of manageable control. Furthermore, take into account the time frame required to make progress relative to the measurement time frame and your organization's ability to mitigate the impact of external factors by adjusting your activities.

## 3.5 ASSESS THE EVALUATION AND MEASUREMENT METHOD

Paying for results requires measuring results. The process includes two decisions. First, the data on the payment metrics must be collected using the measurement method of choice. Second, the evaluation method of choice is used to statistically estimate the results that are attributable to the program and warrant payment.

The measurement method defines by whom, when, and how data is collected. The choice of the method contains many tradeoffs. Data collection by independent third parties and larger samples typically increase the confidence in the results but are also more expensive. But the high costs and reliance on third-party data collection make more reliable methods less likely to be integrated into the implementer's routine practice in the future.

<sup>30</sup> The evaluation method does not protect against external factors that influence the effectiveness of the program. For example, if a fertilizer program is more effective when rainfall is high, then a reduction in rainfall would mean that the fertilizer program would be less effective as compared to the same fertilizer program in a context with more rainfall.

The choice of evaluation method is a function of the desired rigor, of feasibility, and affordability. For instance, well-designed Randomized Control Trials (RCTs) are rigorous but expensive and sometimes not politically feasible. More rigorous methods produce less biased estimates of the results attributable to the program and thus reduce the risk of payments not reflecting the true results achieved (payment risk). They also generate stronger incentives and produce more reliable evidence on program impact. The higher costs associated with more rigorous methods, however, can limit the suitability of the RBF design as a financing mechanism in the long term.

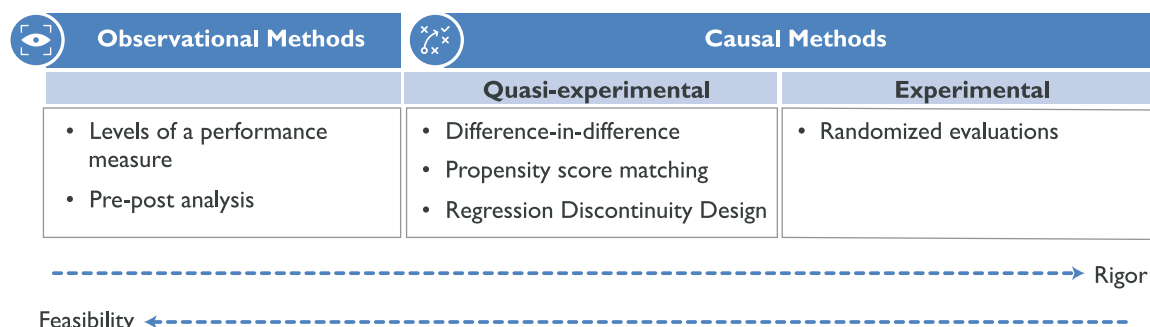
By highlighting the most important methodological tradeoffs, this section provides a starting point for how to assess different evaluation and measurement methods in the context of assessing the RBF design.


### 3.5.1 IMPLICATIONS OF EVALUATION METHODS

#### » Paying for observed versus causal results

A range of evaluation methods can be used and can broadly be categorized into two buckets, observational and causal (see Figure 11).

**FIGURE 11.** Categorization of common evaluation methods



 **Observational methods:** Consider an agricultural intervention that aims to increase harvest yields. Observational methods simply record the level of the desired outcome (e.g., the harvest yield) or the change in outcomes (e.g., the increase in harvest yield). These methods cannot claim that their estimated effect is attributable to the implementer's program. We will not be able to tell if the change in results is due to the program or due partly or entirely to external factors. For instance, a bumper harvest might be attributable to the agricultural intervention or to that year's excellent rainfall. An observational method cannot distinguish between these sources of change.

## THERE ARE TWO MAIN TYPES OF CAUSAL METHODS: EXPERIMENTAL AND QUASI-EXPERIMENTAL METHODS

**Experimental methods** use a 'lottery' (randomization) to assign people to the treatment and comparison group. This 'lottery' assignment equalizes every participant's chance of being placed in either the treatment or comparison group and ensures that differences between the groups are not systematic at the outset of the program. A common term for the experimental method is an RCT.

If a lottery cannot be conducted to assign potential beneficiaries into treatment and comparison groups, quasi-experimental methods can be considered. **Quasi-experimental methods** use statistical methods to build a comparison group that is as similar as possible to the treatment group. While these methods can control for several factors that affect results, they are generally less effective than randomized approaches at producing unbiased results.



**Causal methods:** Conversely, causal methods aim to isolate the effect of the program from other factors, such as economic conditions, demographic characteristics or other programs that might affect results. The program's impact is isolated by comparing the results for program beneficiaries (the treatment group) to the results for a similar group of people who did **not** receive the program (the comparison group). The comparison group is constructed to be as similar as possible to the treatment group in all relevant characteristics except for participation in the program. The differences in results between the groups should then be attributable to the program.

### » Which evaluation method is the gold standard for the RBF instruments?

There is no golden standard for evaluation in RBF. The key for RBF design is to understand the trade-offs of the various evaluation methods. Observational methods are generally cheaper than causal methods, and administratively and politically more feasible. As a result, RBF designs that rely on such methods are easier to scale. However, causal evaluation methods can lower payment risks, generate evidence for learning, and mitigate perverse incentives. Decisions on which method to adopt must be informed by the program's objectives and by what is feasible in the given context.

While causal evaluation methods such as RCTs are often hailed as the "gold standard" for evaluation, they are not necessarily the appropriate choice for a given RBF design. In fact, only the minority of RBF instruments are tied to results measured by RCTs. Most RBF instruments (particularly PBCs) pay for observed results. Observational methods are often favored when only a small portion of a program's funding is tied to the estimate of program impact, and the risk of a biased estimate does not justify using a more expensive causal evaluation method. Observational methods are also typically appropriate when the payment metric is an output, rather than an outcome, since the causal relationship with the program is clearer in this case.

However, in an RBF agreement with a large portion of funding tied to outcomes (such as an Impact Bond) the consequences of the evaluation's finding on payments is high. In these cases, it is more common to pay for a causal estimate of program impact because it lowers payment risk. The choice of the evaluation method is, therefore, tightly linked to the type of RBF agreement you are considering.

Appendix 3 is a primer for technically inclined readers on evaluation methods. The appendix's first table (Table 6) provides an overview of commonly used evaluation methods and discusses each method's key trade-offs in terms of their reliability. Additional threats to their reliability are outlined in Table 7. Implementers should assess if these factors might constitute a risk in their context and if so, discuss mitigation strategies with the evaluation experts. Note that the list of evaluation methods and technical discussion of their trade-offs is not exhaustive. The last section of the appendix (Box 16), discusses to what extent statistical significance of estimated program impact should matter in the context of an RBF agreement. While statistical significance is an important standard for assessing a program's impact, making payments contingent on statistically significant effects may not be a fair way of sharing risks among parties. Implementers should assess if these factors might constitute a risk in their context and if so, discuss mitigation strategies with the evaluation experts. Note that the list of evaluation methods and technical discussion of their trade-offs is not exhaustive.

Lastly, the evaluation method and potential mid-course adjustments – which are required in case of spillovers and attrition – should be pre-specified in the RBF agreement to avoid disputes. For unforeseen circumstances, dispute resolution mechanisms should also be specified contractually (see section 3.7).

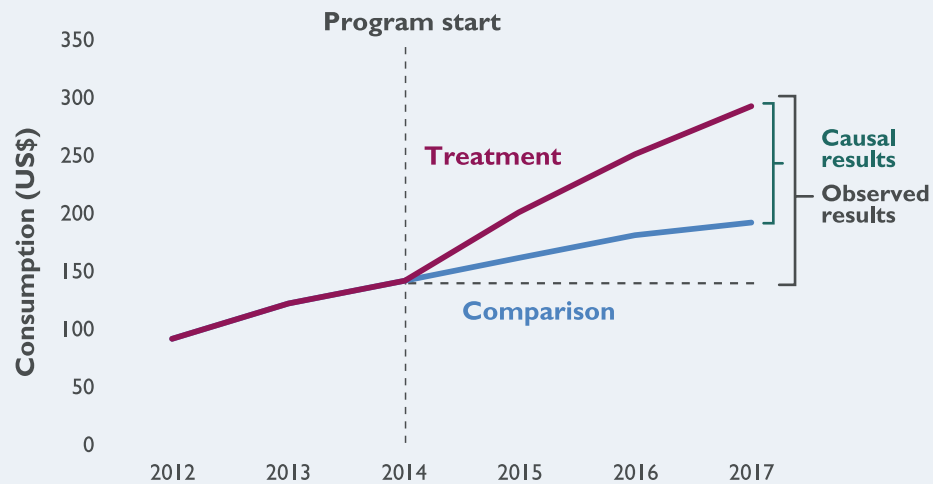
## » Benefits and challenges of increased rigor

### I. Reduced payment risks

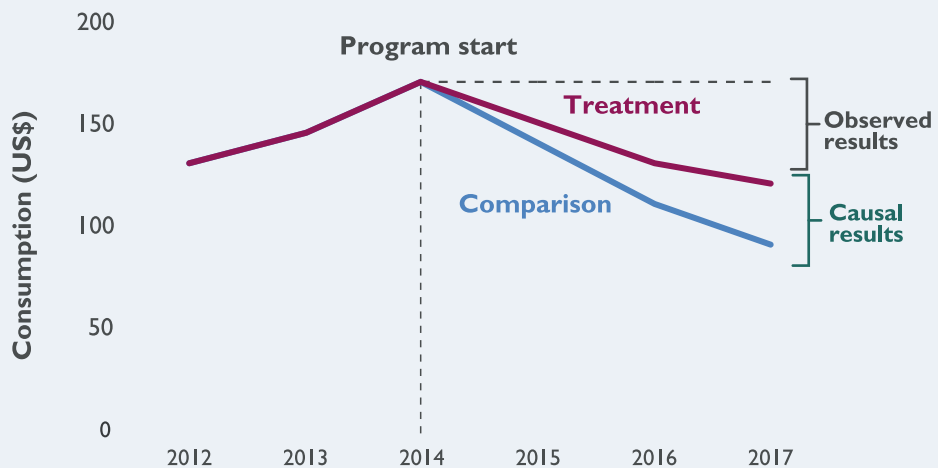
More rigorous evaluation methods can be useful for implementers, investors, and outcome payers because they reduce the risk of under- or over stating results that have been achieved. Increased rigor lowers the payment risks the parties are exposed to. Implementers and investors are protected from not being paid for the results they delivered and outcome payers are less likely to pay for results that would have occurred without the program.

To illustrate, consider the following example that compares estimates obtained through observational methods versus those from an experimental method.

At the start of a poverty graduation program in Mexico in 2014, households were randomly assigned to the treatment and comparison group. In 2017, an evaluation finds that since the start of the program in 2014, average consumptions of households participating in the program increased from \$140 to \$290 or an increase of \$150 (observed results). However, given that the average consumption of households that did not participate in the program also increase by \$50, it is likely that the improvement in average household consumption caused by the program is only \$100 (causal results). The outcome payer would only want to pay for results that can be attributed to the implementers' program (see figure, below).



Now consider the opposite case. Imagine observed results indicate that household consumption decreased after program implementation. On first sight, it appears as if the program was a failure. However, during the program, Mexico also experienced a severe drought. By measuring results causally, the implementer would find that, although household consumption decreased, those participating in the program experienced less of a decrease than those in the comparison group. The implementer would be able to prove that despite a decrease in consumption, the program itself positively impacted beneficiaries' consumption (see figure, below).





The rigor of the evaluation method will be of more concern to implementers, investors, and the outcome payers when: a) the amount of funding that is tied to the achievement of results is high; and b) results are highly sensitive to external factors or are likely to change in the absence of the program.

## 2. Reduced risk of cream-skimming

Where the potential for undertaking ‘cream-skimming’ is high, implementers and outcome payers may prefer to pay for causally estimated results to align incentives towards the desired impact.

Paying for observed results creates incentives for implementers to select those beneficiaries who are most likely to achieve the results in *absence of the program* (i.e., cream-skimming). Investing resources where results would have been achieved without the program, however, adds no value towards achieving the desired impact. Paying for results measured causally mitigates perverse incentives; since payments are only made for results that are caused by the program, financial incentives are aligned towards working with those beneficiaries whose outcomes can be truly impacted by the program.

## 3. Evidence of program effectiveness for learning

Increased rigor of the evaluation method generates valuable evidence of the effectiveness of the program, which might be an additional objective of the parties:

- For implementers, more rigorous evaluations can help improve program design by generating evidence on the program’s overall effectiveness and providing valuable insights on what program components drive effectiveness.
- For the outcome payer (or external parties) rigorous evidence may help to influence public policy and scale up programs.

The closer the payment metrics are related to the desired impact, the more valuable the evidence generated. For example, knowing that the program increased the number of students that received textbooks (i.e., an output) is valuable, but it is more valuable to know to what extent the program improved learning outcomes for students.

## 4. Increased costs and decreased feasibility and scalability

While measuring results causally is technically preferable, it is usually also costlier and at times unfeasible. It is also less suitable for scaling-up programs and long-term integration into the existing evaluation processes of the implementer or external parties.

The extra costs associated with causal methods are due to the additional measurement of the results of the comparison group. However, there may be significant differences in costs between

the various causal methods. In general, randomized evaluations are costlier than quasi-experimental methods, which can be used when suitable data for a comparison group is already available. If the necessary data to measure results is not already available, it makes less sense to use a quasi-experimental method (compared to an experimental method) since they will likely have the same costs but not the same rigor. In this case, the only reason for using a quasi-experimental method is if randomization is not feasible because of ethical or practical considerations, such as problems with maintaining the integrity of the treatment group (see Appendix 3) or because the stakeholders are not able to bear the political risks associated with not offering the program to half of the population.

Given these drawbacks, paying for results measured causally may reduce the short-term affordability and long-term replicability or scalability of the RBF design.

### 3.5.2 IMPLICATIONS OF MEASUREMENT METHODS

Often, the results that are paid for in an RBF agreement can be measured by different actors and with different data. For example, attendance rates could be measured using public school records, based on classroom spot-checks carried out by the implementer. Alternatively, it could be based on an independent evaluator's household survey. This section will briefly describe the implications of how, when, and by whom the payment metrics are measured.

#### » Types of data

There are two broad types of data used to measure results in social programs:

1. **Data reported by beneficiaries:** This includes, for example, survey responses (using a household survey, phone calls, or text messages), customer feedback forms, and citizen scorecards. Self-reported data from beneficiaries is used to measure results such as household income, changes in behavior, beneficiary knowledge or perceptions of program quality.
2. **Data from direct observation:** This includes the recording of a specific transaction, service or status, such as the attendance at a job training, enrollment in school, employment, or construction of a road, as well as assessment using technology (such as measuring hectares managed by an agroforestry system using GPS technology, measuring anthropometric indicators, or conducting blood tests).

Many metrics can be measured using either method; for example, household savings can be measured using self-reported savings or observed savings in a bank account. Other results that can be measured either using self-reported or observed data include changes in behavior, school enrollment and attendance, and employment status.

The choice of the type of data should be driven by considerations around payment risks. Self-reported data might be more prone to underreporting than observed data because, for example, beneficiaries may

underreport savings in the hope of receiving more benefits. Systematic underreporting, especially if more common among beneficiaries of the program than non-beneficiaries, may result in inaccurate evaluation and increase your payment risk. On the other hand, observed data, such as formal savings in a bank account, might result in underreporting of the results if beneficiaries save the greatest part of their additional income informally at home. Consider potential effects of self-reported and observed data, keeping in mind that the accuracy of data affects your organization's payment risk.

### » Entity responsible for data collection

The data for measuring payment metrics can be collected by an independent evaluator, an existing external actor, or by your organization. The choice of the entity will affect the costs of data collection and the ability to replicate the RBF design in the future. If the implementer takes on the data collection role you may be able to build valuable data collection capacity within your organization. Considerations for each entity are listed below. For each actor, consider how frequently measurement can occur and how accurate measurement will be, as these have implications for your organization's ability to learn and course-correct and the risk of the reported results not reflecting your actual impact, respectively.

#### 1. Independent evaluator

Contracting an independent evaluator to collect the data ensures the independence of the evaluation and may be a precondition for the outcome payer. Independent evaluators also have the highest capacity to collect accurate and credible data, reducing the payment risk for all parties.

Where data can only be collected with a low frequency or where the data collection methods (e.g., a household survey) may make frequent data collection too costly, contracting out data collection may be less costly than building your organization's internal capacity. In the opposite case, contracting an independent evaluator to do data collection might increase costs.

Where the need for an independent evaluator is determined from the outset, the frequency of measurement may be reduced for cost purposes, reducing the potential for your organization to learn what is working in your program and adapt where needed to improve results.

The use of an independent evaluator also means that data collection is likely not integrated into existing data collection systems and that your organization's capacity is not built through measurement, reducing the suitability of the RBF design as a long-term financing mechanism.

#### 2. Existing external actor

Data on results may be collected by an existing external actor that does not form part of the RBF agreement; for example, schools, health facilities, and prisons. Usually this data is already being collected by these actors, but in some cases requests can be made for additional data to be collected, particularly where the outcome payer of the RBF agreement is a government and the existing external actor is a public institution.

Using data from existing actors is often the lowest cost option and may increase replicability or long-term suitability of the RBF design. However, where data from existing actors is used, less performance management capacity is built within the implementer through measurement. Furthermore, the parties to the RBF agreement have less control over data quality and frequency of measurement, which are important for payment risks and your organization's ability to learn.

Note that an independent evaluator is often contracted to analyze data from the existing external actor (e.g., matching data to program beneficiaries) and report results.

*Prior to implementation it is important that parties confirm that data is accessible and of the required quality. For example, if teacher absenteeism is high, school records of student attendance may not be of the required quality, which might result in attendance rates not being measured accurately. Consider how quickly data can be gathered and if data quality and accessibility are likely to remain high over time. For example, a school might decide three years in the future to no longer allow access to attendance records.*

### 3. Your organization

Data collected by your organization can be used to determine payments in an RBF agreement. Your organization may already be collecting the relevant data, it may be easy to add new data to existing data collection activities, or you may need to collect data in new ways.

Collecting the data can help to build your organization's performance management capacity, which may help you to improve results included in the RBF agreement and facilitate data-driven decision-making in the long term. If you are already collecting this data, or if frequent data collection requires a permanent team on the ground, having your organization collect the data may have lower costs than contracting an independent evaluator or using data from an existing external actor. Building capacity within your organization and minimizing costs may increase the suitability of the RBF design as a long-term funding mechanism.

Given your organization's potential to benefit from overreported results, data collected by implementers in RBF agreements must always be verified. Verification is most commonly conducted by: a) repeating data collection for a random sample using the same method; or b) crosschecking data using a different source (e.g., verifying your attendance data with administrative records); or c) by doing spot checks or follow-up surveys with a representative sample or outliers (e.g., verifying your attendance data with a beneficiary survey). The RBF agreement should specify how discrepancies between your reported data and the verification will be addressed (see Box 3).

*Where your organization is responsible for data collection, ensure you have, or can build, the capacity to collect high quality data. Where possible, test the verification method on your data before implementation.*

### Box 8. Addressing discrepancies between reported and verified results

When the implementer does the primary reporting of results (which is then verified by the independent evaluator), the parties should agree on how to address any potential discrepancies between the results reported by the implementer and those reported by the independent evaluator.

First, it is common practice that parties establish a threshold for the acceptable margin of error<sup>31</sup> between reported and verified results. For example, if an implementer reports that 100 children attended school, but the independent evaluator finds that 90 children attended, the implementer's report has a margin of error of 10%.

Second, parties benefit from setting protocols and mechanisms to guide them towards resolution in the event that discrepancies exceed the margin of error. Possible mechanisms include:

- a) The results reported by the independent evaluator take precedent (i.e., payment is made according to the independent evaluator's results).
- b) Either the implementer re-measures results or the independent evaluator re-verifies results for a larger sample to ensure that the sample selection did not cause the discrepancy. In such scenarios, parties should agree upfront on who bears the cost for additional data collection and analysis.
- c) Reducing payments by discounting the results reported by the implementer proportional to the discrepancy.

<sup>31</sup> A margin of error is a range of discrepancies between what is reported and what is found through the verification, which will not influence payment.

## 3.6 ASSESS THE PAYMENT STRUCTURE

Where the payment metrics define the results that are paid for, the payment structure specifies the timing of payment and how payment varies with results achieved on the payment metrics. The payment structure is thus an important design component in defining the nature and strength of the financial incentives provided to the implementer and the payment risk transferred to the implementer (and investor(s) in case of an Impact Bond).

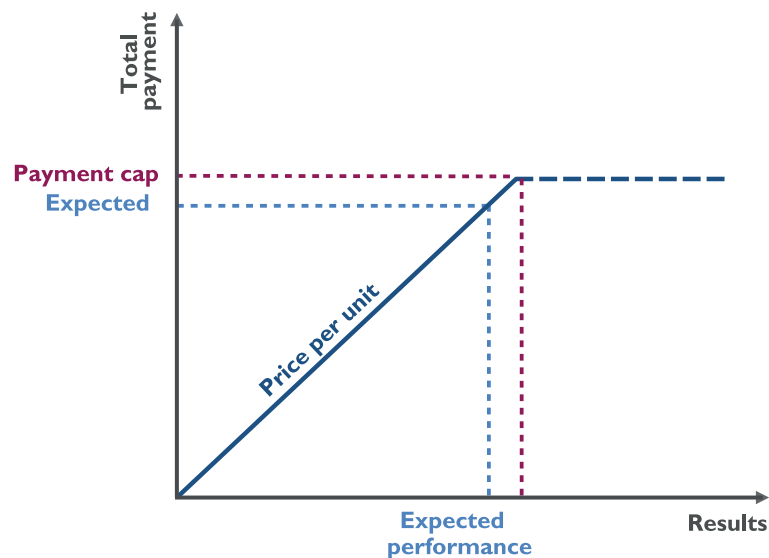
This section outlines the considerations for assessing the two key elements of the payment structure: (1) the payment function and (2) the payment schedule. The payment function determines the total payment as a function of the results achieved. The payment schedule determines the timing of the payments.

### 3.6.1 THE PAYMENT FUNCTION

The payment function should generally be kept as simple as possible so that it is easily understandable for all members of your staff. The blue line in Figure 12 shows a simple and common payment function for an RBF agreement where an increase in the results achieved translates into greater payment in a linear way. The slope of the payment function represents the price per unit of results.<sup>32</sup>

RBF agreements generally set a maximum payment to protect the outcome payer against unlimited expenditure. This payment cap, which determines the largest permissible payment, is often set to correspond to a level of results that is higher than expected, to reward extraordinary results.

**FIGURE 12.** Payment function with cap



<sup>32</sup> Where the RBF agreement has multiple payment metrics, this graphic becomes multidimensional. However, since the payment for one metric usually does not depend on the achievement of another metric (with exception for when there is an overall cap which is lower than the sum of the individual caps, see below) you can draw one payment function for each metric.



### Box 9. Background on how prices are set

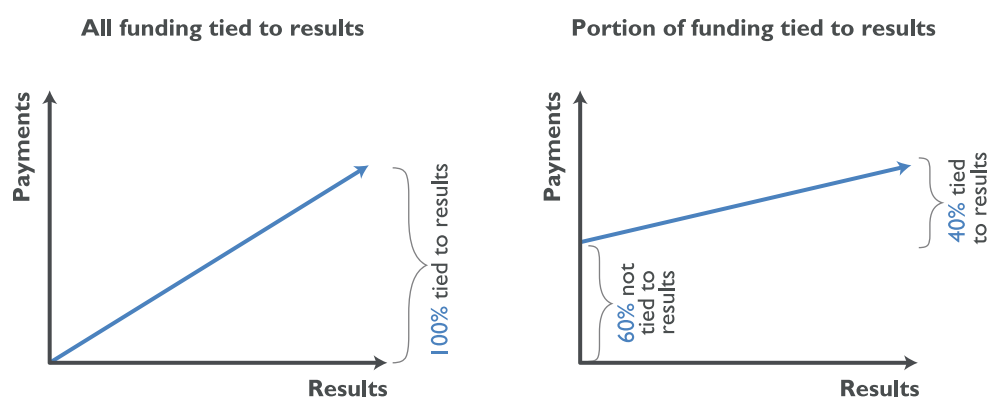
The price per unit of results is often set based on the costs of the program or the benefits it is expected to generate, or some combination of the two approaches. Where the price is set based on the value of the benefits, it may be based on the cashable benefits associated with the achievement of the results (e.g., not paying unemployment benefits or increased tax revenues from higher wages) or the quantified social value (e.g., improved well-being of society due to reduced violence). If the price per unit is based on cost, it is often a result of a cost-plus pricing approach. This approach uses 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 (risk premium).

The price per unit of results is then calculated by dividing the total cost or benefits by the total results expected.

#### » Amount of funding tied to results: incentives, flexibility, payment risk and upfront capital

The amount of funding tied to results is the portion of the total funding provided under the RBF agreement that is conditional on the achievement of pre-defined results. For example, a Performance-Based Contract (PBC) might provide 100% of funding contingent upon the achievement of results (see Figure 13, left) or 60% of funding upon contract signing and the other 40% contingent upon the achievement of results (see Figure 13, right).

**FIGURE 13.** Funding tied to results



The portion of the total funding that is contingent upon the achievement of results affects the strength of the incentives provided and the flexibility your organization has to adjust activities. It also influences the payment risk transferred to your organization and the upfront capital needed for program implementation. The following table outlines these considerations.

**TABLE 3.** Considerations to determine the amount of funding tied to results

Implications of more funding tied to results		Considerations to determine the optimal amount of funding
<b>Strength of incentives</b>	<p><b>Provides stronger incentives to achieve results, but may magnify perverse incentives created by the payment metrics or other design components.</b></p> <p>For example, consider a poverty graduation program with beneficiaries' savings as the sole payment metric. A small bonus payment may not distract your organization from paying attention to other important results such as food security, but a high amount of funding might increase internal pressures to modify the program in ways that increase beneficiaries' savings but do not facilitate a sustainable graduation to higher income.</p>	<p><b>The strength of the financial incentives your organization can productively respond to.</b></p> <p>This includes analyzing potential perverse incentives created by the RBF design and considering your organization's existing incentive environment, including how incentives will be perceived by your staff (e.g., as a reward for good performance or a sanction for not delivering results), their intrinsic motivation, and the capacity of your organization to achieve greater results in response to incentives.</p>
<b>Flexibility</b>	<p><b>Likely increases your organization's flexibility to adjust activities in the pursuit of results.</b></p> <p>More funding tied to results implies less funding contingent upon the completion of the pre-specified activities required under most procurement rules.</p>	<p><b>The flexibility your organization needs and whether the funder would be willing to allow your organization flexibility if the funding is not tied to results.</b></p> <p>Flexibility is especially important if your organization is still testing components of a program's theory of change or needs to adapt it to changing context to improve results.</p>
<b>Payment risk</b>	<p><b>Increases the payment risk transferred to your organization.</b></p> <p>For example, with a PBC where 90% of the total funding is tied to the achievement of a 10 percentage point increase in beneficiaries' income, the organization is exposed to a larger potential financial loss than with a PBC that conditions only 20% of the funding upon the achievement of this target.</p>	<p><b>Your confidence in your organization's capacity to deliver the expected results and ability to absorb a financial loss.</b></p> <p>More funding tied to results is less problematic if you are confident that you can achieve the expected results. See Box 11 below for guidance on how to assess your future performance using past evidence.</p>
<b>Upfront capital</b>	<p><b>Increases the upfront capital your organization needs to implement the program and generate results.</b></p> <p>For example, a PBC where 80% of the funding is tied to results that are measured at the end of the program will require your organization to pre-finance around 80% of the costs of program implementation.</p>	<p><b>Your organization's ability to pre-finance a portion of the program or access other financial funds.</b></p>

The optimal amount of funding tied to results strikes a balance between the potential benefits of providing effective incentives and flexibility, and the costs of potentially magnifying perverse incentives and increasing financing costs for the implementer and outcome payer. As such, the amount of funding tied to results should be the minimum amount that will provide effective incentives for progress on the payment metrics and the necessary flexibility for the implementer.

### Box 10. Implementer incentives in an Impact Bond

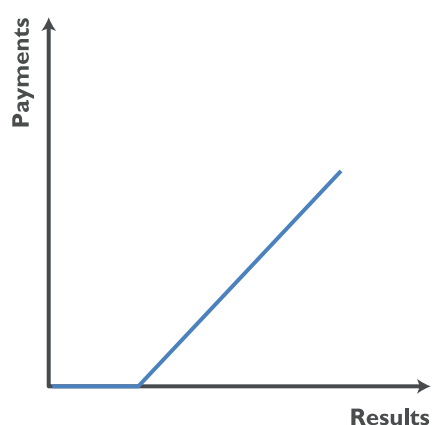
In Impact Bonds, generally all funding is tied to results; however, because investors are the principal bearers of the payment risk, the amount of funding tied to results has far fewer direct consequences for the payment risk of implementers than in a PBC. Implementers still face a reputational risk if results are not achieved. Investors may also pass on financial incentives to the implementer by providing a bonus for the achievement of results or transferring some of the payment risk:

1. **Bonus payments:** The investor may pay the implementer a bonus in the case of achieving or exceeding expected results. For example, in the Educate Girls DIB in India, the investor, UBS Optimus Foundation, will pass on 32% of any payments it receives beyond its initial investment to Educate Girls.<sup>33</sup>
2. **Risk transfer:** Investors could provide only a portion of the total program funding needed to the implementer, requiring the implementer to invest their own funds to implement the program. For example, in the Benevolent Society Social Benefit Bond in Australia, the implementer used its funds to invest in the SIB using the same investment mechanism as other investors<sup>34</sup> and in the Massachusetts Criminal Justice SIB in the United States, the implementer deferred its fees for the implementation of the program and stand to gain success fees.<sup>35</sup>

### » Thresholds: Strong incentives to improve results to reach the threshold

Payment functions may also include a threshold, which is the minimum level of results that an implementer needs to achieve before any payment is made (see Figure 14). Outcome payers may want this feature because they would not want to spend money on a program that does not achieve a minimum level of impact.

**FIGURE 14.** Minimum threshold



To illustrate, imagine an example where payments are not made until 100 units of the payment metric have been achieved. In other words, whether your organization achieves 0, 50, or 99 units does not affect your payments. They remain at zero until 100 is reached.

Thresholds provide strong incentives to improve results, but may reduce staff motivation if set too high and are perceived to be unachievable. Thresholds also introduce a higher risk for implementers since progress below the threshold is not rewarded.

<sup>33</sup> Instiglio (2015).

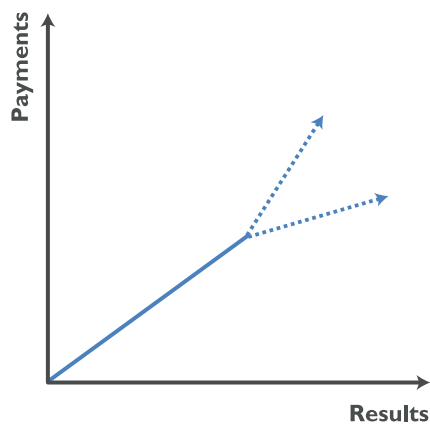
<sup>34</sup> Gustafsson-Wright, Gardiner, and Putcha (2015).

<sup>35</sup> Third Sector Capital Partners (2015).

### » Payment kinks: Adjusting for changing marginal costs

Payment functions may also include payment kinks, where the price per unit of results changes after a specified level of results has been achieved (see Figure 15). Kinks in the payment function reflect variation in the price per unit of results achieved based on the level of results achieved. Prices may increase or decrease as more results are achieved and a payment function may have one or multiple kinks. For example, the price per unit of the payment metric may be different between 0 and 50 units and between 50 and 100 units.

**FIGURE 15.** Payment kinks



An increase in prices for higher levels of results can be a way to compensate for an increase in marginal costs. Often, there are individuals within your target population for whom it is costlier to achieve results. If an implementer targets those individuals easiest to impact first and then moves on to the harder to impact individuals, marginal costs will naturally increase. An increase in prices after a certain level of results compensates your organization for these additional costs. Last-mile problems also justify an upward kink. For example, in a program treating a population for tuberculosis, a portion of the beneficiaries are likely to not comply with the treatment regimen and develop multi-drug resistant or extremely drug resistant tuberculosis and require treatment that is 20 to 1,000 times costlier than the standard treatment.<sup>36</sup> Therefore, average treatment costs would increase over time, and a payment kink could compensate for this.

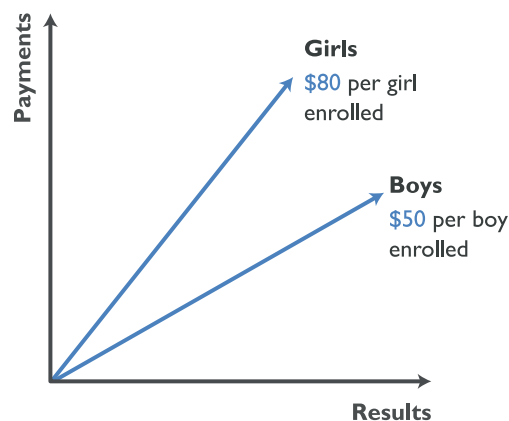
Conversely, economies of scale might mean that marginal costs decline as you expand your program. In that case the outcome payer might argue that the payment function should have a downward kink. However, you should be wary of such claims, because, as mentioned above, there is often a point at which marginal costs of delivering results increase as you expand your program's coverage.

<sup>36</sup> World Health Organization (2011).

## » Differential prices: Drawing greater attention to a subset of your target population

Prices may also vary for different subsets of your target population (see Figure 16). For example, the outcome payer may pay more for each girl enrolled in school than for each boy enrolled.

**FIGURE 16.** Differential prices



Differential prices are used for two reasons. First, it might be costlier to achieve results with one subset of the target population. For example, in an RBF agreement for maternal and child health in Chad, an “equity bonus” has been added to prices for facilities that are located in remote areas because it is costlier to reach remote regions.<sup>37</sup> Differential prices have also been used in RBF with workforce development programs around the world and have limited (but not eliminated) the incentive to focus on those who are easiest to employ. Where the differential prices were less effective, it was often because stakeholders found it challenging to set price variation accurately to achieve the intended effect.<sup>38</sup> Second, results for one subset are valued more by society or the outcomes payer. For example, even where costs for enrolling girls and boys are the same, the outcome payer or your organization may value enrolling girls more than boys for equity reasons or if evidence suggests that the relative impact of enrolling girls on life-time earnings or on future generations is greater.

**To assess the usefulness of kinks and differential prices to your program, address the following questions:**

1. How do prices compare with your cost structure? Are costs per unit increasing or decreasing as coverage expands?
2. Do costs vary for a subset of your target population?
3. Are results for a subset of the target population valued more by society or the outcome payer?
4. Do you have the necessary information on population characteristics and marginal costs or benefits to calibrate kinks and differential prices or will those features simply add complexity and uncertainty in the RBF agreement?
5. Do the proposed prices create incentives that are well-aligned with your organization’s mission?

Given the added complexity of kinks and differential pricing, implementers should assess if it is worthwhile to negotiate kinks and differential prices with the outcome payer.

<sup>37</sup> World Bank (2014).

<sup>38</sup> See, for example, Department for Work and Pensions (2014).

» **Relative prices of the payment metrics: Strong incentives to deliver one metric may lead to a de-prioritization of other key metrics**

Assuming that the payment metrics are well-aligned with your organization's desired impact (see section 3.4), the relative price of these metrics should approximately reflect their relative costs to maintain the alignment of incentives.<sup>39</sup>

To illustrate, consider an RBF agreement with two payment metrics equally important for achieving the desired impact. Both metrics have the same price per unit achieved (e.g., \$50), but it might cost the implementer \$30 to deliver one unit of metric A and \$70 to deliver one unit of metric B. The relative price of the metrics does not reflect the relative effort required to achieve them and your organization might face a financial incentive to focus exclusively on achieving the metric with the greater price per cost ratio (metric A), which may undermine the achievement of the desired impact.

» **Individual payment caps reduce the ability to pool payment risk across metrics but can align incentives with the desired impact**

RBF agreements generally have an overall payment cap to protect the outcome payer against unlimited payments. Where the RBF agreement has more than one payment metric, payment caps can also be set to cap the payment for each metric. Implementers should analyze the implications of these individual payment caps on their incentives and the payment risk.

For example, consider an RBF agreement with payment conditioned upon improvements in numeracy test scores (metric A) and literacy test scores. To limit its financial risk, the outcome payer caps the total payment tied to these metrics at \$100,000. No individual payment cap on either of the metrics is set. Theoretically the implementer can receive the full payment by improving only numeracy test scores or literacy test scores allowing it to pool the risk of not achieving results across the metrics. Not achieving the expected results on one metric can be compensated by overachieving results on the other metric. However, if both improvements in numeracy and literacy test scores are important to reach the desired impact (e.g., progression to secondary school), it might also create perverse incentives to prioritize the metric that is easier to achieve.

As discussed before, assigning a higher price to the metric that is more difficult to achieve is one way to mitigate perverse incentives. Including individual payment caps might be another way to align incentives with the desired impact. For example, including an individual payment cap of \$50,000 for each metric would still limit the overall payment but would incentivize the implementer to focus on both metrics. To allow for

<sup>39</sup> Where payment metrics are sequential (e.g., being retained in job training is necessary to be placed in a job), note that the relative price of the second metric (job placement) can be equal or higher than the required relative effort to ensure incentives are aligned with the desired social impact. For example, it might cost the implementer \$30 per person to conduct the job training and \$10 per person to place this person in a job. As long as the relative price of metric 2 is equal or greater than a quarter of the total price per person (e.g., \$25 per person who completes the training, \$15 per person who is placed in a job), incentives are well-aligned with the desired impact as the implementer could not operate profitably focusing just on providing the job training. If the price of metric 2 is however lower (e.g., \$35 per person that completes the training and \$5 per person that is placed in a job) the implementer has an incentive to focus on training people and not deliver the additional activities to place them into a job.



some additional flexibility, individual payment caps could also be set at \$60,000 while maintaining the overall payment cap at \$100,000. That allows the implementer to make reasonable trade-offs between the metrics.

Where payment metrics independently contribute to the desired impact, individual payment caps might not limit, but rather introduce, perverse incentives. For example, consider an RBF agreement with payment metrics for women delivering babies in health facilities and children receiving malaria treatments, both of which contribute to the desired impact of reducing child mortality. Including an individual payment cap on each metric might force the implementer to deliver on both, even if during program implementation it becomes apparent that treatments for malaria are a more impactful channel to reducing child mortality than in-facility deliveries.

### » Expected level of performance determines payment risks

As described above, the payment function determines the payments your organization receives based on the results achieved. In a scenario of expected performance, the payments should at least cover the costs of implementing the program and provide a premium for the payment risk assumed. Before entering an RBF agreement, your organization should therefore carefully estimate the results it expects to deliver and assess whether the risk of underpayment is acceptable to the organization.

Past evidence of your program's performance can help to inform your organization's estimate of future expected performance. However, performance estimates can be sensitive to changes in the program design or implementation, in the policy environment and general context, and in the target population and evaluation methodology, among other factors. Below we outline a 5-step process for estimating your expected performance in a way that accounts for changes in such relevant factors.<sup>40</sup>

<sup>40</sup> Academics often refer to the internal validity and external validity of evidence. Internal validity of evidence is whether an evaluation method accurately estimated the results that are attributable to the program. External validity of evidence is whether the estimated results are generalizable to a different context or time period. In using past evidence to estimate future performance, internal validity of the evidence is only important insofar as it differs between the two evaluations. Rather than considering internal and external validity separately, we do not explicitly distinguish between the two in this section. Instead we simply ask the reader to reflect on the effects of differences in target populations, context, measurement, and evaluation methods on the replicability of their previous findings.

### Box 11. 5-step process for estimating your expected performance

**1. Identify past evidence.** Identify any information your organization has on results relevant to the payment metrics in the RBF agreement. This may include, for example, outputs your organization has delivered, pre-post data on key results, or quasi-experimental or experimental evaluations (see section 3.5 for more information on evaluation methods).

**2. Identify relevant differences between the past evidence and the RBF agreement.** Consider how the baseline levels of the metric, target population, context, program characteristics, results measurement, and evaluation methods differ between the past evidence and the RBF agreement in ways that might be relevant for the effectiveness of your program. Your organization's theory of change, and its implicit assumptions, can help determine what is relevant to your program's effectiveness (see section 3.4 for more information on theories of change). The table below provides a list of potentially relevant differences.

**TABLE 4.** Potential differences affecting performance

Category	Potentially relevant differences	Example
<b>Baseline levels of the payment metric</b>	Baseline levels of the payment metric	If the payment metric in the RBF agreement is malaria incidence, baseline incidence levels affect your organization's ability to impact the percentage point reduction in incidence. For example, a 10 percentage point reduction in the incidence of malaria might be possible where incidence levels are high at baseline (e.g., 50%) but might be not realistic where incidence is lower (e.g., 30%).
<b>Target population</b>	<ul style="list-style-type: none"> <li>• Number of individuals</li> <li>• Ages</li> <li>• Gender</li> <li>• Birth or death rate</li> <li>• Household size</li> <li>• Health</li> <li>• Education</li> <li>• Occupations</li> <li>• Income</li> <li>• Social norms and culture</li> <li>• Preferences and motivations</li> </ul>	The past evidence may have measured impact on learning among a poor population, while the RBF agreement measures impact on learning among an ultra-poor population. If ultra-poor households are not able to support their children as much as the poor, this might reduce learning outcomes.

Category	Potentially relevant differences	Example
<b>Context</b>	<ul style="list-style-type: none"> <li>• Employment opportunities</li> <li>• Population density</li> <li>• Economic and political conditions</li> <li>• Policy environment</li> <li>• Accessibility and quality of health and education services</li> <li>• Presence of other social programs (particularly any complementary programs your performance is dependent upon)</li> <li>• Infrastructure</li> <li>• Climatic conditions</li> </ul>	The past evidence may have been collected in a region with more employment opportunities, which is relevant to your organization's ability to place individuals in jobs.
<b>Program</b>	<ul style="list-style-type: none"> <li>• Duration</li> <li>• Components</li> <li>• Scale of implementation</li> <li>• Staff capacity</li> </ul>	The past evidence may have been collected before your program was improved in a relevant way or, if your program was implemented at smaller scale when the past evidence was collected, program quality may be lower in the RBF agreement as a result of more thinly spread management.
<b>Results measurement</b>	<ul style="list-style-type: none"> <li>• Metric definition</li> <li>• Data source and data collection method</li> </ul>	If the metric was measured using your implementation data in the past evidence and will be measured using administrative data in the RBF agreement, this is likely to be relevant to your expected performance.
<b>Evaluation</b>	<ul style="list-style-type: none"> <li>• Evaluation method</li> <li>• Treats to causal identification (see section 3.5 for more information on evaluation methods)</li> </ul>	If a randomized control trial was used to estimate your impact in the past but the RBF agreement uses a pre-post evaluation to estimate your impact, external factors that were controlled for in the past evaluation may affect the estimate of your performance in the RBF agreement (see section 3.5).

**3. Adjust estimates of past evidence.** For each relevant difference between your past evidence and the RBF agreement, adjust the results achieved in your past evidence up or down. For example, your past impact evaluation may have indicated that your program can reduce severe malnutrition among children under 5 by 10%. However, the quality of hospitals, which are critical to your program's success, is lower in the context of the RBF agreement, and thus you adjust your estimate of expected performance to 5%. Adjustments that lower your expected performance are particularly important to pay attention to.

It is quite a challenge to accurately predict the change in expected results. There is an array of resources that may be useful in the assessment, including relevant literature and analysis of subsets of your data, where this data is available. For example, within the sample of an impact evaluation, there may have been one district where hospital quality was similar to that of the context of the RBF agreement, and data from this district can be used for the assessment.

If your impact estimates are highly dependent on external factors, you might also want to consider potential mitigation strategies in implementation. For example, if you monitor the prices of agricultural inputs in your performance management system, you could adjust your agricultural intervention accordingly if prices of fertilizers rise unexpectedly.

**4. Assess your certainty in achieving the expected performance.** Consider the following:

a. How confident are you in your original estimate? As stated in section 3.5, a smaller sample size of the evaluation reduces the confidence in the estimated results, that is, it increases the confidence interval. Consider how likely it is that your past results were overstated. For example, consider a scenario where your past impact evaluation found a 10% impact with a 95% confidence interval of 5% to 15%. In other words, with 95% certainty your program had an impact anywhere within that range. The larger the range, the lower your confidence in the estimated results. Assess whether the payments that correspond to the upper and (especially) lower bounds of the confidence interval are acceptable to your organization.

b. How confident are you that you made all the relevant adjustments to the past evidence and were able to adjust accurately? For example, you may have little certainty around your program's performance with a different target population.

Scenario analysis is useful in assessing the potential variation in your expected payments. What are the key factors your performance is dependent upon, and how would your performance differ if these factors vary? What are the best-case and worst-case scenarios, given the prior considerations? Keep in mind that overperformance might not result in higher payments, because of payment caps.

**5. Assess whether expected payments are greater than expected costs.** If your expected payments do not cover costs of implementing the program or if there is a lot of uncertainty, consider negotiating with outcome payers to adjust the payment metrics, target population, context, or measurement methods; reduce the performance targets; compensate your organization to carry out mitigation plans for risks that can be mitigated; reduce the amount of funding tied to results or include a higher risk premium for assuming payment risks associated with factors that are outside your manageable control; or add contract clauses that protect you from particular risks (see section 3.7).

### 3.6.2 THE PAYMENT SCHEDULE

The payment schedule defines when payments will be made. For example, payments may be made monthly, yearly, or once at the end of the contract. Most commonly, payments are either made with the same frequency as payment metrics are measured (see section 3.5.2) or at the end of the RBF agreement. Sometimes the measurement frequency is determined by the desirable payment frequency. The primary advantages and disadvantages of frequent payments and frequent measurement for an implementer are summarized in Table 5.

**TABLE 5.** Advantages and disadvantages of more frequent RBF payments and measurement

	Advantages of more frequent payments	Disadvantages of more frequent payments
Frequent payments	<b>Financing cost:</b> More frequent payments allow implementers and investors to recycle the upfront capital invested to implement the program earlier, reducing the financing cost.	<b>Payment transaction cost:</b> A higher payment frequency requires the attention and resources of all parties involved.
Frequent measurement	<b>Learning:</b> Frequent measurement can generate valuable insights regarding how to improve results in latter periods. It can therefore can reduce payment risk.	<b>Measurement cost:</b> Each additional measurement (data collection and analysis) has an associated cost for the outcome payer and may add operational complexity for the implementer.

Note that outcome payers may also have preferences with regard to the frequency of payments. In some cases, frequent payments, for example annual payments, are administratively easier for their organization because it resembles their status quo practice. In other cases, the outcome payer may prefer later payments.

## 3.7 ASSESS THE RBF CONTRACT FEATURES

When assessing your organization's contract within the RBF agreement, first assess whether the contract accurately reflects the RBF design decisions, including the definition of the target population, the payment metrics, and the details of the evaluation and measurement methods. Next, there are general considerations regarding the contract and risk management that you should assess. This section provides a non-exhaustive outline of these latter contracting considerations as they relate to RBF.

### » General contracting considerations

Important contract details to consider in negotiations which might affect your organization's ability to achieve results include:

**Prescribed activities.** Implementers should assess if the contract includes prescriptions requiring your organization to undertake certain activities or use certain inputs. This may limit your organization's flexibility to adjust your program in the pursuit of results.

**Clear definition of roles, particularly with regards to results reporting procedures.** The contract should clearly lay out the roles and responsibilities of all parties in terms of service delivery, results measurement, reporting, and payment. There should be a clear plan for the measurement and reporting of results, including when they will be evaluated and by whom, to whom they will be reported and what formal documentation will be created.

### » Managing project risks and disputes

The RBF contract should clearly identify key risks associated with factors beyond your organization's control that may impact the achievement of results and, to the extent possible, stipulate a process to address these risks.

#### I. Risk identification

Risks which might have a critical impact on the project include:

- **Evaluation and measurement risks** which affect the verification of results. This includes problems with baseline data, issues with subsequent data collection or measurement, delays and evaluator errors.
- **Non-compliance risks**, including failure to act within laws and regulations or delays in meeting obligations specified in the contract.
- **External factor risks**, such as:
  - Changes in **government policy** or **actions of other agents** which could significantly affect results.
  - **Force majeure** events that significantly affect the achievement of results or performance of any party due to factors beyond their control, such as natural disasters or political upheaval.



- **Currency and inflation risks.** As with any contract involving international transactions and being executed over a period of years, parties should define exchange rate and inflation corrections based on appropriate financial advice or, in the case of exchange rate risks, avoid them using financial products.

## 2. Risk sharing and mitigation

The contract should also establish a clear **process for handling these risks**. This includes considerations such as who monitors what risk, who informs others, and who is responsible for mitigation. As far as possible, decisions on **how risks are shared** among stakeholders should be made at the project's outset and addressed in the contract. Usually risks that are not within the implementers control should not be assumed by the implementer. The risks that your organization will take on will influence the attractiveness of the RBF agreement.

## 3. Decision and dispute resolution mechanisms

Not all risks can be anticipated or specified in the contract and may require decision or dispute resolution processes which should themselves be specified in the contract.

The contract should provide guidelines for risk mitigation, such as: (1) ensuring that all parties are alerted when risks emerge; (2) requiring a decision to be made regarding whether to accept, avoid, control, transfer, or monitor the risk; and (3) establishing a process to identify the party responsible for risk mitigation and resolution. As far as possible, mitigation approaches should be based on consensus-based decision-making.

Allowing for the prospect that certain critical risks cannot be sufficiently mitigated, contracts should also include mechanisms to help with dispute resolutions. Contractual dispute resolution mechanisms may entail nominating a mediator in the contract or stipulating other dispute resolution processes agreed on by all parties (see example below).

### Box 12. Example of a resolution mechanism - Educate Girls

**Evaluation Dispute:** The parties – the outcome payer, investor, and implementer – may, at any time, dispute the evaluation procedure/outcomes in case of a manifest error or failing in regards to:

- Data collection instruments, survey, and methodology;
- Calculation of the Outcomes;
- Data analysis, findings, and interpretation; or
- Sharing methodology, data, and findings outside of the parties.

**Resolution Mechanism:**

- Establishment of a Panel, consisting of three people with expertise. The Panel is jointly appointed by the parties or – in case no agreement is reached – each party appoints one (independent) panelist.
- Each of the parties, plus the independent evaluator, will deliver a statement to the Panel, summarizing their arguments.
- The Panel will review the statements and may send follow-up questions.
- The Panel is then required to reach a majority decision with respect to the dispute within 60 days.

#### 4. Termination

It may be necessary to terminate the contract if the risk cannot be mitigated and parties decide that early cessation of the program is best, or if parties have contractual early termination rights in case of specific events. For example, Impact Bonds often include an early termination right for the investor in case of material underperformance by the implementer. The contract should account for this prospect and specify the timelines, responsibilities, and process for termination - including the communication protocol between the stakeholders and externally, whether the implementer will be compensated for progress to date, and how stakeholders can ensure that the implementer can continue delivering services as seamlessly as possible.



## 3.8 PREPARE FOR NEGOTIATIONS

During negotiations with outcome payers, the implementers' main objectives are to ensure the RBF agreement is aligned with their desired impact and carries a fair and manageable level of payment risk. We have included a checklist summarizing the key considerations covered in this chapter in Appendix 4, which will help you keep track of how the alignment of the agreement to your mission and payment risk evolves through negotiations.

In addition, in our experience participating in RBF negotiations, implementers will benefit from a clear understanding of:

1. their performance and its implications;
2. the outcome payer's objectives and preferences;
3. helpful negotiations practices.

This section provides recommendations regarding each of these three topics.

### » Understanding your performance and its implications

During negotiations, outcome payers may request modifications to the proposed RBF design. They may request to modify, for instance, the basket of payment metrics, the target levels of performance on certain payment metrics, the target population, or the target geography.

In this process, it is important to understand how each modification affects the risk of not meeting performance targets and therefore your payment risks. In our experience, two measures may help you navigate these negotiations:

1. It is useful to have access to analytical backstopping from a member of your team – the M&E team usually – who is intimately familiar with your evidence. This person should be able to analyze the existing evidence and consult the project teams to provide predictions of performance on new metrics, in new geographies, or for new populations (see section 3.6).
2. It is also useful to have an easy-to-use financial model that links the performance predictions to payments. This allows the implementer to quickly translate requested modifications into financial implications and use this information to support your position in negotiations.

### » Understanding outcome payer's objectives and preferences

The negotiations with the outcome payer should be informed by an understanding of their main objectives and preferences. Information on the outcome payer's preferences will be revealed during the negotiations, but implementers should seek to understand them as much as possible beforehand. It is particularly important to distinguish non-negotiable preferences from negotiable ones.

For instance, in a recent negotiation in an education RBF, an outcome payer revealed that they strongly preferred the use of an RCT for the evaluation. Because the implementer was seeking a more scalable RBF

design they strongly preferred a pre-post methodology. The negotiation teams from both parties reached a tense deadlock after weeks of back-and-forth. Digging deeper, we discovered that the outcome payer cared about a rigorous attribution of literacy outcomes, but did not feel strongly about school enrollment outcomes. Conversely, the implementer was mostly worried about its enrollment outcomes. This deep understanding of the outcome payer's interest helped design a creative adjustment: the implementer was able to negotiate a pre-post evaluation on enrollment and agree to an RCT over the literacy outcomes.

Understanding the outcome payer's preferences can inform your prioritization of requests and facilitate the negotiation of agreements which meet both parties' objectives.

### » Practices for RBF negotiations

While we recommend drawing from your experience and the negotiations literature,<sup>41</sup> we wanted to share a few context-specific practices that in our experience can lead to better outcomes in RBF negotiations:

1. **Purpose-driven negotiations:** While it is inevitable that the negotiation will evolve around the interests of various stakeholders, we recommend firmly establishing that the shared goal of all parties is to find an agreement that maximizes the program's impact on beneficiaries. This is typically to all parties' advantage. For your organization, this ensures that the design of the RBF agreement is aligned with your mission. For example, consider a provider implementing a community health intervention. The outcome payer may request to add tuberculosis incidence to the set of payment metrics, but the implementer does not have experience in addressing tuberculosis. Instead of highlighting their increased payment risk, the implementer could argue that this change might distract their attention from the elements of their program that are known to be highly impactful.
2. **Negotiations are seldom a zero-sum game.** This means that collaborative and open negotiations, where each party tries to understand and flexibly respond to the needs of the other, can create greater value for all parties. For instance, as there are different ways of managing risks, parties should creatively explore a range of solutions to find one that meets both parties' needs. For example, if the outcome payer has strong interest in a payment metric which poses high risk for the implementer, this can be managed by negotiating a lower target, a lower amount of funding allocated to this metric, or additional funding for capacity building.
3. **All-at-once versus piecemeal:** It is important to consider the outcome payer's requests all-at-once, as opposed to negotiating specific requests or design features. This will allow you to assess the overall impact on the attractiveness of the RBF agreement, prioritize your proposed changes, and prepare a strong counter-offer. Unidimensional negotiations can lead you to agree to a series of concessions that make sense individually, but together significantly reduce the attractiveness of the RBF agreement.
4. **Getting the right expertise.** Negotiating an RBF design is complicated and implementers may consider seeking advice from a technical assistance provider.

<sup>41</sup> For example, see Wheeler (2000); Lax, Whitener & Sebenius (2007).

## 3.9 KEY TAKEAWAYS

### RBF design and contract components

- The RBF design is critical for your success because it determines the focus, incentives, and risks for the implementation of the program. The design centers on several components outlined below.
- Good contracting is critical to adequately reflect the terms of the agreement, protect parties against risks that are outside their manageable control, and specify protocols for addressing a variety of scenarios.

### Guiding questions for assessing the RBF agreement

- There are two primary considerations when assessing the design of an RBF agreement:
  - Alignment with your organization's desired social impact. The RBF agreement should provide the incentives and space for your organization to prioritize what matters to achieve the desired social impact.
  - Payment risks or the likelihood that you will not receive payment conditioned on results. The RBF agreement should transfer a level of risk that is adequate for your organization's capacity.

### Financial size, duration, and target population

- There are tradeoffs when deciding on the overall financial size of the project. Given the additional demands RBF creates for leaders and organizations, we recommend starting with RBF projects that are small enough to be well staffed and allow for timely adaptation of the delivery model.
- RBF agreements typically have a longer duration to allow for learning, adaptation, and measurement.
- The target population should be defined based on the implementer's desired impact and their experience and track record of achieving results with a particular target population.

### Payment metrics

- The payment metrics are the results that are paid for in RBF agreements. They are the source of many of the incentives that your organization will experience, and thus will strongly shape what your teams do on the ground.
- Developing a clear theory of change will help you assess the payment risk transferred to your organization and how well the proposed payment metrics are aligned with your desired impact.
- An RBF agreement that pays for your organization's desired impact would create the ideal incentive environment for your organization to maximize impact. However, this is often infeasible due to challenges around measurement, the time horizon, and attribution of results. Further, it usually transfers a higher payment risk to your organization.



- Where it is not possible for payments to be tied to your desired impact, payments in an RBF agreement will usually be based on a combination of other outputs or outcomes. You should always ensure that the selected payment metrics are strong contributors to your organization's desired impact (both individually and as a group) and minimize potential perverse incentives.
  - Metrics should be a strong contributor to the desired impact and their causal link with the desired impact should rely on minimal assumptions and have minimal vulnerability to external factors.
  - The set of metrics should collectively capture the key contributors to the desired impact; and implementers should consider the measurement costs involved—fewer payment metrics can reduce measurement costs and help draw attention to the most impactful elements of your program.
  - The payment metrics should minimize potential perverse incentives, which may include: (1) de-prioritization of key results in multi-dimensional programs; (2) de-prioritization of subsets of the population due to differences in the effort required to achieve impact; (3) de-prioritization of subsets of the target population as a result of a binary or continuous metric; or (4) conflicts with the beneficiaries' best interests.
- Payment metrics may adversely affect your flexibility to adapt your program if one metric prescribes a way to reach another. This might not only limit your ability to innovate and maximize the desired impact, but also increase your payment risk.
- The fewer assumptions the causal link between your activities and the metric relies upon, and the less vulnerable the link is to external factors, the lower your payment risk.

#### Evaluation and measurement method

- The process of measuring results includes two elements and decisions: (1) the evaluation method to estimate the results attributable to the program; and (2) the measurement method that defines by whom, when, and how data is collected.
- The range of evaluation methods available to estimate the results vary, in their ability to produce an unbiased estimate of the effect attributable to the program (rigor), their feasibility, and costs.
  - More rigorous methods, such as well-designed Randomized Control Trials, are useful for implementers because they reduce the risk of under- or over-stating results that have been achieved, and thus protect implementers from not being paid for the results they delivered.
  - More rigorous methods also produce more reliable evidence of program impact and reduce the risk of cream-skimming.
  - However, more rigorous methods are usually also more expensive and sometimes not feasible as they require measuring results for a comparison group.
- The rigor of the evaluation method will be of more concern to implementers when: (1) the amount of funding that is tied to the achievement of results is high; and (2) results are highly sensitive to external factors or will likely change in the absence of the program.



- There are two broad types of data to measure results in social programs: (1) data reported by beneficiaries; and (2) data from direct observation. Which type of data is more accurate depends on the results that are measured and the availability of existing data in your context.
- The data to measure payment metrics can be collected by an independent evaluator, an existing external actor, or by the implementer. Collecting data yourself can help build performance management capacity, which may help your organization improve results and facilitate data-driven decision-making in the long term.
- For each option, implementers should consider how frequently measurement can occur and how accurate measurement will be, as these have implications for the implementer's ability to learn and course-correct, and for the risk of under- or overstating results that have been achieved (payment risk).

### Payment structure

- The payment structure specifies the timing of payment (payment schedule) and how payment varies with results achieved on the payment metrics (payment function).
- The price per unit of results is often set using a cost-plus pricing approach. This approach uses the costs of the program as the basis and adds on a predetermined percentage to the costs to account for the risk transferred. The price per unit is then calculated by dividing this amount by the total results expected.
- The optimal amount of funding tied to results strikes a balance between the potential benefits of providing effective incentives and flexibility, and the costs of potentially magnifying perverse incentives, increasing payment risks, and financing costs for the implementer.
- Thresholds in the payment function create strong incentives to improve results around the threshold, but may reduce staff motivation if set too high and increase the payment risk for the implementer since progress below the threshold is not rewarded.
- Payment functions may also include payment kinks, where the price per unit of results changes after a specified level of results has been achieved. Kinks are usually motivated by an increase or decrease in the marginal costs as more results are achieved.
- Differential prices are used to create stronger incentives to focus on a subset of the target population. They can be either motivated by a higher cost of achieving results for some individuals or because results for one subset are valued more by society or the outcome payer.
- The relative prices of the payment metrics, if these are well-aligned with your organization's desired impact, should approximately reflect the relative costs required to achieve them in order to maintain the alignment. Strong incentives to deliver one metric may lead to a de-prioritization of other key metrics.

- RBF agreements generally have a payment cap, which determines the largest permissible payment, to protect the outcome payer against unlimited expenditures. Implementers should carefully analyze the implications of individual payment caps which reduce their ability to pool payment risk across metrics but can help align incentives with the desired impact.
- To assess the payment risk, implementers should estimate their expected performance. Past evidence can help inform this assessment but needs to be adjusted for changes in the program design or implementation, the general context, the target population, and the evaluation and measurement method used. Section 3.6 outlined a 5-step process to guide implementers through this assessment.
- The frequency of payments strikes a balance between the potential benefits – including reduced financing costs and potential for learning – and the associated costs, in particular measurement cost.

### Contract features

- Given that payment is tied to the achievement of results, implementers should carefully assess if the contract prescribes any activities that would limit their organization's flexibility to adapt their program in the pursuit of results. Implementers should also assess if roles and processes related to results reporting are clearly laid out in the contract.
- The RBF contract should clearly identify key risks associated with factors beyond the implementer's control that may influence results and establish a clear process for handling these risks, including how they are shared among stakeholders. The contract should also establish a dispute resolution mechanism to allow for the prospect that certain risks cannot be anticipated or sufficiently mitigated.

### Prepare for negotiations

- During negotiations with outcome payers, the implementer's main objectives are to ensure the RBF agreement is aligned with their desired impact and carries a fair and manageable level of payment risk. Appendix 3 summarizes the key considerations covered in this chapter into a negotiations checklist.
- Analytical backstopping from a member of your team, who is intimately familiar with your evidence and an easy-to use financial model that links performance predictions to payments, may help you assess how requested modifications to the RBF design affect your risk of not meeting performance targets.
- Implementers should also understand the objectives and preferences of the outcome payer. Distinguishing non-negotiable preferences from negotiable ones will help prioritize your asks and negotiate agreements which meet both parties' objectives.
- A series of practices can lead to better outcomes in RBF negotiations. These include: framing and anchoring negotiations around maximizing the desired impact; exploring alternative design solutions that meet both parties' needs; considering all requests at once to lead a prioritization process; and seeking advice from a technical assistance provider.









# 4

## BUILDING THE CAPACITY TO LEVERAGE THE BENEFITS OF RBF

*Because RBF can be adapted to fit an organization's capabilities, RBF agreements can be designed to suit almost all organizations, regardless of their level of capacity. Nonetheless, an organization will benefit from developing certain key capabilities to coordinate and negotiate with other stakeholders in an RBF agreement. Furthermore, the organization will increase its chances of success by improving its ability to monitor, manage and improve results.*

*This chapter will describe each of these capacities and explain their relevance for RBF. It encourages organizations and funders to make a concerted effort to build these capacities as much as possible before the implementation of the RBF agreement.*

# BUILDING THE CAPACITY TO LEVERAGE THE BENEFITS OF RBF

---

## 4.1 WHY IT MATTERS

RBF is often more challenging than using traditional contracts for two reasons:

- (1) Additional stakeholders in the agreement governance structure and the higher stakes between them creates an even greater need for well managed and collaborative relationships.
- (2) The shift in focus from activities to results necessitates organizational processes that drive improved performance and an organizational culture that empowers staff to deliver these improvements.

This chapter focuses on two critical capacities to address these factors: stakeholder management and performance management. Building these capabilities enables better negotiation and implementation of RBF.

## 4.2 STAKEHOLDER MANAGEMENT SKILLS

Implementers will benefit from building trust-based relationships with key stakeholders to facilitate collaborative problem solving. Below we provide key recommendations for stakeholder engagement.

### » Outcome payers

The shift in relationship between implementers and outcome payers enabled by RBF is perhaps the most dramatic. Under RBF, the accountability framework shifts from holding implementers accountable for activities to holding them accountable for results. Building outcome payers' understanding of how these results are achieved is central to building their willingness to respond as risks emerge. Outcome payers will only be willing to renegotiate elements of the contract if they understand why a shortfall in performance was caused by factors outside of the implementer's control. Therefore, implementers will benefit from actively connecting outcome payers to their work, the context they operate in, and the challenges they face on the ground. This may require implementers to increase their openness and transparency regarding challenges and implementation failures. For example, rather than providing outcome payers with the bare minimum of compliance driven reporting, implementers can provide regular updates and reports that give early warnings, highlighting unforeseen implementation challenges and providing insights on course corrections taken.

## » Investors

For SIBs and DIBs, the involvement of an investor may be a completely new experience for implementers. Given the investors' strong incentive to see the program succeed, implementers should consider them an ally and engage them extensively as a source of advice. Investors can bring a valuable outside perspective to the program, challenge your organization in new ways, and provide insights and resources to enable effective course corrections to the program. For example, in the Educate Girls DIB in India, the investor (UBS Optimus Foundation) provided valuable support for developing the performance management system to accelerate Educate Girls' learning.<sup>42</sup> Maintaining alignment between the implementer and investor also helps in negotiations with the outcome payer.

Helping investors develop a strong understanding of the program and challenges on the ground is central to a strong partnership. To this end, we recommend inviting representatives of the investors to spend time on the ground with the implementer before and during program implementation. In the case of the Educate Girls DIB, quarterly reports and conversations regarding results achieved, day-to-day implementation challenges, and proposed course corrections strengthened the implementer's partnership with the investor.

The partnership and engagement should reflect the investors' capacity, preferences, and levels of expertise. Different investors might bring different levels of expertise to the program and might have different preferred levels of involvement. For example, in the Department for Works and Pensions Innovation Fund in the UK, larger investing institutions were more directly involved in the operations of the implementers, with some investors becoming highly involved in building the capacity of implementers over sustained periods of time.<sup>43</sup>

## » Implementation partners

Successful implementation is critically dependent on successful collaboration between implementation partners. This includes the implementer's internal stakeholders (e.g., the organization's staff) and external stakeholders (e.g., sub-contractors engaged for elements of the delivery).

For internal stakeholders, the introduction of RBF serves as an opportunity to redouble the focus on results, and empower team members to continuously improve performance. However, it is not a given that RBF incentives will trickle down to project teams. Strategies to ensure that the high-level incentives of RBF are meaningful to teams include: i) considering changes to internal incentive structures to align teams with the RBF structure – for instance, some implementers provide group-based performance incentives to front-line workers based on the organization's performance on payment metrics; ii) strengthening the process of information sharing around results to keep staff engaged and enable them to engage in problem-solving; and iii) promoting a performance focused culture that embraces change and continual learning. Section 4.3 on performance management provides examples and recommendations on how to build these capacities. Without careful preparation and capacity building, there may be a risk that staff cannot deliver the expected results, that perverse behaviors develop, or that job satisfaction declines.

<sup>42</sup> OPM (2014).

<sup>43</sup> Department for Works and Pensions (2016).



Your organization might subcontract portions of your program delivery responsibilities to other organizations. The same principles of incentive alignment and adequate transfer of payment risk described in the previous sections of this chapter apply to any subcontracts your organization may hold. Your organization may extend the terms of the RBF agreement to the sub-contractors by making some portion of their payments contingent on the results they achieve. This could help align incentives around the same results that your organization is paid for. However, if this is not carefully designed it may also introduce perverse behaviors or jeopardize the partnership by transferring more payment risk to the implementation partners than they can bear. Beyond providing financial incentives to implementation partners, implementers should consider how to empower implementation partners to support the objectives of the RBF agreement.

### » Independent Evaluator

Unlike the other parties, the role of the independent evaluator is not to help improve the program's results, but to accurately measure them as the basis for payments. Given this policing function, there might be some level of tension in the relationship with the evaluator. However, this tension should not create a barrier to working effectively with the evaluator: it is in all parties' interest to ensure timely and reliable access to the data needed to robustly verify the results achieved and to make this process as smooth and productive as possible.

## 4.3 CAPACITY FOR PERFORMANCE MANAGEMENT

Performance management is the data-driven iteration of the program design and delivery practices of an organization that improve their performance on key payment metrics in the RBF agreement. Strong performance management capabilities maximize impact and minimize payment risks.

This section presents an overview of performance management, offering key recommendations and examples to inform organizations engaging in or considering RBF.

### 4.3.1 PERFORMANCE MANAGEMENT VERSUS MONITORING OR IMPACT EVALUATION

Performance management represents a change from the traditional monitoring and evaluation practices of the sector. Traditional monitoring tracks inputs and activities to understand if the implementation is going according to the plan, usually agreed upon with the donor. Impact evaluation usually takes place after program implementation to understand if the program had the expected impact, not offering on-time insights for program managers. Some insights can be drawn from monitoring and evaluation for future program design and implementation, but neither aim to inform rapid iteration of program design and delivery practices to improve their outcomes during implementation.

As a result, organizations operate with insufficient insights on whether and how their inputs and activities are leading to the expected outcomes. Performance management shifts the measurement focus towards generating near real-time managerial insights for program managers to make dynamic adjustments to improve outcomes.

“ One of the benefits of the DIB has been building an organization that has performance management in its DNA. Regardless of the scale, every part of the organization is delivering to outcomes.”

Safeena Hussain  
Executive Director at Educate Girls

**FIGURE 17.** Comparison of monitoring, performance management, and impact evaluation

	Monitoring	Performance Management	Impact Evaluation
<b>Scope</b>	Inputs, activities and outputs	<b>Results</b>	Impact
<b>Time frame</b>	Short term – during implementation	Short and medium term – <b>throughout implementation</b>	Long term – after implementation
<b>Insight</b>	Are we implementing according to plan?	Are we on track to impact and <b>how to improve?</b>	Did it work?
<b>Purpose</b>	Report	<b>Learn</b> how we can improve the program	Transparency and accountability

A well-functioning performance management system relies on three pillars: i) relevant and well-timed data on program implementation and results; ii) an information system that collects and processes the data to generate real-time and clear insights for decision makers; and iii) a team with the skills, culture, and incentives to successfully transform performance insights into improved practices.

### 4.3.2 RELEVANT DATA

Performance management is not about collecting a lot of data, but the right data, which drives your understanding of whether and how the program is achieving critical results. There are two different types of data: performance indicators and indicators of external factors.

- **Performance indicators** should measure activities, outputs, and outcomes:
  - a) that are relevant for achieving the desired impact;
  - b) where there is less certainty about your performance, as this data will shed more light on how results are achieved; and
  - c) that are under your control, as these are the ones your organization will be able to influence during program implementation.
- **Indicators of external factors**, such as individual or village-level characteristics, provide information of factors that are not related to the implementation of the program, but might have an impact on the program's results. They contextualize performance indicators and provide a more nuanced understanding of the drivers and barriers to impact.

### Box 13. Performance management case study: Out of home care

Casa del Cerro is an implementer in Chile that works with families and foster care institutions to ensure children can return home and to protect their rights. The ultimate objective is to ensure children return to their families as soon as the appropriate conditions are met and to reduce the re-entry rate into the child care system. Two years ago, Instiglio and Fundación San Carlos de Maipo partnered to build the performance management capacities of Casa del Cerro. Since then, Casa del Cerro has been implementing its program with a focus on results with continuous support from Fundación San Carlos de Maipo.

In their performance management system, Casa del Cerro tracks: a) key activities, such as the frequency of sessions with children and the children's family; and b) key results, such as the improvement in children's wellbeing, the time children stay in foster care institutions, and the number of children that re-enter institutions.

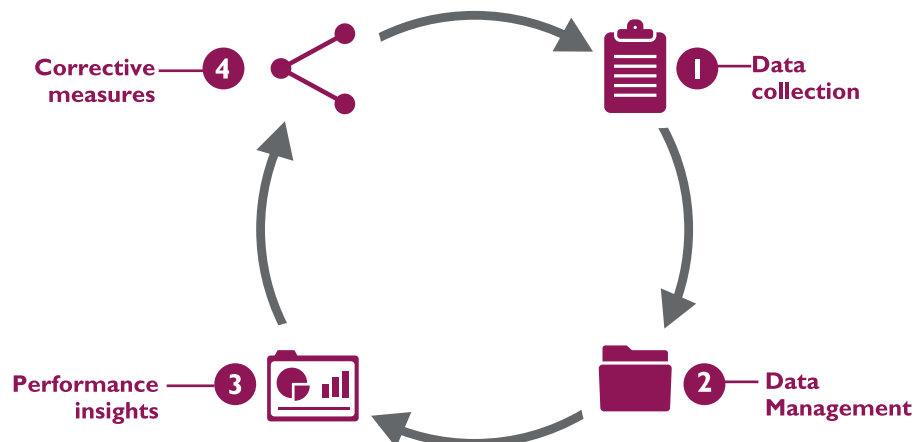
By analyzing the performance indicators by individual characteristics (such as age and gender), family characteristics (such as family composition), and institutional characteristics (such as infrastructure and size of the staff), Casa del Cerro has been able to identify groups that need a more intensive follow up or different approaches to achieve results.

### 4.3.3 AN INFORMATION SYSTEM THAT DELIVERS NEAR REAL-TIME PERFORMANCE INSIGHTS

Performance management requires an efficient process that provides near real-time insights for decision-making. As data might come from different sources and at different times, it is critical to define and map data flows, timelines, roles, and responsibilities to ensure that the system is of high quality, cohesive, and matches capacity. To ensure adoption of the system by the program staff, it is critical for the process to align with the program implementation realities on the ground: timelines, infrastructure, and front-line worker's work load, among others.

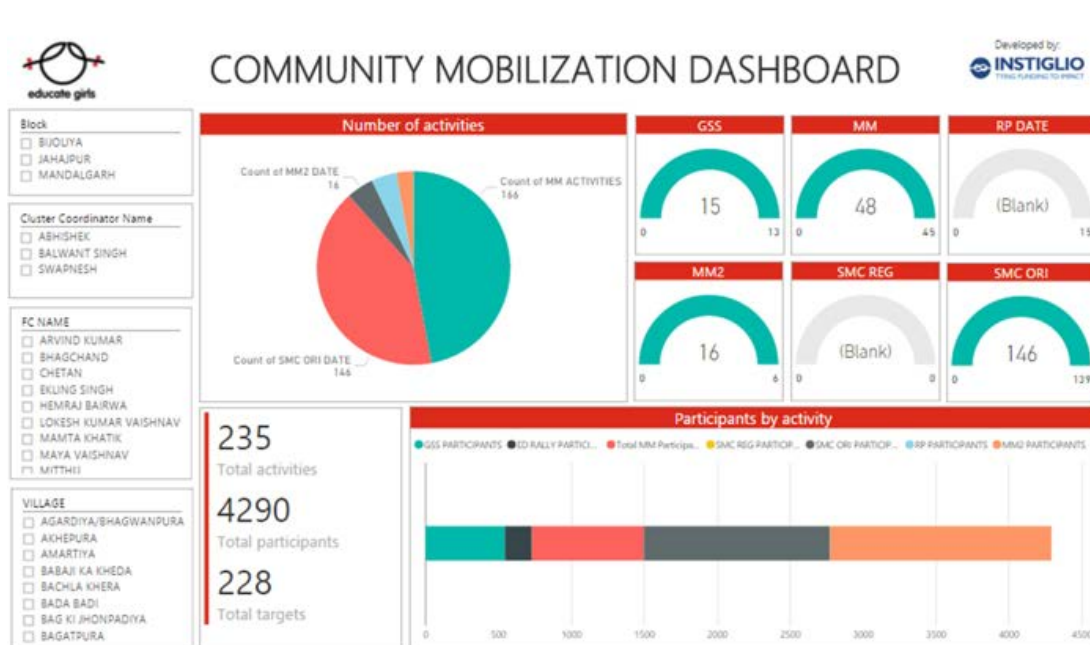
The performance management process has four main stages: i) data collection, ii) data management, iii) performance insights, and iv) course correction.

**FIGURE 18.** Performance management process



- **Data collection:** the purpose of this stage is to capture timely and reliable data. The data collection forms, sources, timelines, and owners need to be laid out clearly to ensure data collection is timely. A good data entry system can further facilitate near real-time data and reduce data entry mistakes. For instance, collecting data on mobile devices that continuously synchronize data with the storage location allows for the simultaneous collection and entry of data.
  - **Data management:** the purpose of data management is to prepare the data for analysis. This includes aggregating data from different sources in a single database, correcting data errors, and performing any additional data calculations to build the performance indicators. A good IT system can help reduce the time from data entry to data analysis by automatizing data aggregation, cleaning, and transformation.
  - **Performance insights:** the purpose of this stage is to transform data into actionable insights. This follows a two-stage process: transforming data into information and transforming information into insights. For both stages, a good data visualization system is key to facilitate the interaction of the team with the data to draw their own performance insights.
- i) Transforming data into information: data are unprocessed facts that need to be aggregated and put into context to become information. This is usually done through dashboards and reports that contain data visualizations of performance. Through these visualizations, the program team is able to easily identify if there is a problem with the program implementation or the expected results. The image below is a sample dashboard from Educate Girls' performance management system.

**FIGURE 19.** Educate Girls performance management dashboard

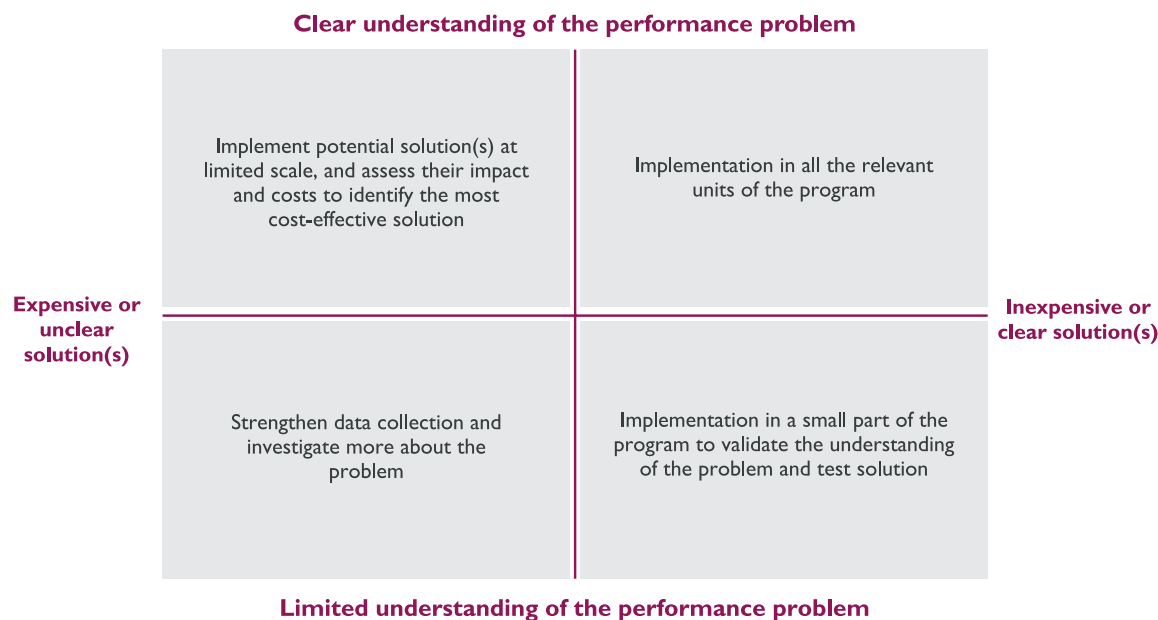


- ii) Transforming information into insights: insights are generated by analyzing information and drawing conclusions. Usually, before potential solutions to an identified problem can be formulated, further analysis is needed to diagnose the causes of the problem. This may require digging deeper into the existing program data or collecting additional quantitative and qualitative data. In both cases, the final goal is to obtain relevant and specific insights that are useful to formulate recommendations on how to improve program implementation.

#### Box 14. Data, information, and insights

Bank statements provide data on financial activity. Without much context, information on the current balance of a bank account, for example \$1,000, is not very useful. Looking at the bank statements for multiple months it is possible to understand if the current balance is too low or too high. Let's assume that, for the last three months, the balance had been increasing by \$200, but the last month it increased by \$100, meaning it is lower than expected. This information can now be used to generate insights into how much effort is needed to achieve a result: for example, a balance of \$1,300 in the next two months. Furthermore, this data provides the user with the information needed to course correct and design a spending strategy to achieve the target in the target time period. An example of this can be identifying budget items where cuts can be made in the next two months.

- **Corrective measures:** the purpose of this stage is to translate insights into clear actions. Particularly when implementing RBF, it is important to analyze the financial implications of adjusting implementation to ensure the results are delivered within the price per unit of result achieved. Depending on your understanding of the performance problem, the cost of proposed solution, and your certainty that it will work, you can use different strategies to implement the solution. For example, where there is a clear and inexpensive solution to a well-identified problem, you might want to directly implement it across your program to improve results as quickly as possible. On the other hand, if you are uncertain that the solution will work or there are alternative potential solutions, you might want to first test all solutions at limited scale before investing a lot of resources to implement one. Figure 20 provides a guiding framework of recommended strategies.

**FIGURE 20.** Classifying corrective measures

#### 4.3.4 GETTING THE RIGHT TEAM IN PLACE

To successfully implement performance management, it is critical to have the right team in place. In addition to having sufficient staff to fill the required roles, the team also needs to have the right skills to fulfill their responsibilities and be guided by a performance culture. Two positions are key to running the performance management system:

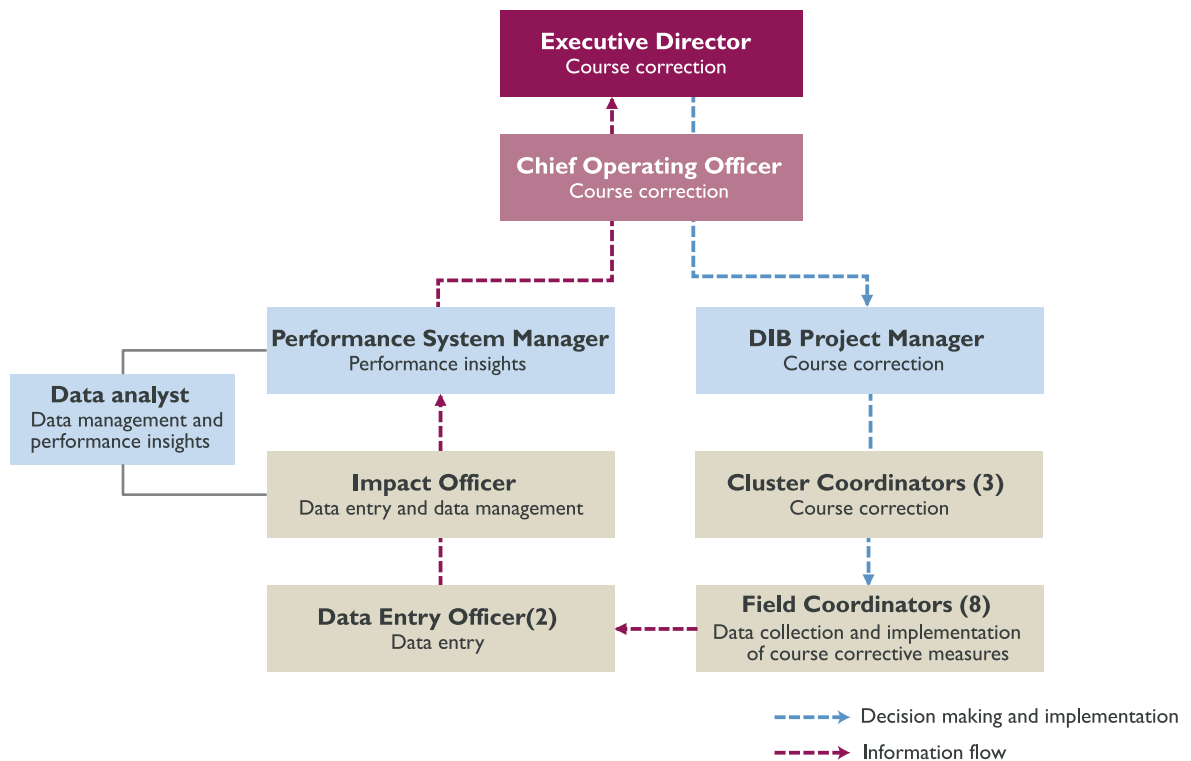
- **The Performance Management System Manager** leads and coordinates the implementation of the performance management system (PMS). The manager is the key link between project managers and decision-makers in the organization. The manager's primary role is to facilitate timely and accurate performance management. The manager works with the leadership to improve the organization's performance culture and use of evidence in decision-making. As the lead of the PMS, the manager will identify and present performance insights to key decision makers and propose clear corrective measures that can increase the efficiency of program implementation and the effectiveness of the program design. To successfully fulfill their role, the manager should have strong communication and analytical skills and be results-oriented, solution-driven, curious, a team player, and familiar with statistical analysis.
- **The Data Analyst** works with the Performance System Manager to draw insights from the performance management data. They are responsible for supporting data entry and storage, preparing and analyzing data, building data visualizations that facilitate the interpretation of performance insights at all levels of the organization, and supporting the manager in the formulation of recommendations to improve the impact of the project based on findings. The ideal analyst has good knowledge of quantitative statistical techniques, is proficient in different data analysis IT solutions, and can analyze and present data in a structured manner.



Other team members involved in performance management should also have a good understanding of qualitative and quantitative data collection techniques and analytical and data analysis skills, depending on their specific roles.

Figure 21 below describes the structure of the Educate Girls' performance management team that collects real-time data and implements course corrective measures to achieve the desired results in the DIB.

**FIGURE 21.** The Educate Girls performance management team structure



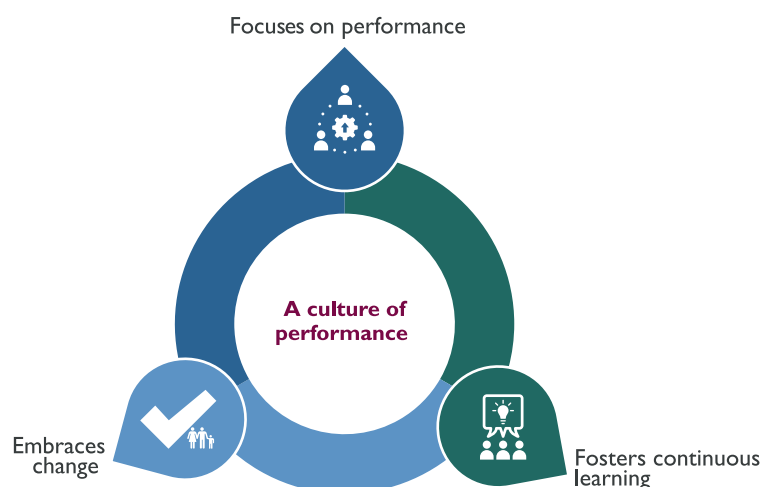
### » Performance culture

Data is often not enough to make an organization outcomes-oriented: 98 percent of nonprofit organizations collect a lot of information but a third of them are unable to reflect on this information and integrate it in a meaningful way into program activities.<sup>44</sup> According to Mario Morino, the author of *Leap of Reason*, "...the organization should have the mental mindset to do what it does as well as it possibly can and continually seek to do even better."<sup>45</sup> Clearly, the culture of the organization is important. Based on others' experience and our own, we have identified three key characteristics of a performance culture.

<sup>44</sup> Smith Milway and Saxton (2011).

<sup>45</sup> Morino (2011).

**FIGURE 22.** Characteristics of a performance culture



- **Focus on performance:** Regular tracking and communication of results and performance targets helps to keep staff motivated and engaged with their impact. Tracking results at the staff level is useful to highlight areas to strengthen and opportunities for sharing best practices, insights, and learnings. For instance, Roca, an implementer in the Massachusetts recidivism SIB, consistently collects data to inform their decision-making. Insights from data are incorporated in their meetings in a weekly, monthly, quarterly, and annual basis, and are used to assess and adapt their model.<sup>46</sup>
- **Continuous learning:** Opening space for the team to suggest areas for improvement, detect challenges early, and problem solve with the aid of data is critical. The Urban Institute suggests a number of ways to involve staff in a continuous learning process. These practices include reflective supervision, data review meetings, and the presentation of data insights in regular staff meetings. Cure Violence, an organization working to stop the spread of violence in communities has integrated formal weekly staff meetings and regular supervisor reviews to discuss and update their current understanding of violence in communities and identify strategies for interrupting it.<sup>47</sup>
- **Open to change:** Flexibility to adapt your program is a major benefit of RBF. Organizations cannot benefit from the flexibility provided by RBF if they are not prepared to change their approach. Willingness to challenge assumptions and test innovations in programmatic design is key to improving performance. A good example of this is the International Planned Parenthood Foundation. In pursuit of greater impact, they implemented a Branch Performance Tool that allows them to track activities and results at the branch level. Following the implementation of this tool, the Uganda branch dramatically increased the number of mobile clinics, tripled the number of contraceptives delivered, and eliminated multiday waiting lines. In Cambodia, their branch was able to identify and close low-performing clinics, move clinics closer to women who needed services, and expand the services offered.<sup>48</sup>

<sup>46</sup> Pierce (2009); Roca (2013).

<sup>47</sup> Cure Violence (2017).

<sup>48</sup> Green, Fahnestock, and Blau (2015).

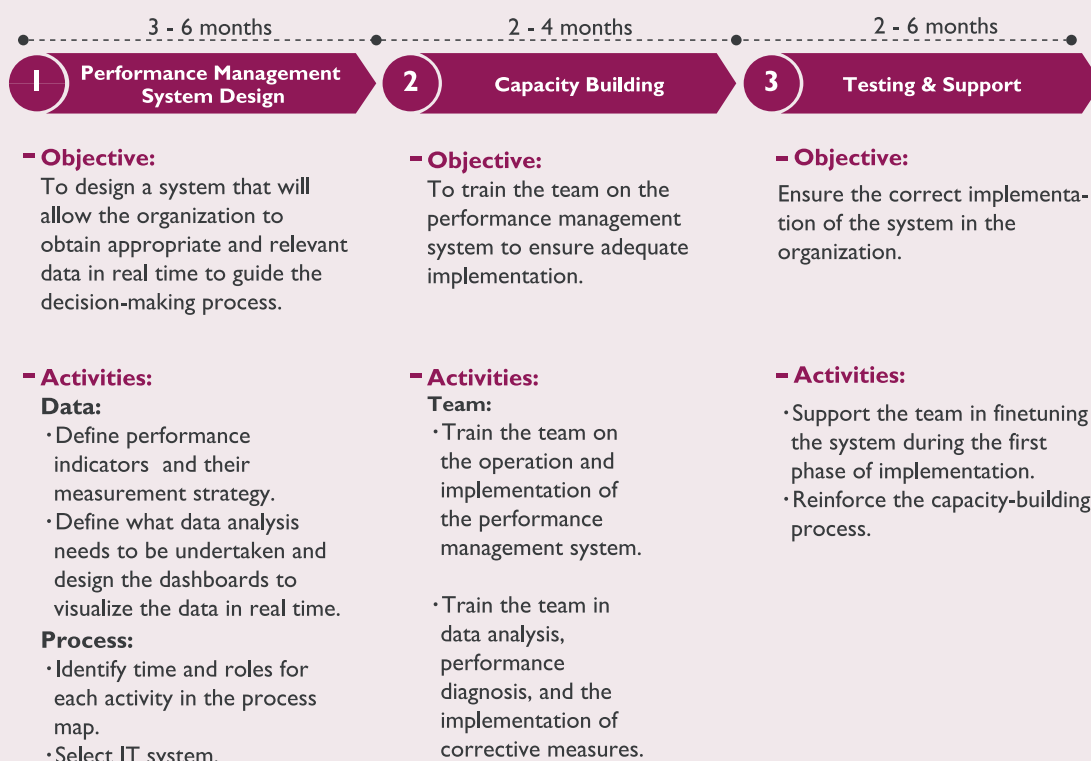
### 4.3.5 WHAT IT TAKES TO BUILD YOUR ORGANIZATION'S CAPACITY FOR PERFORMANCE MANAGEMENT

There are three potential ways to build the capacity to manage outcomes in your organization, if the organization does not already have this capacity: i) hire someone who has the technical knowledge and the experience to set up and manage performance management; ii) train the organization staff to design and implement performance management systems; or iii) hire an external organization to design the system and build the internal capacity of the organization to implement it. The time and cost of setting up and implementing the performance management system will depend on:

- The scope and scale of the system, which depends, among other factors, on the complexity of the program, the number of indicators to measure, the number of dashboards to design, and the number of individuals in the team implementing performance management who need training.
- The existing M&E infrastructure of the organization: designing something from scratch is often more expensive than building on an existing system.
- The extent of the capacity building required to allow your team to manage the performance management system, including trainings and accompaniment during implementation.
- The IT infrastructure: depending on the IT solutions used, significant investments in hardware and software might be required. However, there are also free or very low-cost solutions available.

### Box 15. Process to set up a performance management system

Setting up a performance management system is a three-step process that starts with the design of the system, is followed by a capacity building phase, and finishes with testing to fine-tune elements while assisting in the generation of insights and course corrections in the first performance management cycles. The figure below provides more details on the phases, activities, and estimated times to set up a performance management system.



If your organization is considering engaging in RBF, we recommend budgeting the time and cost to design, implement, and test the performance management system before implementation of RBF starts.

## 4.4 KEY TAKEAWAYS

- An organization implementing RBF will benefit from developing certain key capabilities to coordinate and negotiate with other stakeholders and increase its ability to monitor, manage, and improve results.

### Stakeholder management skills

- Implementers should actively build the outcome payer's understanding of the program to increase their willingness to respond flexibly as implementation challenges emerge.
- Investors can bring a valuable outside perspective to the program, challenge your organization in new ways, and provide insights and relationships to enable effective course corrections to the program.
- The introduction of RBF serves as an opportunity for implementers to re-engage their teams, redouble their focus on results, and empower staff to find solutions that enhance performance.
- When engaging other organizations in the implementation, implementers might explore ways of extending the incentives and flexibility provided in the RBF agreement to their delivery partners.
- The implementer's relationship with the evaluator responsible may have elements of tension, but this should not create a barrier to working productively together and using the evaluation process as a learning opportunity.

### Performance management capacity

- Performance management is the data-driven iteration of the program design and delivery practices of an organization that improve their performance on key payment metrics in the RBF agreement.
- A well-functioning performance management system relies on three pillars: i) relevant and well-timed data; ii) an information system that collects and processes the data to generate real-time insights; and iii) a team with the skills, culture, and incentives to successfully transform performance insights into improved practices.
- Relevant and well-timed data. Performance management is not about collecting a lot of data, but the right data, which drives understanding of whether and how the program is achieving critical results. This includes collecting data on:
  - Performance indicators that measure activities, outputs, and outcomes a) relevant for achieving the desired impact; b) where there is less certainty about performance, and c) that are under the implementer's control.
  - Indicators of external factors, such as individual or village-level characteristics, that contextualize performance indicators and provide a more nuanced understanding of factors that influence results.

- An information system that collects and processes the data to generate real-time and clear insights for decision makers. The performance management process has four main stages: i) data collection, ii) data management, iii) performance insights, and iv) course correction. The purpose of the third and fourth stages is to transform data into actionable insights and derive clear actions from them.
- A team with the skills, culture, and incentives to successfully transform performance insights into improved practices. Data is often not enough to successfully implement performance management and achieve results.
  - In addition to having sufficient staff to fill the required roles, the team also needs to have the right skills to fulfill their responsibilities and be guided by a performance culture.
  - We have identified three key characteristics of a performance culture: i) focus on performance; ii) continuous learning, and iii) openness to change.

# REFERENCES

- Abdul Latif Jameel Poverty Action Lab. (n.d.). *Experimental Methodology Table*. Online. Retrieved September 7, 2016 from <https://www.povertyactionlab.org/sites/default/files/documents/Experimental%20Methodology%20Table.pdf>.
- AgResults. (n.d.). *Vietnam GHG Emissions Reduction Pilot*. Online. Retrieved September 6, 2017 from <http://agresults.org/en/313/Vietnam-GHG-Emissions-Reduction-Pilot>.
- Andrews M., Pritchett L., & Woolcock M. (2017). *Building State Capability: Evidence, Analysis, Action*. Oxford University Press.
- Bain, K. (2016). *Doing Development Differently: Updating the Plumbing to Match the Architecture*, London: Overseas Development Institute.
- Banerjee, A., E. Duflo, N. Goldberg, D. Karlan, R. Osei, W. Pariente, J. Shapiro, B. Thuysbaert, & C. Urdy. (2015). *A multifaceted program causes lasting progress for the very poor: Evidence from six countries*. *Science* 348 (6236): 1260799. DOI: 10.1126/science.1260799.
- Barder, O. (2014). *12 Principles for Payment by Results (PbR) in the Real World*. The Center for Global Development. <https://www.cgdev.org/blog/12-principles-payment-results-pbr-real-world-0>.
- Basinga, P., P. Gertler, A. Binagwaho, A. Soucat, J. Sturdy, C. Vermeersch. (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. [http://dx.doi.org/10.1016/S0140-6736\(11\)60177-3](http://dx.doi.org/10.1016/S0140-6736(11)60177-3).
- Birdsall N. & Savedoff, W. (2010). *Cash on Delivery: A New Approach to Foreign Aid*. Washington, DC.: Center for Global Development.
- Chonghaile, C.N. (2016). *Less aid money, less influence: Brexit's likely hit to the UK's development role*. The Guardian. <https://www.theguardian.com/global-development/2016/jul/07/brexit-uk-development-role-less-aid-money-less-influence>.
- Cure Violence. (2017). *Implementation – Essential Steps to Implementing Health Model*. Online. Retrieved September 4, 2017 from <http://cureviolence.org/the-model/implementation/essential-steps-to-implementing-health-model/>.
- Department for Work and Pensions. (2014). *Work Programme evaluation: operation of the commissioning model, finance and programme delivery*. London: Department for Work and Pensions. [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/425301/rr893-report.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/425301/rr893-report.pdf).
- Department for Work and Pensions. (2016). *Qualitative evaluation of the DWP Innovation Fund: Final report*. London: Department for Work and Pensions. <https://www.gov.uk/government/publications/qualitative-evaluation-of-the-dwp-innovation-fund-final-report>.
- Departamento Administrativo para la Prosperidad Social. (2017). *Convenio de Cooperación 282 de 2017*.
- Department for International Development (DFID). (2013). *Business Case and Intervention Summary*. Inclusive Growth and Prosperity Grant. Budget Support to Sierra Leone 2013-2015.
- DFID. (2014). *Payment by Results Strategy: Sharpening incentives to perform*. Online. Retrieved September 6, 2017 from <https://www.gov.uk/government/publications/dfids-strategy-for-payment-by-results-sharpening-incentives-to-perform>.
- DFID. (2016). *Raising the standard: the Multilateral Development Review 2016*. Online. Retrieved September 6, 2017 from <https://www.gov.uk/government/publications/raising-the-standard-the-multilateral-development-review-2016>.
- DFID. (2017). *Girls' Education Challenge*. Online. Retrieved September 6, 2017 from <https://www.gov.uk/guidance/girls-education-challenge>.



- Engineer, C. Y., Dale, E., Agarwal, A., Agarwal, A., Alonge, O., Edward, A., Gupta, S., Schuh, H.B., Burnham, G., & Peters, D. H. (2016). *Effectiveness of a pay-for-performance intervention to improve maternal and child health services in Afghanistan: a cluster-randomized trial*. *International Journal of Epidemiology* 45 (2): 451-9. <https://doi.org/10.1093/ije/dyv362>.
- Global Learning XPRIZE. (n.d.). Online. Retrieved September 6, 2017 from <http://learning.xprize.org/>.
- Global Partnership on Output-Based Aid (GPOBA). (n.d.). Online. Retrieved September 6, 2017 from <https://www.gpoba.org/>.
- Government of Lao PDR. (2016). *District Development Fund (DDF) Final Report on Impacts and Lessons Learnt*. National Governance and Public Administration Reform programme – Strengthening Capacity and Service Delivery of Local Administrations (SCSD) 2012-2016.
- Green, L., Fahnestock, M. & Blau, J. (2015). *Planned Performance*. Stanford Social Innovation Review. Winter 2015.
- Gustafsson-Wright, E., Gardiner, S., & Putcha, V. (2015). *The Potential and Limitations of Impact Bonds: The First Five Years of Experience Worldwide*. The Brookings Institution. <https://www.brookings.edu/research/the-potential-and-limitations-of-impact-bonds-lessons-from-the-first-five-years-of-experience-worldwide/>.
- Instiglio. (2015). *Educate Girls Development Impact Bond: Improving Education for 18,000 Children in Rajasthan*. <http://instiglio.org/educategirlsdib/wp-content/uploads/2016/03/EG-DIB-Design-1.pdf>.
- Instiglio. (n.d.). *Youth in Mexico*. Online. Retrieved September 12, 2017 from <http://www.instiglio.org/en/education-chiapas/>.
- Kenny, C. (2017). *Results Not Receipts: Counting the Right Things in Aid and Corruption*. Brookings Institution Press.
- Killefer, N. (2011). *Doing more with less: A government roundtable*. McKinsey & Company. [https://www.mckinsey.com/~media/mckinsey/dotcom/client\\_service/Public%20Sector/PDFS/McK%20on%20Govt/Change%20under%20pressure/TG\\_MoG\\_6\\_Doing\\_more\\_with\\_less.ashx](https://www.mckinsey.com/~media/mckinsey/dotcom/client_service/Public%20Sector/PDFS/McK%20on%20Govt/Change%20under%20pressure/TG_MoG_6_Doing_more_with_less.ashx).
- Lannes, L. (2015). *Improving health worker performance: The patient-perspective from a PBF program in Rwanda*. *Social Science & Medicine* 138, August 2015.
- Lax, D.A., Whitener, B., & Sebenius, J. K. (2007). *3D Negotiation*. Gildan Media Corporation.
- Massachusetts Institute of Technology, (n.d.). *Evaluating Social Programs*. Online. Retrieved September 7, 2017 from <https://www.edx.org/course/evaluating-social-programs-mitx-jpal101x-6>.
- Morino, M. (2011). Leap of reason: Managing to outcomes in an era of scarcity. *Innovations* 6(3): 167-177.
- New Way to Finance Health in World's Less Developed Nations. (2013). Online. Retrieved September 6, 2017 from [http://www.who.int/pmnch/media/news/2013/rbf\\_press\\_release.pdf?ua=1](http://www.who.int/pmnch/media/news/2013/rbf_press_release.pdf?ua=1)
- O'Carroll, L. (2016). £1.3bn troubled families scheme has had 'no discernible impact'. *The Guardian*. August 8, 2016. [https://www.theguardian.com/society/2016/aug/08/13bn-troubled-families-scheme-has-had-no-discernible-impact?CMP=Share\\_iOSApp\\_Other](https://www.theguardian.com/society/2016/aug/08/13bn-troubled-families-scheme-has-had-no-discernible-impact?CMP=Share_iOSApp_Other).
- OPM. (2014). *Evaluation of the Essex Multi-Systemic Therapy Social Impact Bond: Interim Evaluation Report*. Online. Retrieved September 6, 2017 from <https://www.opm.co.uk/wp-content/uploads/2015/10/Interim-report-Essex-MST-SIB-Evaluation.pdf>.
- Perakis R., & Savedoff W. (2015). *Does results-based aid change anything? Pecuniary interests, attention, accountability and discretion in four case studies*. Center for Global Development Policy Paper 52.
- Pierce, B.A. (2009). *Roca's High Risk Youth Intervention Model: Initial Implementation Evaluation Report*. Crime and Justice Institute at Community Resources for Justice.
- Popper, N. (2015). *Success Metrics Questioned in School Program Funded by Goldman*. *The New York Times*. November 3, 2015. Print.
- Ramalingam, B. (2013). *Aid on the edge of chaos: rethinking international cooperation in a complex world*. Oxford University Press.

- Roca. (2013). *2013 Annual Report – Truth, Trust and Transformation: The Next 25 Years*.
- Rondinelli, D. (1993). *Development Projects as Policy Experiments: An Adaptive Approach to Development Administration*. New York: Routledge.
- Saldinger, A. (2017). Development impact bonds gain momentum. Devex. <https://www.devex.com/news/development-impact-bonds-gain-momentum-90591>.
- Sithara K, Grossman J, & Overholser G. (2014). *The Massachusetts Juvenile Justice PFS Initiative: Lessons Learned*. Boston: Third Sector Capital Partners.
- Smith Milway, K. & A. Saxton. (2011). *The Challenges of Organizational Learning*. Stanford Social Innovations Review. Summer 2011.
- Soeters, R., Peerenboom, P.B., Mushagalusa, P., & Kimanuka, C. (2011). *Performance-Based Financing Experiment Improved Health Care in The Democratic of the Congo*. Health Affairs 30 (8): 1518-27. <http://content.healthaffairs.org/content/30/8/1518.abstract>.
- Stephens, B. (2017). *Driving development effectiveness: Exploring the frontier of innovation in social spending*. Harvard Kennedy School Policy Analysis Exercise [unpublished].
- Sturgess, G. & Cumming, L. (2011). *Payment by Outcome – A Commissioner's Toolkit*. 2020 Public Services Trust at the RSA. [http://www.2020publicservicestrust.org/downloads/3\\_Payment-by-Outcome-A-Commissioners-Toolkit-2020-Public-Services-Trust-2011.pdf](http://www.2020publicservicestrust.org/downloads/3_Payment-by-Outcome-A-Commissioners-Toolkit-2020-Public-Services-Trust-2011.pdf).
- Third Sector Capital Partners. (2015). *The Massachusetts Juvenile Justice PFS Initiative: Project Brief*. Online. Retrieved September 6, 2017 from [http://www.thirdsectorcap.org/wp-content/uploads/2015/02/TSCP\\_MAJJ-PFS-Project-Brief.pdf](http://www.thirdsectorcap.org/wp-content/uploads/2015/02/TSCP_MAJJ-PFS-Project-Brief.pdf).
- Webster, R. (2016). *Payment by Results: Lessons from the Literature*. Online. Retrieved September 6, 2017 from <http://www.russellwebster.com/payment-by-results-lessons-from-the-literature/>.
- Wheeler, M. (2000). *Negotiation Analysis: An Introduction*. Harvard Business School Background Note 801-156.
- World Bank. (2007). *GPOBA Commitment Paper: Vietnam EMWF GPOBA Rural Water Supply Development Project*. Online. Retrieved September 6, 2017 from [https://www.ewb.org.au/resources/download/514P2010-01-13\\_18:50:45/L](https://www.ewb.org.au/resources/download/514P2010-01-13_18:50:45/L).
- World Bank. (2008). *Project Appraisal Document on a Proposed Loan in the Amount of US\$50 Million to the Republic of Botswana for a National HIV/AIDS Prevention Support Project*. Report No: 43112-BW.
- World Bank. (2010). *Helping Botswana Confront a Challenging Future*. IBRB Results. 91813.
- World Bank. (2013). *Program Appraisal Document on a Proposed Loan in the Amount of SDR65.1 Million (US\$100 Million Equivalent) to the Federal Democratic Republic of Ethiopia with a Proposed Grant from the Health Results Innovation Trust Fund (HRITF) in the Amount of US\$20 Million in Support of the Millennium Development Goals Program-for-Results*. Report No: 75101-ET.
- World Bank. (2014). *Project Appraisal Document on a Proposed Grant in the Amount of SDR 10.22 and a Proposed Grant in the Amount of US\$5 Million from the Multi-Donor Trust Fund for Health Results Innovation to the Republic of Chad for a Mother and Child Health Services Strengthening Project*. Report No: PAD885.
- World Bank. (2016a). *Using RBF to Improve the Delivery of Maternal and Child Health Services in Zambia* (webinar). Impact Evaluation Results; April 26th 2016.
- World Bank. (2016b). *Vietnam Rural Water Webinar Presentation*. Online. Retrieved September 6, 2017 from <https://collaboration.worldbank.org/docs/DOC-20571?sr=stream>.
- World Bank. (2017). *Program-for-Results Financing (PforR)*. Online. Retrieved September 6, 2017 from <http://www.worldbank.org/en/programs/program-for-results-financing>.
- World Health Organization. (2011). *Global Tuberculosis Control 2011*. Online. Retrieved September 6, 2017 from [http://apps.who.int/iris/bitstream/10665/44728/1/9789241564380\\_eng.pdf](http://apps.who.int/iris/bitstream/10665/44728/1/9789241564380_eng.pdf).





# APPENDIX

# APPENDIX I: RESULTS-BASED FINANCING (RBF) CHEAT SHEET

“Funding tied to **pre-agreed** and **verified results**”

## Value-add of RBF for implementers

1

### IMPROVING IMPACT



#### DRAWING ATTENTION TO WHAT MATTERS

Paying for results requires investing in data systems and regularly tracking results. Thus, RBF gives implementers continuous feedback on their impact trajectory.



#### ALIGNING INCENTIVES TO BENEFICIARIES' WELFARE

By tying funding to agreed-upon results, RBF aligns the objectives of outcome payers and implementers with the beneficiaries' welfare. Thus, RBF financially rewards the implementer for achieving its mission, rather than for meeting any other stakeholders' interests.



#### PROVIDING FLEXIBILITY TO MAXIMIZE RESULTS

RBF tightens the control over achieving results; allowing outcome payers to relax their control over activities and grant implementers the flexibility and freedom to pursue a range of strategies to achieve greater impact. Implementers are free to invest in necessary overheads, to monitor and adapt their program to changing contexts, and to experiment with new delivery strategies without lengthy funder consultations or approvals.

2

### IMPROVING FUNDING PROSPECTS

#### IMPLEMENTERS CAN ENGAGE NEW RISK-AVERSE FUNDERS

Given funders only pay (at least the totality) if the program achieves predefined results, risk-averse funders may be more willing to fund RBF programs – particularly relatively innovative ones. This can help attract more funding and create opportunities to expand programs and generate more impact.

#### IMPLEMENTERS ARE MORE COMPETITIVE IN SECURING FUNDS

Since RBF requires implementers to bet on their performance, only high-performing implementers tend to respond. Therefore, funders develop greater confidence in implementers who are ready to engage with RBF.

### RBF design components



**Financial size:** total funding provided under the RBF agreement, including funding not tied to results and any payments made if pre-defined results are achieved.

**Duration:** the period of time in which the RBF terms apply.

**Target population:** the geographic and socio-demographic specifications of the individuals for which the payment metrics will be measured.

**Payment metrics:** results that are paid for and apply to the defined target population.

**Evaluation method:** the evaluation method used to determine results.

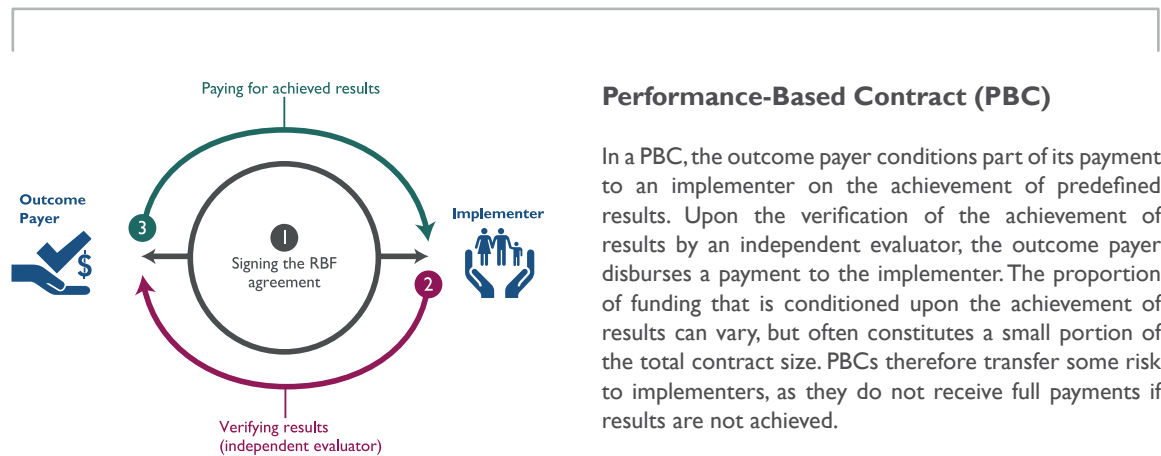
**Measurement method:** how, when, and by whom data is collected and results measured.

**Payment function:** the function that determines payment given the results achieved.

**Payment schedule:** the timing of the payments to the implementer (or investor(s)).

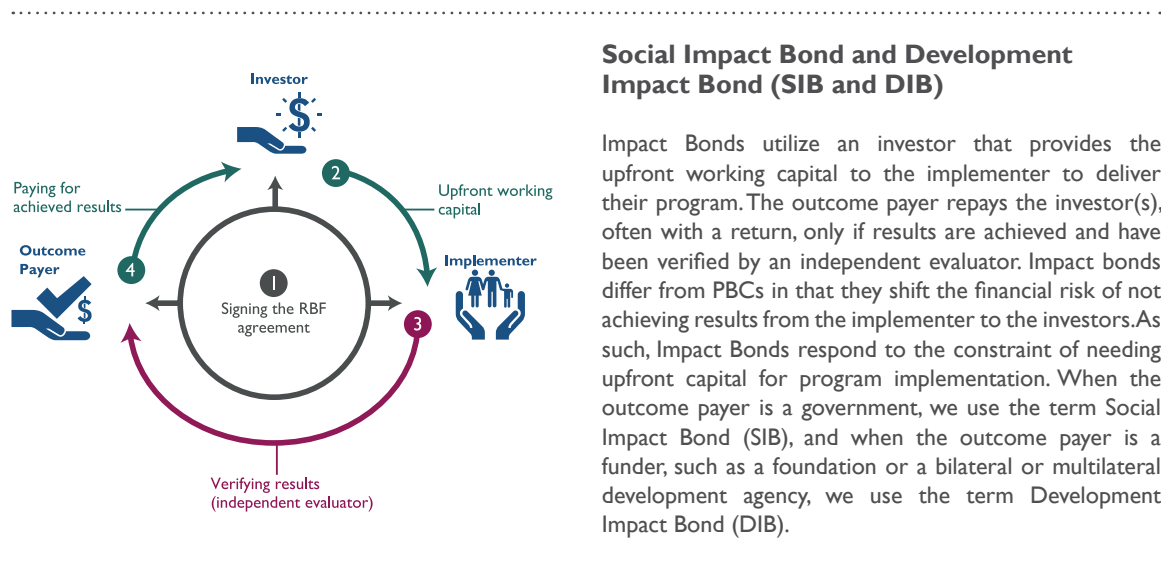


## RBF instruments



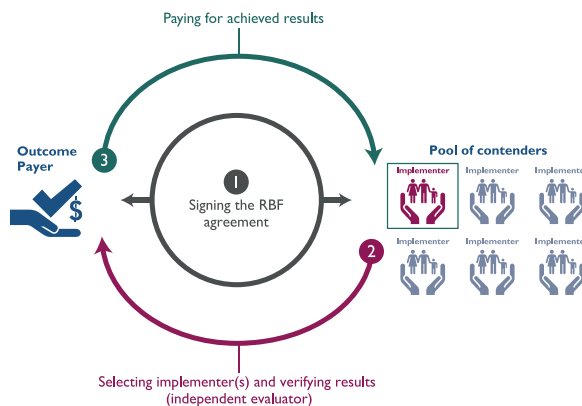
### Performance-Based Contract (PBC)

In a PBC, the outcome payer conditions part of its payment to an implementer on the achievement of predefined results. Upon the verification of the achievement of results by an independent evaluator, the outcome payer disburses a payment to the implementer. The proportion of funding that is conditioned upon the achievement of results can vary, but often constitutes a small portion of the total contract size. PBCs therefore transfer some risk to implementers, as they do not receive full payments if results are not achieved.



### Social Impact Bond and Development Impact Bond (SIB and DIB)

Impact Bonds utilize an investor that provides the upfront working capital to the implementer to deliver their program. The outcome payer repays the investor(s), often with a return, only if results are achieved and have been verified by an independent evaluator. Impact bonds differ from PBCs in that they shift the financial risk of not achieving results from the implementer to the investors. As such, Impact Bonds respond to the constraint of needing upfront capital for program implementation. When the outcome payer is a government, we use the term Social Impact Bond (SIB), and when the outcome payer is a funder, such as a foundation or a bilateral or multilateral development agency, we use the term Development Impact Bond (DIB).



### Prize-Based Challenge

A Prize-Based Challenge is an open bid competition that awards a financial prize for the best innovation developed within a predefined timeframe. Although similar to other forms of competition, Prize-Based Challenges reward competitors for delivering the best solution rather than producing the best proposal, with competing solutions often assessed based on the results they deliver. Note that a Prize-Based Challenge differs from Impact Bonds and Performance-Based Contracts in that it engages many competing implementers who usually develop a new solution, rather than focusing on improving the impact and funding prospects of implementers' existing programs.

■ Governance flow (agreement) ■ Financial flow ■ Verification of results



## APPENDIX 2: RESULTS-BASED FINANCING INSTRUMENTS

Instrument	Definition	Risk transferred to:	Example
<b>Performance Debt Buy-Down</b>	A third party buys down all, or part, of a loan between a government and a lending organization if predefined results are achieved.	Central government	<b>National HIV/AIDS Prevention and Support Project in Botswana.</b> In 2008, the Government of Botswana agreed to finance efforts to tackle its HIV/AIDS epidemic through a US\$ 50 million International Bank for Reconstruction and Development (IDRB) loan, and concurrently reached an agreement with the European Commission to buy down the IDRB loan contingent on meeting certain targets. The European Commission agreed to provide up to EUR15 million which, if targets were reached, would reduce the effective interest rate on the loan to 0%, and contribute roughly US\$ 3 million towards the principal repayment. Performance targets for the government included increasing condom use and awareness of HIV transmission among young people. <sup>49</sup>
<b>Performance-Based Loan</b>	A development bank provides a loan to a government with disbursements conditioned upon the achievement of predefined results.		<b>Health Millennium Development Goals Program-for-Results Project in Ethiopia.</b> The World Bank supports the Government of Ethiopia's Health Sector Development Program (HSDP) by disbursing credit against achievement of a subset of the HSDP's key objectives. Through the project, the government can receive a maximum of US\$ 100 million, depending on its performance against a set of eight indicators. The indicators include health outcomes, such as the proportion of deliveries attended by skilled birth providers, and outcomes of institutional reforms, such as increased transparency in pharmaceutical procurement processes. <sup>50</sup>
<b>Cash on Delivery / Results-Based Aid</b>	A multilateral agency or a foreign government awards funding to a government if predefined results are achieved.		<b>Poverty Reduction Budget Support to the Government of Sierra Leone.</b> Between 2013 and 2015, the UK's Department for International Development incentivized improved public financial management in Sierra Leone by partially tying budget support to the outcomes of institutional reform. Funding was provided in the form of a fixed annual amount plus an additional payment contingent on reaching targets for a range of financial management indicators, from discrepancies between actual and budgeted expenditure on health, education, and road maintenance to the proportion of tax payers filling tax returns on time. <sup>51</sup>

<sup>49</sup> World Bank (2008); World Bank (2010).

<sup>50</sup> World Bank (2013).

<sup>51</sup> DFID (2013).

Instrument	Definition	Risk transferred to:	Example
<b>Performance-Based Grant</b>	A central government awards resources to a local government if predefined results are achieved.	Local government	<b>The District Development Fund in Lao.</b> The District Development Fund is the mechanism through which the Government of Lao transfers funds from the central level to the district level. Since 2012, amounts transferred to districts have been based on performance in areas, such as planning and budgeting, financial management, and procurement, accountability, and transparency. <sup>52</sup>
<b>Performance-Based Contracts</b>	An outcome payer conditions part of its payment to a service provider based on the achievement of predefined results.	Providers (public or private)	<b>Escalera in Mexico.</b> Escalera delivers high school preparation curriculum kits to public schools one year before high school enrolment to help students make informed decisions regarding high school enrollment. In 2014, Escalera signed a Performance-Based Contract with the Government of Chiapas, which paid for each additional student that enrolled in high school because of the program. <sup>53</sup>
<b>Prize-Based Challenges</b>	An outcome payer opens a competition that awards a predefined financial prize for the best innovation developed in a predefined timeframe.		<b>Greenhouse Gas Emissions Reduction Pilot in Vietnam.</b> AgResults, a partnership between several bilateral funders, the World Bank, and the Bill and Melinda Gates Foundation, is incentivizing innovation to reduce greenhouse gas emissions and increase yields in Vietnamese agriculture by awarding cash prizes to the best solutions as part of a competition. In the first phase of the pilot, selected implementers test their tool, product or agronomic technique, and compete to win prizes of up to US\$ 50,000 linked to emissions reductions and yield increases. In the second phase, implementers scale their solutions, competing to win prizes of up to US\$ 1 million based not only on emissions and yields, but on the number of smallholders reached. <sup>54</sup>
<b>Social Impact Bonds (SIB) &amp; Development Impact Bonds (DIB)</b>	An investor provides upfront capital to a service provider and only gets paid back by the government (SIB) or funder (DIB) if the predefined results are achieved.	Investor	<b>The Educate Girls Development Impact Bond in India.</b> In 2015, Educate Girls, a high-performing service provider, received upfront investment capital from UBS Optimus Foundation to expand education services to about 15,000 children. UBS Optimus Foundation's investment will be repaid with a return by the outcome payer, Children's Investment Fund Foundation, if Educate Girls' intervention reaches targets for enrollment among out-of-school girls and improvement in test scores. <sup>55</sup>

<sup>52</sup> Government of Lao PDR (2016).

<sup>53</sup> Instiglio (n.d.).

<sup>54</sup> AgResults (n.d.).

<sup>55</sup> Instiglio (2015).

## APPENDIX 3: OVERVIEW OF EVALUATION METHODS

The table below describes common evaluation methods and presents the required data and assumptions under which they produce an unbiased estimate of the effect of the program.

**TABLE 6.** Common evaluation methods

Experimental Method				
Methodology	Description	Comparison Group Assignment	Required Assumptions for Causal Attribution	Required Data
<b>Randomized Control Trial (RCT)</b>	Measures the impact of the program by comparing the results of beneficiaries to the results of non-beneficiaries.	Individuals are <b>randomly</b> assigned to the comparison group.	Randomization “worked:” the treatment and comparison groups are statistically identical in relevant observed and unobserved factors.	Data on the results of interest for both comparison and treatments groups at the end of the program.

Quasi-Experimental Methods				
Methodology	Description	Comparison Group Assignment	Required Assumptions for Causal Attribution	Required Data
<b>Regression Discontinuity Design</b>	Exploits the random cutoff that determines whether an individual is eligible to participate, e.g., admission test scores. The method compares the results of beneficiaries and non-beneficiaries close to the cutoff.	Individuals who are close to the cutoff but not eligible to participate in the program.	Individuals just above or below the random cutoff do not differ significantly in relevant observed and unobserved factors.	<p>Data on the results of interest for both comparison and treatments groups at the end of the program.</p> <p>The list of individuals just below and above the cutoff.</p>

Quasi-Experimental Methods				
Methodology	Description	Comparison Group Assignment	Required Assumptions for Causal Attribution	Required Data
<b>Propensity Score Matching</b>	Compares the results of beneficiaries and non-beneficiaries who have been selected because they share similar observable characteristics	Individuals not participating in the program who have a mix of characteristics that predict that they would be as likely to participate in the program as beneficiaries.	The <i>excluded</i> factors (characteristics not considered in the matching exercise because they are unobservable or have not been measured) do not bias results. This is the case if they are either uncorrelated with the measured results or do not differ between beneficiaries and non-beneficiaries.	Data on the results of interest at the end of the program for both beneficiaries and non-beneficiaries.  Data on relevant variables that allow for selecting similar non-participants.
<b>Difference in Differences</b>	Compares the observed change in results of beneficiaries to the observed change of non-beneficiaries.	Individuals not participating in the program, but for whom data on the results of interest has been collected both before and at the end of the program.	If the program did not exist, the two groups would have had identical trajectories over the period of the program.	Data on the results of interest before and at the end of the program for both beneficiaries and non-beneficiaries.

Observational Methods				
Methodology	Description	Comparison Group Assignment	Required Assumptions for Causal Attribution	Required Data
<b>Pre-post analysis</b>	Measures the observed change in the results of beneficiaries.	Observational methods do not include a comparison group that consist of non-beneficiaries.	If the program did not exist, beneficiaries' results would not have changed. That is, the program was the only factor causing a change in results.	Data on the results of interest before and at the end of the program for beneficiaries.
<b>Levels of performance measure</b>	Measures the results of beneficiaries.		If the program did not exist, beneficiaries would not have achieved the measured results.	Data on the results of interest at the end of the program for beneficiaries.

Source: This table is an adaptation of a similar resource provided in Massachusetts Institute of Technology. (n.d.). The complete table can be found in Abdul Latif Jameel Poverty Action Lab. (n.d.), which provides information on other impact evaluation methods.

Table 7. Threats to causal identification describes threats to reliability of the evaluation methods outlined above. Implementers should assess if these factors might constitute a risk in their context and if so, discuss mitigation strategies with an evaluation expert.

**TABLE 7.** Threats to causal identification

Factors	Description
<b>Assumptions for causal attribution are not met</b>	All evaluation methods rely on an assumption associated with causal attribution (see Table 6. Common evaluation methods above). If these assumptions are not met, the selected methods will underestimate or overestimate the true impact of the program.
<b>Attrition</b>	Attrition occurs when results cannot be measured for participants in the treatment or comparison group. This occurs when people move, are not at home when enumerators call, or refuse to cooperate. Differential attrition across groups can result in an over- or underestimation of the program effect. For example, consider a program to improve learning gains for children: If low-performing students in the comparison group are more likely to not show up on test day compared to low-performing students in the treatment group (given a positive impact of the program on attendance), then differences in average learning gains between the two groups include the effects of i) the program and ii) the differential compositions of the two groups. In such a scenario, this would lead to an underestimation of the true impact of the program.
<b>Spillovers</b>	Spillovers occur when the program has indirect effects on individuals in the comparison group. Positive indirect effects increase the results in the comparison group and cause an underestimation of the program effect, while negative indirect effects may cause an overestimation of the program effect.  For example, consider the following: Chicken pox vaccinations are randomized within schools (e.g., half of the students get vaccinated and half do not). The vaccination of the students in the treatment group prevents the transmission of the disease, which reduces the probability of getting chicken pox for the students in the comparison group (spillover). The direct effect of the vaccination can no longer be measured separately from the spillover effects, and we would underestimate the true impact of the program.
<b>Crossovers or contamination</b>	Contamination happens when participants from the comparison group move into the treatment group, or when other service providers or the government introduce a similar program in the comparison group. For example, imagine if parents attempt to move their children from a control school into a treatment school after learning of the implementer's education program. Most likely contamination will result in an underestimation of the true impact of the program.

### Box 16. Why considering statistical significance when making payments may not be a fair way of sharing risks

#### Minimum detectable effect

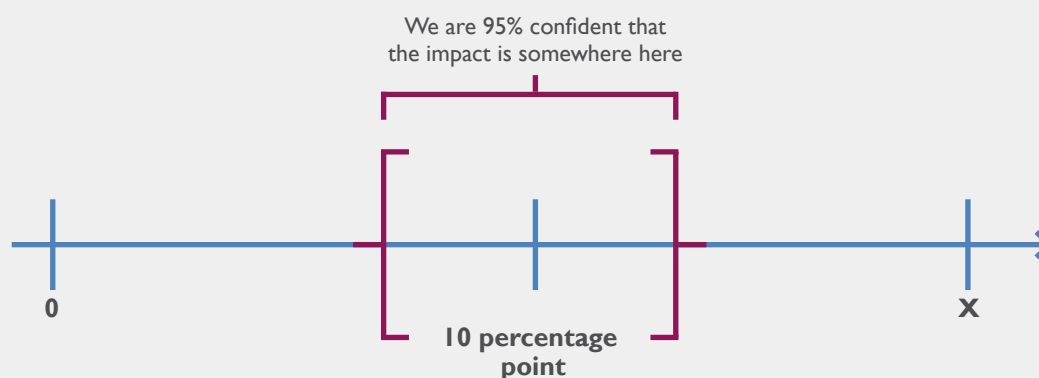
The Minimum Detectable Effect (MDE) represents the minimum level of impact that a certain evaluation design can detect at a certain level of statistical significance. It is primarily determined by an evaluation's sample size. The smaller the impact to be detected, the larger the required sample size. Larger sample sizes typically increase the costs of data collection but also increase the confidence in the estimates and reduce the payment risk for the parties.

#### Statistical significance should not be considered when making payments

Setting the MDE at, for example, a 15 percentage point increase means that any improvements in results smaller than 15 percentage points cannot be identified with the chosen degree of statistical significance. Should this mean that no outcome payments are made if the evaluation finds an impact of 10 percentage point, which will be statistically insignificant?

Statistically speaking, no. Considering statistical significance as a factor for calculating outcome payments may not be a fair way of sharing risks with the implementer or investor. Even if statistically insignificant, the 10 percentage point estimate is still the **most likely estimate** of the program's impact and the MDE was set ex-ante so that a 10 percentage point increase would not be statistically significant.

Consider the graph below which shows the impact estimate (10 percentage points) of a program with a 95% confidence interval. Statistically, we can reject the hypothesis ( $H_0$ ) that the impact is equal to zero. However, we cannot ensure that the true impact is really 10 percentage points, nor can we prove that the impact is different from 10 percentage points plus or minus 1.96 x standard error (the 95% confidence interval). As such, the implementer might reasonably request to be paid based on the higher value of the confidence interval. Similarly, the outcome payer could reasonably request to pay based on the lower value of the confidence interval.



In any evaluation, the outcome payer will never be absolutely certain that they are paying for the true impact (e.g., paying for the true 10 percentage points of improvement), if they overpaid (e.g., paid for 10 percentage points, but the true impact 9), or if they underpaid (e.g., paid for 10, but the true impact was 11). What we can be certain of (under the most commonly used specifications), is that the estimated impact of the program is the impact that is most likely given the available data; hence calling these specifications "Maximum Likelihood Estimations." Even if the estimate of 10 percentage points is statistically insignificant at 5%, 10% or any arbitrary percentage, it is still true that 10 percentage points is the **best estimate**.



## APPENDIX 4: RBF DESIGN NEGOTIATIONS CHECKLIST

The following checklist summarizes key considerations covered in Chapter 3 to help your organization during negotiations in determining whether the RBF agreement is aligned with your organization's desired impact and carries a fair and manageable level of payment risk. Your organization may consider other criteria; for example, the extent to which the RBF agreement helps your organization develop evidence and build capacity, its suitability as a long-term financing mechanism, and the administrative and financing costs associated with the RBF agreement.

- ☐ The **financial size**, and consequently implementation scale, allows for the program to be well-staffed and adapted in a timely manner.
- ☐ The **contract duration** is sufficient for results to be achieved and measured.
- ☐ The **target population** is aligned with your organization's desired impact and experience.
- ☐ The **payment metrics** are collectively closely related to your organization's desired impact, efficiently capturing the most impactful elements of your program.
- ☐ The **payment metrics** do not take away any necessary flexibility to adapt the program in order to improve the desired impact or achieve progress on another payment metric.
- ☐ The **payment metrics and payment function** (including prices and payment caps) do not provide incentives to de-prioritize key results or subsets of the population important to your desired impact, or to conduct activities that are not in the best interest of the beneficiaries.
- ☐ The **amount of funding tied to results** creates effective incentives, transfers adequate payment risk and you are able to pre-finance the portion of the program until payments tied to results are made.
- ☐ The rigor of the **evaluation** is appropriate given the amount of funding tied to results and the vulnerability of results to external factors.
- ☐ The **measurement and evaluation process** is credible.
- ☐ Taking into account your evidence of past performance and changes in the context, target population, and evaluation methodology (among other factors), you are confident that you can reach the **expected level of results**.
- ☐ The likelihood of achieving a lower than expected level of results is acceptable to your organization, and your organization stands to earn a **fair compensation for the risk** you take on.
- ☐ **Measurement** occurs as frequently as possible given the constraint of associated costs, maximizing your organization's ability to learn and course-correct.
- ☐ Your organization can raise the necessary funding to supplement the expected cash flows stipulated by the **payment schedule**.
- ☐ The **contract** does not prescribe any activities that restrict flexibility that would be beneficial to your organization and offers sufficient protection against risks beyond your organization's control that might affect the achievement of results.

## APPENDIX 5: IT SYSTEMS FOR PERFORMANCE MANAGEMENT

The table below identifies some IT solutions that can be used for performance management purposes. This is not an exhaustive list of programs and applications, and Instiglio does not endorse any of these solutions. The best IT system is the one that best fits the program implementation realities on the ground and the IT capabilities of the team.

IT Solution	Data collection and entry	Data management	Data visualization	Key features
Excel	X	X	X	<ul style="list-style-type: none"> <li>• Computer-based data entry</li> <li>• Basic data validations</li> <li>• Offline data entry</li> <li>• Basic data management functions</li> <li>• Basic data visualization functions</li> <li>• Free for nonprofit organizations</li> </ul>
Google Forms	X	X		<ul style="list-style-type: none"> <li>• Computer- and mobile-based data entry</li> <li>• Basic data validations</li> <li>• Basic data management functions</li> <li>• Basic data visualization functions</li> <li>• Online data entry</li> <li>• Free</li> </ul>
Survey CTO	X		X	<ul style="list-style-type: none"> <li>• Mobile-based data entry</li> <li>• Advanced data validations</li> <li>• Offline and online data entry</li> <li>• Data can be downloaded in multiple formats</li> <li>• Basic data visualization functions</li> <li>• Free and paid plans</li> </ul>
Open Data kit	X			<ul style="list-style-type: none"> <li>• Mobile-based data entry</li> <li>• Advanced data validations</li> <li>• Offline and online data entry</li> <li>• Data can be downloaded in multiple formats</li> <li>• Free</li> </ul>
Stata		X		<ul style="list-style-type: none"> <li>• Automated data management</li> <li>• Data can be uploaded and downloaded in multiple formats</li> <li>• Limited data visualization function</li> <li>• Paid plans only</li> </ul>

IT Solution	Data collection and entry	Data management	Data visualization	Key features
SPSS		X		<ul style="list-style-type: none"> <li>• Automated data management</li> <li>• Data can be uploaded and downloaded in multiple formats</li> <li>• Limited data visualizations function</li> <li>• Paid plans only</li> </ul>
Power BI		X	X	<ul style="list-style-type: none"> <li>• Automated data management</li> <li>• Data can be uploaded and downloaded in multiple formats</li> <li>• Advanced and intuitive data visualizations</li> <li>• Offline and online data storage</li> <li>• Free and paid plans</li> </ul>
ETO Social Solutions	X	X	X	<ul style="list-style-type: none"> <li>• Specific for case management</li> <li>• Computer-based data entry</li> <li>• Advanced data validations</li> <li>• Online data entry</li> <li>• Cloud-based</li> <li>• Basic data visualizations</li> <li>• Paid plans only</li> </ul>
Tableau		X	X	<ul style="list-style-type: none"> <li>• Automated basic data management</li> <li>• Advanced and intuitive data visualizations</li> <li>• Data can be uploaded and downloaded in multiple formats</li> <li>• Offline and online data storage</li> <li>• Paid plans only</li> </ul>
Salesforce	X	X	X	<ul style="list-style-type: none"> <li>• Computer and mobile based data entry</li> <li>• Advanced data validations</li> <li>• Online data entry</li> <li>• Advanced data management functions</li> <li>• Data can be uploaded and downloaded in multiple formats</li> <li>• Advanced and intuitive data visualizations</li> <li>• Free for nonprofit organizations</li> </ul>





[www.instiglio.org](http://www.instiglio.org)