ESMAP
Energy Sector Management Assistance Program

TECHNICAL REPORT 005/15

RESULTS-BASED AID IN THE ENERGY SECTOR
An Analytical Guide

ESMAP MISSION

The Energy Sector Management Assistance Program (ESMAP) is a global knowledge and technical assistance program administered by the World Bank. It provides analytical and advisory services to low- and middle-income countries to increase their know-how and institutional capacity to achieve environmentally sustainable energy solutions for poverty reduction and economic growth. ESMAP is funded by Australia, Austria, Denmark, Finland, France, Germany, Iceland, Lithuania, the Netherlands, Norway, Sweden, and the United Kingdom, as well as the World Bank.

Energy Sector Management Assistance Program (ESMAP) reports are published to communicate the results of ESMAP's work to the development community. Some sources cited in this report may be informal documents not readily available.

Written by | Vivid Economics and William Savedoff for Oliver Knight

Energy Sector Management Assistance Program | The World Bank

TABLE OF CONTENTS

List of Figures and Tables

ACRONYMS AND ABBREVIATIONS

BAU	Business-as-usual
DfID	Department for International Development (UK)
EDI	Energy Development Index
GAVI	the Vaccine Alliance
GDP	Gross domestic product
GHG	Greenhouse gas emissions
GPOBA	Global Partnership on Output-Based Aid
HDI	Human Development Index
IEG	Independent Evaluation Group
IT	Information technology
MC	Marginal cost
MCC	Millennium Challenge Corporation
MEPI	Multi-dimensional Energy Poverty Index
MRV	Monitoring, reporting, and verification
NGO	Non-governmental organization
P4R	Program-for-Results (WB)
RBA	Results-based aid
RBF	Results-based financing
REDD	Reducing emissions from deforestation and forest degradation
SE4ALL	Sustainable Energy for All Initiative
T&D	Transmission and distribution
UK	United Kingdom
WTA	Willing to accept
WTP	Willing to pay

All currency in United States dollars (USD or US$), unless otherwise indicated.

EXECUTIVE SUMMARY

This reports considers whether and how to use results-based aid (RBA) to support development objectives in the energy sector. Its objective is to explain the circumstances in which RBA may best be used, as well as to provide practical suggestions to recipients, funders, and development practitioners on how to design an RBA agreement so as to increase the likelihood of success. Although its primary focus is on the energy sector, and this sector provides the bulk of the contextual examples used in the report, much of the analysis is also relevant to other sectors.

In its purest form, three features define an RBA approach. These are:

- **an agreement between a funder and a national or regional government about payments for pre-defined results** (not inputs);
- **an agreement only about a set of results,** with few or no stipulations about inputs or methodology, which gives the recipient freedom in how to deliver the results; and
- **the *verification* of results by an independent third party,** with payments disbursed as soon as possible after results are verified.

The key difference between RBA and results-based financing (RBF) is the recipient of the funding. RBA refers to a program in which a results agreement is made between a funder and a government, either national or regional. Governments typically have a number of different policy levers available to them to achieve results, but are often not directly responsible for implementing projects that deliver those results. RBF, by contrast, relates to agreements with bodies who do not have this range of policy and regulatory levers, but who are directly responsible for the implementation of projects and activities.

HOW IS RBA EXPECTED TO WORK?

RBA relies on a radically different approach to conventional aid programs. RBA aims to support development goals through two channels: first, enhancing the visibility and importance attached to the agreed results within the recipient government and, second, by providing funding to support the delivery of these results. Beyond this, it provides the recipient government with the autonomy necessary to identify barriers and experiment with alternative interventions. Independent monitoring and verification provides feedback on the successes and failures of different approaches, creating an environment conducive to experimentation. It, therefore, aims to increase the effectiveness of aid through exploiting the benefits of recipient adaptation and autonomy. This approach is supported by research which has found aid is more effective when recipients are provided with autonomy.

Four factors need to be in place for RBA to be feasible:

1 | **The recipient must share the funder's desire to improve the result.** A shared commitment by funder and recipient to the attainment of a particular objective is required for any aid program to be effective. The autonomy provided to recipients in an RBA agreement makes

this even more important. While RBA may help to sharpen the incentives for a recipient to deliver the agreed results, it is neither likely to be able to, nor should it try to, fundamentally realign incentives generated by domestic political economy considerations.

2 | **The recipient must have appropriate levels of capacity.** To make maximum use of the autonomy provided by RBA, recipients require a level of capacity that makes it plausible that the results can be delivered cost effectively while respecting appropriate environmental and social standards. At the same time, an enhancement in institutional capacity is often one of the key benefits sought by an RBA agreement.

3 | **The recipient in the RBA agreement must have access to sufficient upfront resources.** For RBA to be feasible, it is important that the lack of upfront funding is not a constraint. Whether this is the case will depend on the scale and ambition of a particular results agreement and the resources available to the recipient. Recipients will typically have a number of different ways of finding these initial inputs, including improving efficiency within the sector, shifting resources from alternative programs within or outside of the energy sector, passing on the upfront costs to other parties, as well as raising additional taxation revenue and/or undertaking greater levels of borrowing.

4 | **RBA is only suitable for long-lived funders with flexible access to capital and who can handle uncertainty over payment amounts.** For the financial support and visibility provided to lead to change, recipients need to be able to rely on the commitments made by the funder, even if there is a lag between the initial agreement and the first results being achieved. This may be a particular challenge for some bilateral donors.

When and whether RBA may be preferred to conventional aid modalities will remain an open question until more evidence is available, but for RBA agreements to be effective, they need to be designed well. The criteria listed above identify some minimum conditions that would need to be in place before RBA can be considered viable. Addressing the question of when RBA might be preferred will only become fully possible when a robust evidence base is available. However, designing RBA agreements intelligently can increase the probability of them effectively delivering on development goals. This report aims to support the robust design of RBA agreements in the energy sector through considering the following issues:

- What result should be targeted in a RBA agreement?
- When might a result be best delivered by a government (RBA) rather than by an implementer (RBF)?
- How should payment level, structure, and baseline be determined?

DESIGNING AN RBA AGREEMENT IN THE ENERGY SECTOR: RESULT AND INDICATOR CHOICE

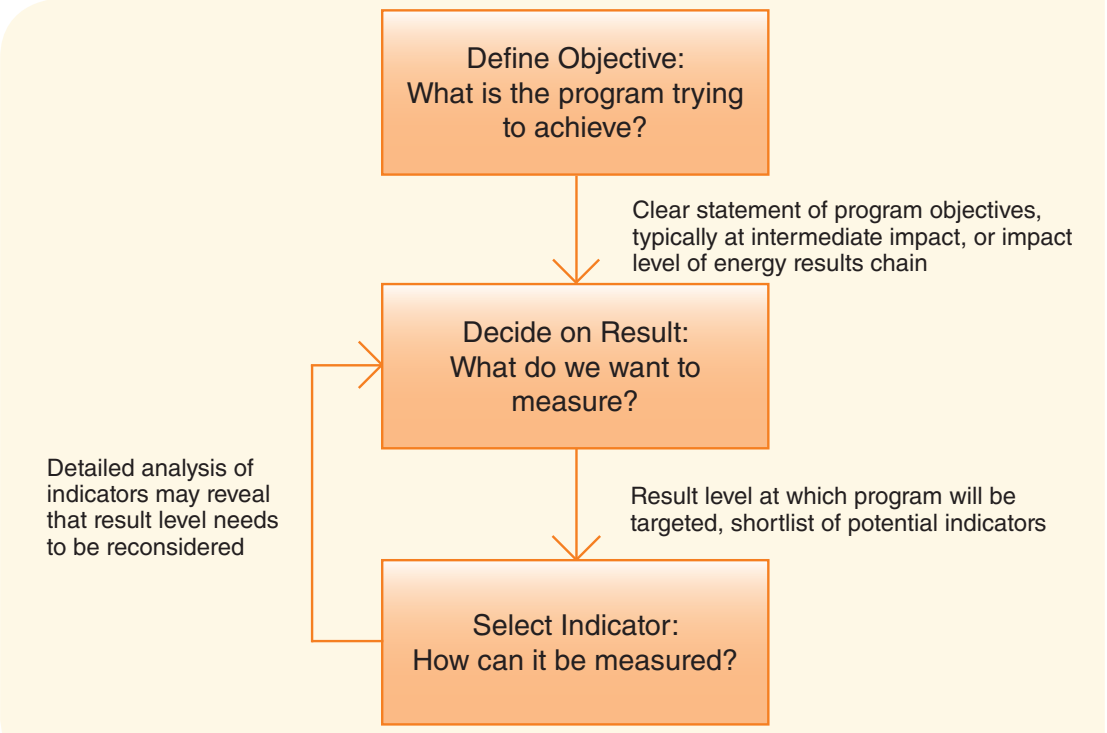
The energy results chain provides a useful tool for analyzing results and indicator choice for energy sector RBA agreements. The results chain provides a clear system for analyzing how interventions in the energy sector are expected to work, by mapping each intervention through a series of steps. Some of the key stages/definitions in the energy results chain are as follows.

- **Inputs:** the resources provided to an energy sector intervention, in terms of both money and time.
- **Outputs:** direct changes within the energy sector expected from the intervention. These include changes in generation, transmission, and distribution of energy, as well as changes in policy and regulation and the capacities of private, public, and nonprofit sectors.
- **Outcomes:** usability of energy services. The usability of energy includes a number of dimensions, such as capacity, duration, reliability, quality, affordability, legality, convenience, health and safety characteristics, and environmental sustainability. An overall increase in usability of energy represents the expected outcome from many energy sector interventions.
- **Impacts:** development goals, such as improvement in education, health, or income levels.

In more sophisticated versions of an energy results chain, there can be additional stages associated with intermediate outputs, intermediate outcomes, and intermediate impacts. The framework and terminology of the energy results chain is used throughout this report.

Using the energy results chain, funders and potential recipients can follow a clear process to identify whether there is a suitable result and indicator that can be used within an RBA agreement. Figure 1 shows the process that can be followed for selecting a result and associated indicator.

FIGURE 1
Three-Stage Process for Choosing Result and Indicator



Source: Vivid Economics

First, the precise objective from the intervention should be specified, and placed within the energy results chain. Funders and potential recipients need to agree on the precise objective of the intervention. For example, the program may seek to reduce the energy intensity of the manufacturing sector by 10 percent. This objective should then be mapped onto the energy results chain to identify whether that objective is an output, outcome, or impact. This can provide helpful clarification and stress-testing; in cases where the objective of the intervention is an output, there are likely to be a number of (possibly implicit) supplementary assumptions about how development impacts will be realized. For example, the objective of a program might be to distribute a certain number of improved cookstoves, with a series of implicit assumptions linking the distribution of cookstoves to development benefits such as reduced exposure to indoor air pollutants and a healthier population. However, these assumptions—such as availability of fuel or cultural acceptance of the improved cookstove—may or may not hold. This stage can help to test how robust such assumptions are and, as necessary, lead to a redefinition of the objective at a higher point in the energy results chain.

The second stage addresses 'what to measure?' in the RBA agreement. By mapping the intervention onto the energy results chain, potential results that could be used to structure an RBA agreement will be apparent at each point in the chain (outputs, outcomes, impacts). Choosing between results at different levels of the energy results chain require a trade-off between three factors.

- **The proximity of the results level to the desired objective.** The chosen result will ideally be closely linked to the desired objective of the intervention. This will mean that satisfactory delivery of the result makes progress towards the intervention's objective (close to) inevitable.
- **The effectiveness of the incentive provided by the result.** A result should be substantially under the influence of the potential recipient. Choosing a result for an RBA where there are a wide number of contingencies beyond the control of the potential recipient will weaken the incentive effect of the agreement. The RBA program will also be less effective at increasing the visibility of the result if it is difficult to communicate; the selected result should offer a simple and well-articulated goal.
- **The ease and costs of monitoring the result.** Results at some levels of the energy results chain are typically easier to measure and monitor progress against than others. For RBA to be cost effective, it is important that results are relatively easy to observe and verify.

The art of RBA design is to find, for a particular objective, a result at the appropriate level of the energy results chain that balances these competing factors. Results at different levels of the energy results chain will typically do better or worse when assessed against these different factors. For instance, output-level results (distribution of cookstoves) are typically easier to observe and verify than outcome or impact results (improved health outcomes). Outcome and impact-level results may also be more difficult for potential recipients to influence (the use of improved cookstoves once distributed will depend on a range of social and cultural factors). At the same time, outcome results will likely be much closer to the appropriately defined objective of the intervention than results at the output level.

It will often be desirable, and it is increasingly becoming possible, to use outcome results in RBA agreements in the energy sector. When the objective of the intervention has been set sufficiently high

up the energy results chain, outcome-level indicators are likely to offer the considerable advantage of being proximate to the desired objective. At the same time, new tracking initiatives in the energy sector (discussed below) offer the possibility of more easily measuring outcome results.

The third stage addresses 'how to measure' a given result. A number of different indicators might purport to measure a particular result. Choosing between those indicators will require consideration of the same three factors: for instance, some results indicators may be much easier to monitor than others, or some indicators may be easier to explain than others. Examples of these issues are provided below. In practice, the choice of result and result indicator will need to be undertaken either in concert and/or iteratively. In a specific intervention, a particular result may appear attractive, but if there are difficulties in finding an appropriate indicator for that result, then an alternative will be needed.

DESIGNING AN RBA AGREEMENT IN THE ENERGY SECTOR: RBA OR RBF?

It is also necessary to assess whether support should be provided to a government (RBA) or an implementer (RBF). This depends on three key factors:

1 | The choice of indicator

2 | The characteristics of the intervention

3 | The characteristics of the (potential) recipients

RBA is likely to be preferable when there is a strong reason to focus on an outcome; RBF is likely to be more appropriate if the targeted result is an output. The successful delivery of the outcome of an intervention depends on a wide range of different factors. For instance, an increase in the proportion of grid connected electricity consumption from renewable sources depends, not only on the successful delivery of additional renewable power generation capacity, but also on a network system operator who is willing to take that power; a reliable transmission and distribution network to deliver that power to households; and a regulatory and policy regime that ensures that there is not a countervailing increase in fossil fuel power generation and consumption. These factors are not all easily within the influence of an individual renewable power plant operator—an implementer—but can be influenced by a regional or national government. Therefore, in cases where there is a strong preference for an outcome based result, it is also likely to be appropriate to structure the agreement as an RBA.

Some interventions will be much more cost effective if a funder works through a national government–and, hence, uses RBA–rather than through a series of implementers in RBF programs. National or regional governments will typically have much better networks within a country than a funder. Improving performance on some results will require action by a large number of smaller implementers; agreeing to separate contracts with each can be costly. Working through a national or regional government, in this case, may be much more cost effective. For example, in programs focused on improving household energy efficiency, the implementers would most likely be individual households. Even if the program was set up to reward improvements in a series of outputs, such as increased penetration of various energy efficient household appliances, the agreement may be best

concluded with a government to exploit the relative ease with which it could reach households. By contrast, when the results only rely on the actions of a small number of implementers, RBF agreements with each individual implementer and RBA agreements with the government are both possible.

The relative capacities and capabilities of different recipients should also be considered. The relevant implementers may have much higher capacity than the government, or vice versa. Given the reliance on the recipient for program delivery, there may be merit in working through the higher capacity recipient. Similarly, access to finance may differ significantly between implementers and different levels of government. This latter consideration may be important, for instance, when achieving a set of results is expected to require a significant investment in capital intensive, grid-connected renewable power capacity: if the national government backing a national utility has limited financial capacity, then the government may struggle to deliver the agreed results (as investors may not consider a power purchase agreement or PPA offered by the utility/government as bankable). It may then be more appropriate to conclude a series of RBF agreements with individual firms. As discussed above, these RBF agreements may need to focus on output results consistent with the desired outcome.

DESIGNING AN RBA AGREEMENT IN THE ENERGY SECTOR: INSTRUMENT DESIGN

The specific design of an RBA instrument involves consideration of at least four further factors:

- The payment type
- The payment level
- The baseline against which improvements are measured
- The conditions under which the agreement is open to renegotiation

The financial instrument used in an RBA can be structured as either a grant or a loan. Under a grant-based RBA, additional grant resources would be provided as results are achieved; under a loan RBA, new loans, that allow refinancing of more expensive debt, could be drawn down as results are achieved. This choice will often be dictated by the nature of the funding organization. To the extent it is not, a grant-based RBA will provide a stronger incentive per currency unit of funding, but the absolute volume of loan resources available from funders will often be considerably greater. This latter consideration may be particularly important in the energy sector given that many results will require significant capital investment.

This report provides new insights on how expected costs and benefits from achieving results should inform the choice of payment level. Crucially, funder and recipient are equal parties to the contract which, given their opposing interests, means that there is no single (theoretically) correct payment amount. Nonetheless, the payment should not be set above the value that the funder places on results, nor below the level that recipient requires to cover the marginal cost of the intervention (broadly understood), taking into account additional sources of income and the financial resources that the recipient can provide. Choosing payment levels becomes more challenging when a results agreement covers multiple indicators. In this case, the difference between payment amounts should

take into account the relative value of achieving improvements of each result. In practice many or all of these variables will be highly uncertain and a process of negotiation will be needed to iterate towards a payment amount acceptable to both parties.

The third feature is the baseline against which improvements in the result are rewarded. Ideally, an RBA should isolate the effects of the program. This is most difficult to achieve, but also most important, when there is a background trend of improvement in the result. In these cases, which might be pronounced in RBAs focused on renewable energy penetration, baselines might be set by reference to existing policy commitments or through using statistical analysis to identify the pre-RBA trend in the result.

Finally, attention is needed to design an RBA that is future-proof. There may be cases where either the funder or recipient may wish to renegotiate an existing RBA agreement. In the case of the funder, this renegotiation may either be not to make a payment when a payment is due according to the terms of the agreement or to make a payment when the result has not been achieved. If recipient countries are able to predict the behavior of the funder then they may reduce their efforts to deliver the result in the first place. Features can be incorporated into RBA design to mitigate this risk, including diversifying the portfolio of RBA agreements and having a credible and legitimate alternative way to spend unused RBA resources.

APPLYING RBA IN THE ENERGY SECTOR: USING THE MULTI-TIER FRAMEWORK TO STRUCTURE AN RBA AGREEMENT

Historically, it would have been difficult to design a robust RBA to promote household electricity access. Household electricity access has conventionally been measured in relatively simple, binary terms. In particular, the percentage of households in a country with an electrical grid connection has been widely used. This output result is a long way from the objective of most interventions aimed at tackling household energy access. For instance, it does not consider whether the electrical grid provides high quality, reliable electrical supply. It also ignores the possibility that off-grid electricity solutions might provide a similar or better level of quality and reliability. As such, it would be a poor indicator to use in an RBA. Given the absence of an alternative, this would likely render an energy access RBA inappropriate.

The multi-tier framework for measuring household access to electricity offers a unique opportunity to link aid payments to outcomes. The multi-tier framework measures the usability of electricity supply along multiple dimensions through representative household surveys. These dimensions are aggregated into a tier score for each household that ranges between 0 (no electricity access) to 5 (high quality accessibility of electricity). In the energy results chain, each of the dimensions of the usability of household electricity supply is an intermediate outcome result indicator, while the tier score for a household is an outcome result indicator. They are much closer to the objective of most household electricity access interventions than alternative measures of energy access, such as household connection rate.

There are three main options for using multi-tier framework for measuring household access to electricity within an RBA agreement. These can be assessed using the approach set out above.

- **Payments could be tied to changes in the average tier score.** For instance, $5,000 could be disbursed for each 0.1 increase in the index score in a particular country or region. This links well to how the indicator may be used in the UN's Sustainable Energy for All initiative, but, as it pays no attention the distribution of households across different tiers, it may not always be close to the objectives of the funder and recipient.
- **Payments could be tied to individual tier improvements.** For instance, $200 could be disbursed for any household moving from Tier 1 to Tier 2, $150 for any household moving from Tier 2 to Tier 3, and so on. This allows the reward structure to be more carefully targeted, potentially allowing for greater proximity to funder and recipient objectives. However, it would somewhat complicate presentation and explanation of the program, possibly compromising the strength of the incentive effect.
- **Payments could be tied to individual dimensions of electricity access.** For instance, a payment of $100 could be made for each household that experiences improved quality of supply. This approach provides a simple, easy-to-explain incentive that could be used to target particularly salient barriers. However, it has poor proximity to impact, as it falls short of a holistic measure of electricity access.

In many cases, the second approach, tying payments to improvements in individual tier improvements, will be the most attractive.


80's
ROCK!

1

INTRODUCTION

KEY FINDINGS

- Results-based aid (RBA) is the provision of financial support to a national or regional government based on the achievement of independently verified results.
- This contrasts with results-based financing (RBF), which provides financial support to implementers or service providers rather than governments.
- RBA is yet to be applied at scale in the energy sector, but could potentially offer good value for money, improved ownership by recipients, and increased accountability.
- This report explores the opportunities for RBA in the energy sector. It seeks to help practitioners understand when it could be effective and how it should be designed.

CONTEXT AND OBJECTIVE OF THE REPORT

Results-based approaches are becoming increasingly important, in both developed and developing countries. With government and agency budgets under pressure and with growing emphasis on accountability, tying disbursement explicitly to achievement of independently verified results has become increasingly popular. In developed countries, results-based schemes have been used to structure mental health services, maternity services, and offender rehabilitation schemes, among other initiatives. The approach is also being increasingly adopted in bilateral and multilateral aid programs, with initiatives such as the Global Partnership on Output- Based Aid (GPOBA) and the World Bank's Program for Results (P4R) making extensive use of results-based agreements.

Within the energy sector, there are already several large-scale results-based schemes that make payments to implementers and service providers. Results-based schemes that focus on paying implementers or service providers are generally referred to as *results-based financing* (RBF). For instance, EnDev has developed a RBF facility that is operating across Sub-Saharan Africa and South Asia, offering payments to firms who achieve results in improving access to modern energy services. GPOBA has used results-based programs to promote rural electrification and renewable energy, with examples including programs supporting solar development in Bangladesh.

However, support to national or regional governments based on the achievement of results in the energy sector remains unusual. Results-based schemes that focus on paying governments, whether national or regional, are referred to as results-based aid (RBA). Potentially, similar programs could offer support to other public bodies, such as rural energy agencies. In contrast to RBF, the option of using results-based payments to governments or other public bodies remains very rare.[1]

RBA schemes are being implemented in many other sectors. There has been particular focus in the health and education sectors. For example, and as discussed in Box 1.1, the UK Department for International Development (DfID) is implementing an initiative of this sort to promote education outcomes in Ethiopia.

This report seeks to explore the opportunity of utilizing RBA within the energy sector. Although the analytical framework has been developed with the intention of being applicable within any sector, the examples throughout the report will be focused on the energy sector and the report deals with a number of sector specific issues, such as measuring energy access.

It complements earlier work by Vivid Economics and the Energy Sector Management Assistance Program (ESMAP) on RBF. The report constitutes one output in a broader research agenda on results-based approaches in the energy sector being undertaken by ESMAP of the World Bank. In particular ESMAP previously commissioned a report on the role of RBF (Vivid Economics, 2013). The earlier report explored the issues around applying results-based schemes when payments are made to implementers or service providers, but paying governments raises a wide range of additional or different issues that will be discussed in this report. For instance, paying governments broadens the range of indicators that you may consider tying to payment levels. The conditions under which RBA might be most effective are different to the conditions under which RBF may be appropriate.

This report is intended to provide a high-level guide to results-based aid in the energy sector. It should be useful for a wide range of interested parties, including recipients, donor agencies, development practitioners, and academics. By providing a clear and justified analytical framework, the report ensures that the analysis is rigorous and open to critique. Given the paucity of RBA programs within the energy sector, unfortunately, there is limited experience and evidence to draw upon. Nonetheless, the report has integrated relevant experience where possible, and draws together the findings to form a series of steps that will be useful to practitioners considering or designing RBA programs.

WHAT IS RESULTS-BASED AID?

Results-based aid is an aid modality where disbursal is linked to the achievement of pre-defined 'results'. RBA has two key features. First, payments are provided after the achievement of particular, pre-agreed, and independently verified results. Second, the recipient is free to determine how to achieve the results. For instance, a traditional aid program within the energy sector may provide $10 million to be spent by the recipient government on off-grid energy systems. An RBA program, by contrast, would start with the recipient and the funder reaching an agreement whereby the funder agreed to provide, say, $200 for every new household provided with reliable, high-quality electricity access, regardless of the system type. The agreement would specify the result to be incentivized, the amount provided for achievement of the result, as well as the way in which achievement will be measured. The RBA instrument would then leave the recipient country free to decide how to achieve this goal, secure in the knowledge that they will receive funding upon the verified achievement of the agreed result.

BOX 1.1

Existing Results-Based Aid and Similar Initiatives

A number of initiatives outside of the energy sector have either fully or partially trialed the use of RBA approaches. One of the most notable examples is a DFID initiative–inspired by 'Cash on Delivery' Aid—to promote education in Ethiopia. This program pays the Ethiopian government between £50 and £100, depending on region and gender, for each child sitting/passing exams. The GAVI Immunization Services Support program also has many RBA characteristics, although, to date, independent verification of results has not been mandatory. Other programs, such as the World Bank's Program-for-Results (P4R), also approximates the above definition, although as it requires some degree of oversight and monitoring of inputs, this also departs somewhat from the 'hands-off' philosophy of RBA.[2] Likewise, the Millennium Challenge Corporation (MCC) has also sought to increase the accountability and contingency of aid, but the program differs from RBA as defined above: there is no specific agreement between funder and recipient tying quantity of aid to the achievement of particular results.[3] A summary of the extent to which a selection of existing programs conform to the definition of RBA used in this report is provided in the table below.

P4R and MCC's interventions share many features of RBAs, but do not conform to all of our criteria for a 'pure' RBA

'PURE' RBA	CASH ON DELIVERY AID (E.G., DFID EDUCATION PROGRAM IN ETHIOPIA)	PROGRAM-FOR-RESULTS	MILLENNIUM CHALLENGE CORPORATION	GAVI IMMUNIZATION SERVICES SUPPORT
Agreement tying disbursal to achievement of pre-defined results	✓	✓	X	✓
Focus on results, with no limitation on inputs	✓	X	✓	✓
Independent verification of results	✓	✓	✓	X

Note: P4R places limitations on the methodology that recipients can use to achieve results, and therefore is not fully consistent with the second criterion. MCC grants a compact to countries that perform well against verified indicators, but does not directly tie the amount of aid disbursed to results. GAVI Immunization Services Support program encourages but does not require independent verification of results.

Source: Vivid Economics

This report adopts the following working definition of RBA, following earlier work by Klingebiel (2012a):[4]

- an agreement between a funder and a recipient 'aggregator' (normally a national, regional, or municipal government) about payments for pre-defined results (not inputs);
- an agreement *only* about a set of results, with few or no stipulations about inputs or methodology, which gives the recipient freedom in how to deliver the results; and

- the *verification* of results (which are initially measured by the recipient) by a third party, with payments disbursed as soon as possible after results are verified.

There are a number of initiatives that are consistent with most or all of these criteria, discussed in Box 1.1.

STRUCTURE OF ANALYTICAL GUIDE

The subsequent chapters of this guide follow the structure illustrated in Figure 1.1. The report can be roughly divided into three sections, seeking to answer three interlinked questions:

1 | When will energy sector RBA be effective?

2 | How should energy sector RBA programs be designed?

3 | How can RBA be applied to promote energy access?

Chapter 2 considers the theory of change underlying results-based aid, and uses this to derive four pre-conditions for RBA to be effective. If these pre-conditions are absent, RBA is unlikely to lead to development benefits and other approaches should be considered instead. The theory of change also provides some guidance on when RBA is likely to be more or less effective, which is especially valuable given the limited empirical evidence available at this stage.

Chapter 3 provides a discussion of the energy results chain, which can be used to frame the selection of results level and indicator. The energy results chain developed by ESMAP allows for the mapping of most interventions in the energy sector through a series of steps in a way that provides a clear system for analyzing the theory of change in many energy interventions. It provides a useful tool for discussing different RBA design options.

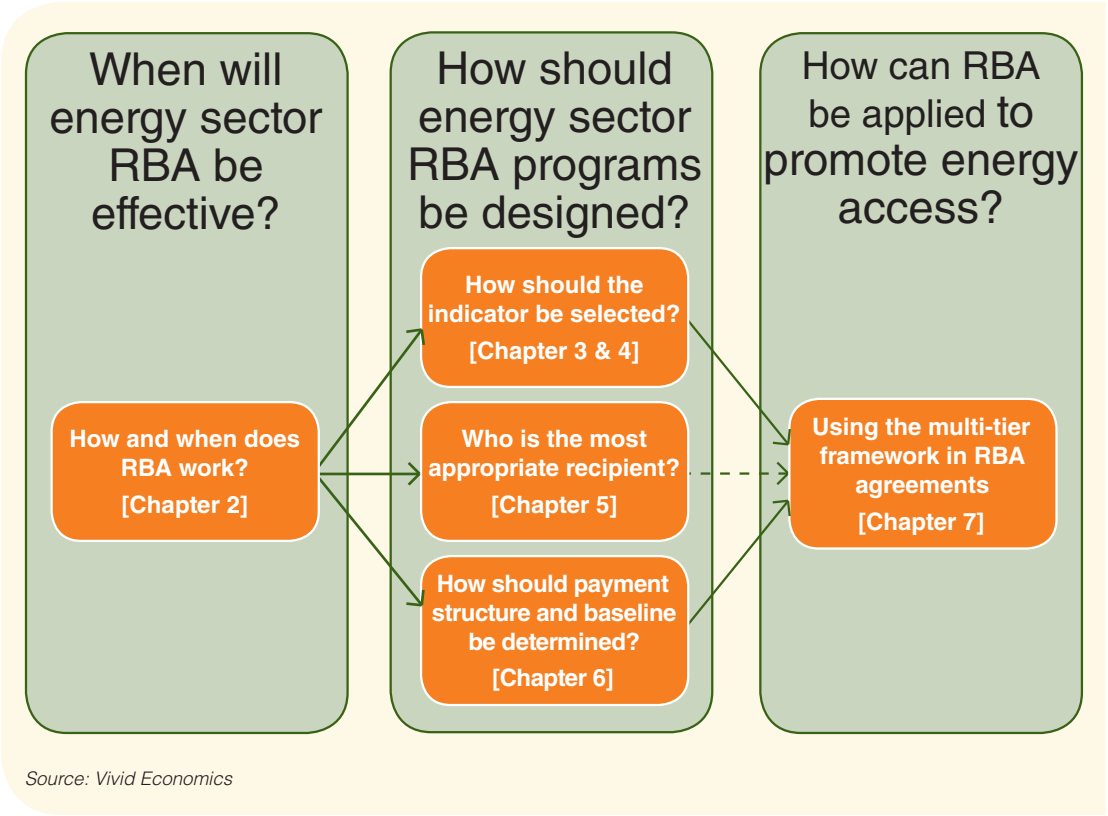
Chapter 4 outlines the factors to consider when selecting a results level and indicator for RBA. RBA is only likely to work well if an appropriate indicator is available. This indicator should be closely related to the objectives of the program and relatively easy to measure. If no such indicator is available, it may be better to consider conventional approaches to structuring an aid agreement.

Chapter 5 considers when it is appropriate to target governments as the recipient in an aid program, rather than implementers or service providers. Some goals are likely to be best pursued by working directly with those who deliver results on the ground, rather than high-level state actors. For instance, state actors may have weak influence over the result or extremely limited access to finance.

Chapter 6 discusses the remaining key design issues within an RBA agreement. This includes setting the payment type, payment structure, payment level, and the baseline against which progress is measured.

A final section, Chapter 7, considers the intersection between RBA and ongoing work on tracking and measuring energy access under the Sustainable Energy for All initiative. While this report is intended to be relevant to the use of RBA in relation to any aspect of the energy sector, promoting

FIGURE 1.1
Structure of the Report



Source: Vivid Economics

energy access is a key development priority: under the UN Secretary General's Sustainable Energy for All Initiative (SE4ALL), the international community has pledged to achieve universal energy access by 2030. Recent pioneering work linked to this initiative focuses on developing a more robust framework for measuring energy access. To provide greater understanding of how RBA approaches might be adopted in the energy sector, Chapter 7 looks at how the household electricity access components of this framework might be used within an RBA agreement.


Forward

2

HOW AND WHEN DOES RESULTS-BASED AID WORK?
RBA Promotes Autonomy, Ownership, and Accountability

KEY FINDINGS

- Results-based aid provides recipient governments with assistance, both financial and promotional, to support the achievement of a shared goal.
- The recipient government is given the autonomy necessary to identify barriers and experiment with alternative interventions, while independent monitoring and verification provides feedback on the successes and failures of different approaches.
- Provided that the recipient has the capacity to improve performance and internal incentives within government to support achievement of the goal, RBA programs could be effective in achieving development goals.
- The recipient will need to be able to access at least some finance to fund the initial stages of the program.
- The funder must be able to deal with both long-term commitments and uncertainty over annual disbursement amounts.

HOW DOES RESULTS-BASED AID WORK?

RBA relies on a radically different theory of change to many conventional programs.

Conventional programs often depend on identifying an obstacle or set of obstacles that prevents the achievement of a goal, and then designing interventions that specifically tackle these obstacles. RBA, on the other hand, is intended to create the space and support necessary for recipient governments to identify and design interventions themselves, adapting their approach to changing conditions on the ground.

For instance, the design process for a conventional aid program may have the following steps:

1 | Articulate the goal or objective that the program seeks to promote. For instance, an energy program may seek to reduce greenhouse gas emissions from industry.

2 | Specify the high-level problem preventing this goal from being achieved. Energy intensive industry may have limited uptake of energy efficient processes.

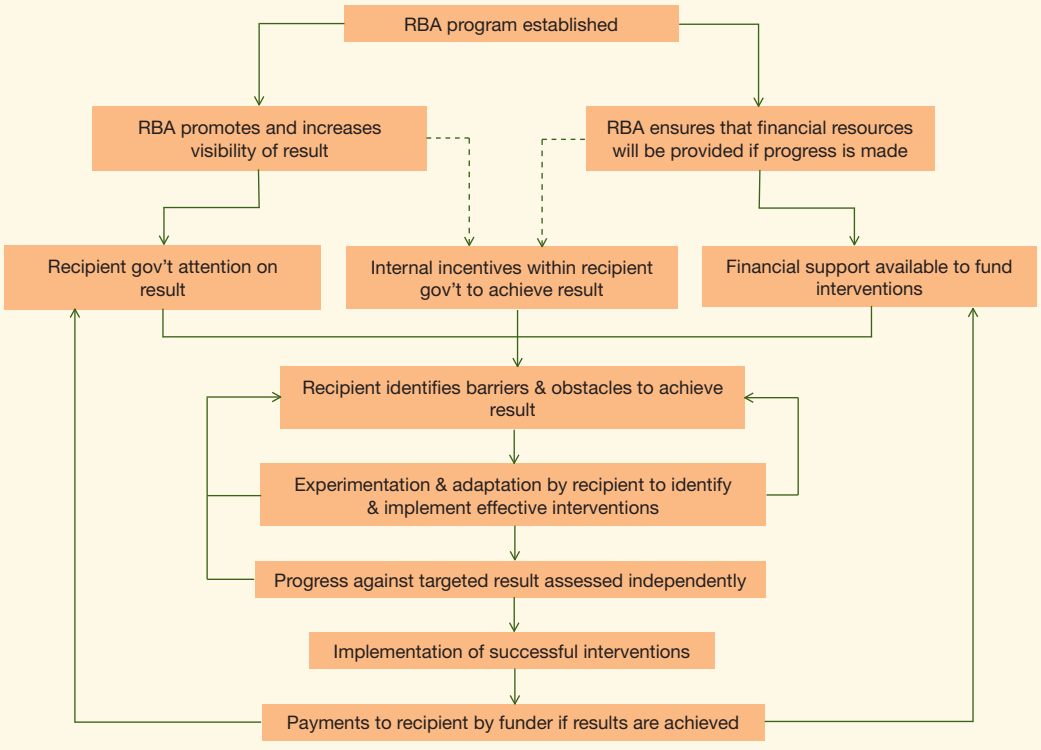
3 | Identify specific obstacles, such as market or government failures, that contribute to this problem. Industry may be unaware of the existence of alternative production techniques and underestimate potential financial savings. Information on potential cost-saving measures may diffuse slowly due to poorly developed local supply chains and low technical capacity.

4 | Design an intervention or set of interventions that tackle these obstacles. Providing technical assistance, in the form of energy audits, free of charge to interested companies could ensure that they are well informed about potential cost savings from adoption of energy efficient production techniques.

RBA, on the other hand, leaves the identification of obstacles and the design of interventions to the recipient government. From the funder's perspective, the process enumerated above need not proceed beyond the second step. The third and fourth steps are left to the recipient. A simplified and generalized theory of change, presented as Figure 2.1, shows the process by which this is intended to create the conditions for achieving program objectives.

The RBA program promotes visibility of the result within the recipient government. Specifying a simple, well-articulated, and measurable goal can focus uncoordinated efforts throughout departments. In this sense, the RBA program could perform a similar role at the national or regional

FIGURE 2.1
Three-Stage Process for Choosing Result and Indicator



Source: Vivid Economics

scale as the Millennium Development Goals performs at the global scale. Also, by ensuring that progress against the result will be closely and independently monitored, the RBA program may increase oversight on the departments responsible for delivery. Information about their performance will be widely available for civil society organizations, other branches of government, and the wider public. The possibility of being called to account for the success or failure of the program, therefore, may increase the scrutiny and attention that the result receives internally.

It also provides the financial support necessary to fund interventions aimed at improving performance against the result. Payments from the RBA program could form a substantial source of revenue for interventions aiming to improve performance against the targeted result. Verified improvements act as an additional revenue stream alongside commitments by the recipient government and other donors. This additional revenue stream will support interventions that otherwise could not go ahead and allow successful interventions to be scaled up once their effectiveness has been verified.

These two effects (increased visibility and greater funding) could create the conditions for achieving the program's goals, if combined with a strong internal incentive structure encouraging government action on the result. As discussed above, the existence of an RBA program will draw attention to the result that is being promoted and the financial reward will mean that better performance will increase government budgets, both of which may incentivize action. However, governments do not act to maximize profit. For a number of reasons, discussed extensively in Chapter 2, "When Will Results-Based Aid Work?," the incentive provided through the RBA program may not filter down to the individuals required to take action within government. Therefore, it is important that the drive to improve the targeted result is already strong within the recipient and that internal incentives are aligned in a manner that makes progress possible. The RBA program can then provide the financial support and attention required to bolster internal efforts.

With greater funding and visibility attached to results, and where internal incentives are aligned, the recipient can pursue the targeted result, identifying obstacles and experimenting with interventions to remove these obstacles. Funders are not involved in the approval of intervention plans, nor are payments conditional on making particular policy changes. Rather, the recipient has full autonomy and ownership over the measures that they want to adopt. Recent research based on a data set of 14,000 projects conducted by international development organizations has found that aid is consistently more effective when recipients are provided with autonomy, especially in difficult environments (Honig, 2014). Similar results have been found in other contexts, such as within the Nigerian Civil Service (Rasul & Rogger, 2013).

Progress against the targeted result is verified independently. A third party is contracted to measure the indicator to which payments have been tied (and, potentially, to perform the monitoring and evaluation for the program as a whole). Monitoring and verification should occur regularly enough to allow rapid disbursement and feedback to the recipient. In combination with autonomy, this creates the ideal conditions for experimentation and adaptation. Recipients can respond to the changing barriers that emerge as the interventions develop, and will be able to track where and how they are succeeding.

If performance improves, as measured by the contracted indicator, payments are disbursed to the recipient. Disbursement may further reinforce the attention on the result and could provide the financial resources necessary to extend successful interventions. This may result in a virtuous circle, with success drawing in attention and resources, leading to further success. On the other hand, if performance does not improve, the financial exposure of the donor is limited and the recipient is provided with detailed information on the performance of the program, allowing them to adjust interventions or identify and tackle further barriers.

ADVANTAGES AND DISADVANTAGES OF RESULTS-BASED AID

The theory of change outlined above helps identify some of the possible advantages of RBA. Proponents of RBA have argued that it can deliver a number of benefits, most of which have been discussed above (Birdsall & Savedoff, 2011; Perakis & Savedoff, 2014). These include:

- RBA may make aid flows more effective by focusing attention on the delivery of results, rather than paying for inputs, and by creating a conducive environment for adaptation and experimentation;
- RBA may increase the ownership of development programs among recipient countries and, in so doing, could promote institutional development and increase the accountability of recipient governments to their citizens;
- RBA may reduce the time and resources that development partners typically spend monitoring processes and checking that resources are being spent appropriately; and
- RBA may reduce the risk of corruption as resources are only disbursed when results are delivered.

On the other hand, there are several potential disadvantages of RBA, compared to conventional programs. The disadvantages could include:

- RBA may increase the total cost of the program, because of the high cost of borrowing for the recipient;
- RBA may require greater expenditure on monitoring and verification; and
- RBA may be less effective than conventional approaches if the recipient lacks the capacity, incentive, or access to capital necessary to deliver.

Costs of borrowing will often be higher for recipient governments than funders, implying that using an RBA could increase the total costs of the program. Recipient governments will generally have worse access to finance than funders. Raising pre-finance to fund the interventions used to achieve program goals will require borrowing at high interest rates. As a result, the total costs of interest will be higher than if the funder borrowed on behalf of the recipient and provided the funds directly. These increased costs are a necessary consequence of the responsibility for delivery lying with the recipient; they may be more than outweighed by the benefits of increased autonomy and ownership, which funders may be unwilling to allow unless the recipient bears financial risk.

Monitoring costs may be higher under RBA, though the costs of monitoring under conventional programs can also be significant. Regular, independent monitoring and verification can be costly,

especially when payments are tied to program outcomes requiring household surveys. The monitoring requirements of RBA could be expected to push up total costs. On the other hand, the value of regular monitoring goes far beyond enabling payments to be based on results. Recipients and funders will be able to track the successes and failures of the program, adapting interventions and approaches based on high-quality information. Accountability will also be enhanced by the availability of good data. Finally, conventional programs often include conditionality based on inputs, such as the provision of receipts that demonstrate particular purchases have been made. The cost of monitoring and verifying these inputs can also be costly, and provides much less useful information than the monitoring of outputs and outcomes.

Whether RBA should be preferred to conventional approaches depends on these potential disadvantages being outweighed by the benefits, but the current evidence base is too weak to provide strong guidelines on when RBA will be most effective. Without more evidence it is difficult to determine where RBA is likely to be more or less advantageous. The literature review provided in our earlier report on results-based financing (Vivid Economics, 2013) summarizes the limited evidence available so far. It is likely, however, that RBA will perform better in some sectors than in others and that different ways of structuring RBA (different indicator choices, for instance) will have implications for how well it works. Unfortunately, the evidence base at the moment is weak. Wider experience is necessary before it is possible to state confidently the conditions under which RBA should be preferred.

However, the theory of change can be used to identify several preconditions that are likely to be required for RBA to be effective. The next section outlines these preconditions. If they do not hold, RBA cannot work as envisaged and adopting a conventional approach may be preferable.

WHEN WILL RESULTS-BASED AID WORK?

The theory of change implies four conditions should hold if RBA is to be effective:

- The recipient government wants to improve performance against the result;
- The recipient government has sufficiently high capacity to design and implement effective interventions;
- The recipient government has access to pre-finance to fund interventions; and
- The funder is trusted to fulfil the terms of the contract.

Recipient Desire to Promote the Result

The success of aid programs depends on the recipient government's commitment to the shared result. The development aid literature has regularly demonstrated that recipient government commitment–often called "ownership"–is key to the effectiveness of aid (see, for instance, Dollar & Svensson, 2000, and Booth, 2011). But conventional approaches to aid are not very good at predicting when a recipient government is truly committed. RBA programs "reveal" recipient

government commitment to the goal by disbursing against progress *ex post* and will therefore channel funding where it is being effective.

RBA differs from RBF in terms of how recipients could be expected to respond to incentives. RBF is targeted at service providers and implementers, who can reasonably be expected to pursue results provided that the incentive payment exceeds the technical costs of the intervention (minus any additional revenue sources that may be available). For private companies, this allows them to maximize profit. For non-governmental organizations (NGOs), this ensures that the incentive payment is high enough to cover their costs while allowing them to pursue a socially beneficial goal. RBA, on the other hand, explicitly targets government actors. These actors are not primarily motivated by monetary incentives, even if the payment exceeds the technical cost. Instead, committed governments will view the RBA payments as a flexible subsidy for doing what they intended, while uncommitted governments will have little interest in RBA programs.

Decision-makers responding to an RBA will be embedded in a system of internal political incentives that may prevent the pursuit of 'profitable' goals. Suppose that an RBA is targeted at a particular government department. The individuals (civil servants, political appointees) within that department will ultimately be accountable for the achievement of political goals, defined by the domestic political economy environment, and associated institutions. Maximizing department revenue, by pursuing the results defined in an RBA, may or may not be beneficial for achieving these goals. Even if the technical costs of the intervention are below the incentive payments, the intervention may still distract limited human or financial capital from the pursuit of more politically salient objectives. This effectively raises the marginal cost of the intervention above the technical cost.

Further, departments targeted through the scheme may not be able to retain revenue, and it may not be possible to make payments to departments that are charged with overseeing the program. Treasuries may reduce central budget support in response to the receipt of RBA revenues by another department. In some countries, foreign funding will need to be channeled through a central department who could choose not to provide it, in turn, to the department achieving results. Either of these institutional factors may undermine the incentive effect for the individuals involved in coordinating the program. Alternatively, the incentive may be enforced indirectly, through pressure from the Treasury or through the prestige attached to generating additional revenue.

Finally, recipient desire to promote the result is essential to ensure sustainability. A recurring concern is that RBA programs will promote the pursuit of low-cost and unsustainable approaches to achieving results. To some extent, this may be dealt with by contractually requiring minimum standards for the quality of infrastructure; for instance, a program seeking to promote solar lanterns could require the Lighting Africa minimum quality standards are met for all lanterns purchased. However, such approaches are likely to either limit recipient autonomy or be costly to monitor. In the solar lantern example, monitoring the quality of purchased lanterns implies that: (i) the program is required to improve access to lighting through solar products and (ii) the program is required to purchase and distribute the solar products itself, rather than supporting their proliferation by encouraging growth in private markets. Both aspects limit autonomy, and highlight that input monitoring will only be well suited

to programs that have heavy involvement in purchase, installation, and implementation. Recipient ownership and commitment provides a much better guarantee of sustainability—if the recipient government wants to promote the result for their own, internal, political reasons and are dedicating their own, internal, resources to do so, they are much less likely to seek to 'game' the scheme by artificially inflating results.

These considerations all suggest that RBA will be effective when it targets a mutually agreed result and reinforces, rather than undermines, accountability. The financial structures and incentives within a government department are likely to be complex and the amount of money involved in an RBA is unlikely to override them. Indeed, attempts to fundamentally alter incentives in government risk weakening domestic accountability. Given this, a shared goal should be chosen as the subject of an RBA agreement. Additionally, both parties require a deep appreciation of how political economy considerations might mediate the incentive impact of the RBA. This may indicate, for instance, that RBA payments will need to be set higher than technical costs to function as an 'attention grabber' and generate interest within the Treasury, or that the regional departments need to be targeted separately from national departments in response to federalized political systems.

Recipient Capacity

Funders may find that RBA is unsuitable for use if agent risk is very high. Agent risk refers to the risk of a recipient reneging, failing to meet the specified results, or otherwise behaving in a manner that the funder believes is inappropriate. A key feature of RBA is that it places the responsibility for delivery of results onto the recipient, so low levels of competency can have a significant impact on the effectiveness of the program. While RBA may allow funders to be insulated from the *financial* impacts of an ineffectively delivered program, there are latent reputational risks that cannot be avoided. For instance, if a multi-year RBA program failed to make any disbursements, funders may be scrutinized by domestic legislatures and other stakeholders on why they adopted a program that failed to deliver development gains. There have been examples of criticism of this kind within REDD+ initiatives (a form of RBA aimed at reducing deforestation), where disbursements are widely perceived to have been slow (see Caravani et al., 2013; CIFOR, 2012). To a significant extent, this has been due to difficulties in establishing measurement systems and other problems with 'readiness', leading to a several funds and programs being established with an explicit focus on building capacity.

A particular concern may be the potential for violation of the rights of individuals within the recipient country. For instance, a recipient government may react to the introduction of an RBA scheme rewarding increased renewables generation by fast tracking the construction of a large hydroelectric power plant. Without the proper safeguards, this could potentially result in the displacement of large numbers of people, destruction of valuable habitats, loss of land necessary for the livelihood of the local population, and so on. The possibility of rights violations forms a developmental risk, as the gains in wellbeing resulting from the RBA scheme may be more than offset by the damage caused through the hydroelectricity project. It also forms a reputational risk for the

funder. To some extent, the danger can be mitigated through intelligent design of results agreements; in the example described above, renewable energy generated through hydroelectricity could be excluded from the agreement or the validity of the agreement could be conditional on favorable human rights monitoring reports. The danger of rights violations will also vary depending upon the type of result being promoted through the program. Ultimately, however, the source of this concern is high agent risk, and the funder needs to be able to trust that the recipient will not violate basic safeguards in the delivery of programs supported by aid.

These concerns may be particularly prevalent while RBA is still in the early stages of development, suggesting caution regarding agent risk is advisable. As noted above, examples of the application of RBA within the energy sector are particularly limited. Until there is experience to draw upon regarding the design and implementation of RBAs, caution on the choice of agents may be particularly prudent. This would provide a means to experiment with different design aspects and collect preliminary information concerning the schemes' effectiveness without exposing either funder or recipient to significant risk. Screening for eligibility should occur with reference to clear, objective criteria, ensuring that decisions are transparent.

At the same time, it should be noted that one of the advantages of RBA is to encourage development of institutional capacity in the recipient country. There are three main ways in which RBA can improve institutional capacity:

1 | **By giving recipients discretion over design and implementation,** RBA can create the opportunity for recipient countries to take ownership over the program and exploit their greater knowledge of the local context. This can create room for experimentation, innovation, and development of policies that are well suited to the country in question.

2 | **By requiring independent verification of outcomes,** RBA increases transparency. Widely available and reliable information on the successes and failures of the program allows local stakeholders to hold recipients and funders to account, encouraging institutional improvement.

3 | **By generating accurate and reliable measurements for outcomes,** RBA programs also create an opportunity for feedback regarding which policies and strategies make a real difference.

Technical assistance should be available for the recipient if desired and can help deal with capacity limitations, but it needs to be demand led. Making provision of funding conditional on accepting technical assistance or on involvement in the design of interventions will limit the (real or perceived) autonomy of the recipient. If foisted on an unwilling recipient, it is also unlikely to be effective. Instead, technical assistance should be available if required, but provided upon request of the recipient. Similarly, providing a roadmap for recipients (for instance, by specifying inputs required for disbursal) may seem to help with limited capacity, but will also obstruct autonomy and achieve little more, in terms of capacity building, than is possible with demand-led technical assistance.

High agent risk will have an adverse impact on the effectiveness of other types of aid; the key issue is which approach manages this risk better. Conventional instruments, such as budget or

project support, are also less likely to be successful in the context of higher agent risk. Similarly, while RBA risks the manipulation of results data to attract extra disbursements, conventional approaches risk the manipulation of input records such as invoices. The key issue is whether funders consider that conventional support allows them to manage these risks more effectively. However, although individual circumstances will vary, in cases where agent risk is very high, many funders may prefer the possibility of greater hands-on management provided by upfront support.

Access to Pre-Finance

In order for an RBA instrument to be a success, the recipient will need to have access to sufficient financial resources to cover upfront costs prior to the delivery of results. There are many potential sources of pre-finance. Sometimes, for instance, enhancing the efficiency with which existing inputs are used will allow an increase in outputs or outcomes. Indeed, one of the key attractions of RBA is that it allows exploitation of the recipient government's knowledge about where such efficiency gains are most abundant. Alternatively, it may also be possible to use mandates or regulations to pass down even a relatively large funding requirement to numerous implementers—each of whom individually can meet their portion of the funding needs through their own capital resources. This, for instance, is a typical approach used by many governments to increase renewable energy generation. However, when additional public resources are needed, this will require either surplus government revenues, access to capital markets, or the capacity to raise additional tax revenue. The importance of this barrier will depend on the difficulties of increasing efficiency and on the scale of the upfront costs that need to be incurred, as well as on the duration between incurring costs and producing measurable results. In particular, smaller, less capital intensive, and faster to deliver programs may face fewer challenges.

Depending on the program being considered and the recipient country in question, this could rule out the use of RBA. Countries that are excluded from borrowing in capital markets, or can only do so at punitive rates, will be limited to their domestic tax base, international aid or, possibly, borrowing (on concessional terms, as appropriate) from multilateral development banks. For resource intensive programs, this may render a 'pure' RBA infeasible. In such a situation a funder and recipient may wish to discuss the use of an alternative instrument, or combine an RBA with a significant tranche of upfront support.

Even if pre-finance is available for the government as a whole, it is necessary to consider the ability of the recipient line ministry, regional, or municipal government to secure finance. An RBA may be targeted at a particular line ministry or regional government. In this case, access to finance may be limited even if the government has good access to capital. For instance, the priorities of the line ministry may conflict with the priorities of the finance ministry or the government as a whole, and the line ministry therefore may experience difficulties with securing sufficient funds. This reinforces the importance of understanding the political economy context, as emphasized above.

Pre-finance is typically less of a concern within RBA than with RBF. The ability to access finance can often be a significant concern at the level of an individual implementer or service provider.

Governments, especially at the national level, will invariably have some access to financial resources even though different governments will benefit from different levels of flexibility. Even if a government has little capacity to raise additional taxes and no access to international capital, they will be able to reroute expenditures from elsewhere, although the feasibility of achieving this rerouting in a particular political economy context needs to be established. The key difference is the scale of the required investment relative to the recipient's capacity: within RBF, the necessary investment might require a significant portion of the recipient's existing revenues, whereas this will rarely be the case with RBA.

Funder Ability to Commit to Contract

RBA is only suitable for long-lived funders with good and flexible access to capital. The primary institutional requirement for the funder is the capacity to credibly commit to disbursing funds following the achievement of results. Given that RBAs are considered to be most effective when framed as long-term agreements, over five or more years (Birdsall & Savedoff, 2011), this constraint may rule out many potential funders. There are two primary concerns:

1 | **The funder must be able to provide funds soon after the achievement of results is verified.** This requires the funder to have flexible access to sources of funds.

2 | **The funder must be able to credibly commit to honoring long-term commitments.** This is particularly relevant in the context of bilateral government agreements. Political change in the funder country may undermine international aid agreements, regardless of the intentions of the original parties to the agreement. Constitutional arrangements may also make multi-year commitments, even by the same administration, challenging. Without a credible commitment the recipient government may reasonably doubt whether funds will be forthcoming, which could discourage them from investing in effective interventions. As explored in Chapter 6, "Avoiding Time Inconsistency," various aspects of the design of an RBA instrument could be used to enhance their credibility.

The funder must also be able to handle uncertainty over annual disbursement amounts. Disbursal will vary year to year depending on performance. This can create difficulties for some funders, especially donors with annual disbursal targets. Similarly, donors may need to structure programs to disburse annually and, as a result, can have difficulties with programs where disbursement needs to occur less frequently. This could be the case if, for instance, measurement of the result requires household surveys that are too expensive to conduct every year.

In the absence of these conditions, it may be preferable to use a more conventional modality or investigate alternative delivery vehicle options, such as multilateral institutions, that can provide the required flexibility. To some extent, these difficulties can be overcome if the program uses creative financial structures; one option may be the creation of a trust fund that receives regular and fixed amounts from the funder each year, and then disburses to the recipient dependent on results. Balances could either be returned at the end of the program or used to fund other results-based schemes.

DESIGN AND EFFECTIVENESS OF RESULTS-BASED AID

The preconditions discussed above determine whether RBA could be successfully applied, but the success of a specific RBA program also depends on its design. Some contexts and development goals will be poorly suited to any results-based approach, no matter how well designed. However, in many cases, the suitability of RBA will depend upon the specific design choices taken. Choosing the correct indicator, baseline, and payment amount could make a huge difference to the effectiveness of the program.

Crucially, RBA depends on the availability of an indicator to use as the basis for disbursement. In some instances, suitable indicators may not be available or creating the necessary monitoring frameworks may be prohibitively expensive. Identifying an indicator, therefore, is essential to determining whether RBA is suitable for use within a particular context. Chapter 3 explores this issue, providing guidance on what considerations should affect indicator choice and what features make an indicator unsuitable.

The suitability of RBA also depends upon how appropriate it is to approach a developmental problem through providing support for government rather than implementers or service providers. Some developmental problems are best tackled by directly interacting with service providers, rather than working through government. This could be because of the particular barriers and market failures that are preventing development or it could be because of the competencies and capacities of different government bodies and service providers on the ground. Chapter 5 discusses the features that determine whether working with service providers may be preferable to an RBA approach.


d.light
8 hrs

3

ENERGY RESULTS CHAIN AND FRAMEWORK FOR CHOICE
A Framework For Mapping Interventions

KEY FINDINGS

- The objectives of energy sector interventions in developing countries tend to fall into at least one of three categories: improving energy access, promoting renewable energy, or increasing energy efficiency.
- Energy interventions explicitly or implicitly rely on a theory of change, linking a change in inputs (such as investments in the energy sector) with developmental impacts (such as low carbon economic growth).
- The energy results chain, developed by ESMAP, can be used to understand the chain of cause and effect required for an intervention to achieve its objective.
- The classification of different indicators into input, output, outcome, and impact levels is used throughout the report to assist with the design of RBA programs.

This chapter identifies the typical overarching objectives of many energy sector interventions and explains how many of the interventions that aim to meet these objectives can be analyzed through the energy results chain, developed by ESMAP. This provides a framework and language that is utilized in the rest of the report.

ENERGY SECTOR OBJECTIVES

Most energy sector interventions primarily focus on one of three broad objectives:

- Improving energy access for household, productive, and community uses
- Promoting renewable energy
- Increasing energy efficiency

Many interventions in the energy sector seek to increase energy access for household, productive, and community uses. The set of potential interventions intended to promote this objective is extremely wide, ranging from improving the availability of appliances (cookstoves, televisions, etc.) to building large-scale generation and network assets. Grid extensions and off-grid generation projects can also be used to improve access, as can a wide variety of education, awareness, private sector development, stakeholder engagement, and community empowerment projects.

There is parallel focus on interventions that aim to reduce greenhouse gas emissions from the energy sector through renewable energy. In many cases, the expansion of renewable energy offers further benefits such as improved energy security. Renewable technologies can be used for

power, heating, cooling, and cooking. As with energy access, the scope for potential interventions is very large, ranging from directly funding renewables capacity to providing credit for local financial intermediaries to encouraging development of private markets in solar lanterns.

Some interventions also focus on improving energy efficiency. In many cases, interventions to increase industrial (including the energy sector), transport, and residential efficiency are pursued as a means of reducing greenhouse emissions. However, they can also be targeted for other reasons, including a desire to enhance energy security or to boost international competitiveness. The range of interventions is as diverse as those discussed above and may include improving credit availability for industrial investments, providing information about energy saving opportunities, or reducing transmission and distribution losses.

Other objectives are also relevant, but tend to be promoted through achievement of the three objectives identified above. Energy programs may seek to build domestic supply chains, though usually this is achieved through a focus on either renewable energy or energy efficiency. They may intend to reduce deforestation, normally by improving energy access or promoting (non-traditional biomass) renewables. They may try to increase gender equality by reducing the time spent on the (typically gendered) task of fuel collection through increased access to modern energy sources. Other objectives are also possible, but most interventions within development will fall into at least one of the three categories above. The rest of this report, will address these three categories for practical reasons, though the principles that are outlined may be applicable to other objectives.

LOGICAL FRAMEWORK FOR ENERGY INTERVENTIONS

Interventions seeking to improve energy access or increase use of renewable energy can be mapped into the logical framework provided by the energy results chain, shown in Figure 3.1.[5] The energy results chain provides a clear system for analyzing the theory of change implicit in energy interventions. It splits each intervention into a chain of inputs, output, outcomes, and impact, providing a means of identifying the (implicit) assumptions about cause and effect within each intervention, as well as a way of understanding how measurement, monitoring, and evaluation should take place.

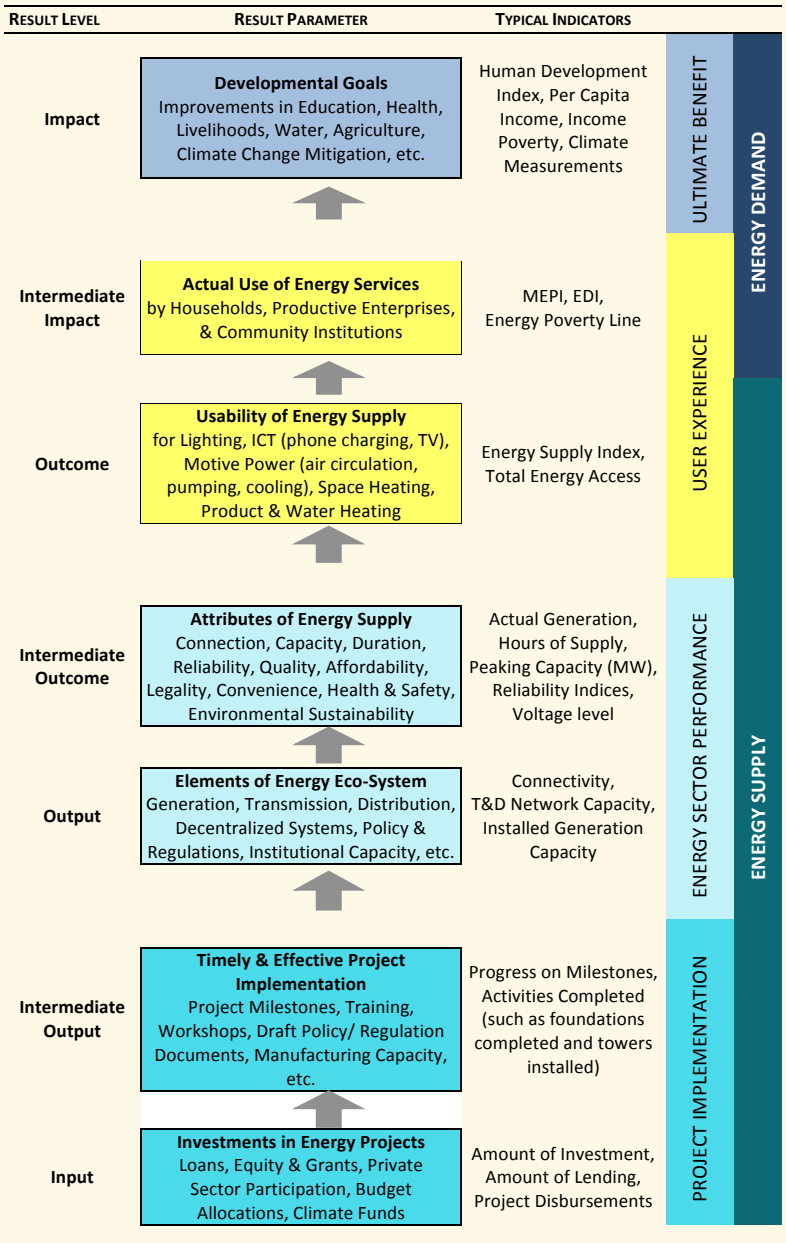
As detailed in Figure 3.1, an energy intervention in the areas of energy access or emission reductions can be broken down into as many as seven stages.

All interventions start with a series of 'inputs'; investments in energy projects, in terms of both money and time. Funds may be used to directly issue loans, equity, and grants to support the development of energy infrastructure, or budget may be allocated to fund technical assistance, training, or other activities.

If implemented and planned well, the use of these inputs will allow achievement of project milestones, referred to as 'intermediate outputs'. Intermediate outputs may include any instrumental goals of the project that do not directly involve changes in the broader energy ecosystem. For instance, they may include the production of draft policy documents, improvements in manufacturing capacity for appliances, the successful running of workshops or training activities, the identification of investment opportunities, or the production of business plans.

FIGURE 3.1

The Energy Results Chain is Split into Seven Levels



Source: ESMAP

As a result of effective implementation, the project should produce 'outputs' in terms of changes in the energy ecosystem. The energy ecosystem includes the infrastructure involved in generation, transmission, and distribution of energy, as well as the wider institutional framework within a particular country or region. Changes in policy and regulation are included as outputs, as are changes in the capacities of private, public, and nonprofit sectors.

The changes in outputs will be expected to affect attributes of the energy supply, which are 'intermediate outcomes'. For instance, increased renewables capacity may both reduce the carbon intensity of energy supply and increase the available capacity on the grid, leading to fewer unscheduled outages. However, this is not a foregone conclusion; it depends upon the renewables capacity being utilized, operational, and a sufficiently robust transmission and distribution infrastructure. Other attributes of energy supply include the number of households with electricity connections, the affordability of energy, the convenience of energy supply, the legality of energy access and the health and safety associated with supply.

These attributes jointly determine the usability of energy, which is the 'outcome'. The usability of energy depends on all of the attributes mentioned above. A high capacity, safe, and legal electricity connection to the grid, for instance, will not increase the usability of energy unless it is available when required and not subject to frequent, unscheduled outages. Similarly, improvements in access to biogas do not lead to better usability unless the relevant appliances are available and functioning.

Improved usability of energy should cause more use of energy services, by households, communities, and productive enterprises. The actual use of energy depends on both the underlying usability of the supply and the preferences of energy users. Even a highly usable supply with many desirable attributes may not be utilized if consumers are unwilling to shift their behavior or if it is unsuitable for their needs. Increases in the actual use of energy are the intermediate impacts of the intervention.

Finally, greater use of energy services should result in a development 'impact', such as improvements in education, health, livelihoods, water, agriculture, or climate change mitigation. This impact is (or should be) the ultimate goal of the program. Many factors beyond the control of a particular program will influence performance at this level. Nonetheless, it is important to bear in mind that the ultimate goal of any development program is to improve the wellbeing of people, and there should be a clear idea of how improving use of energy is likely to lead to this.

As the effects of a program are traced to higher levels in the results chain, attribution becomes more difficult. As mentioned above, results further from inputs may be hard to attribute to the program. If the program is involved in making direct subsidies or investments, outputs can often be traced fairly reliably to the increased expenditure and purchases by the program. For instance, a program distributing solar lanterns can directly trace the increase in number of lanterns available to consumers to purchases made by the program. But even this step may be difficult in some programs that work through the private sector, as it can be unclear as to whether some of the renewables projects funded through the program would have happened in the absence of additional funding. By outcome level, reliable attribution will require considering many of the other factors that could be influencing usability

of energy supply. Observed changes in the usability of electricity could be the result of new capacity constructed through a renewables program, but may also result from simultaneous changes that are occurring in regulation or distribution that are unrelated to the program. At the impact level, changes in, for instance, education will be a function of many factors much broader than just energy outcomes, and establishing the effect of the program will be difficult.

To demonstrate how an intervention can be mapped onto the results chain, Table 3.1 shows potential examples of each results level for a cookstove program.

TABLE 3.1

The Energy Results Chain Can Be Used to Map Out the Theory of Change Implicit in a Cookstove Program

RESULT LEVEL	DESCRIPTION	EXAMPLES IN A COOKSTOVE PROGRAM
Input	Investments in energy projects	– Investment in purchasing cookstoves – Loans or subsidies provided to cookstove suppliers, distributors, or users – Technical assistance provided to businesses or government to support market development
Intermediate output	Timely & effective project delivery	– Providing $50,000 in loans to distributors – Training 200 local manufacturers
Output	Elements of the energy ecosystem	– Improved cookstoves available to 10,000 additional households
Intermediate outcome	Attributes of energy supply	– Improved health and safety with lower indoor air pollution from cooking – Better access to fuel, with less time required for collection
Outcome	Usability of energy supply	– Better usability of improved cookstoves, due to improvements across a number of attributes (access to fuel, quality, reliability of heat, etc.)
Intermediate impact	Actual use of energy services	– Increased incomes and wellbeing due to time savings from fuel collection
Impact	Developmental goals	– Increased use of improved cooking solutions – Increased health due to less indoor air pollution

Source: Vivid Economics


ENERGY,
THAT

4

RESULTS LEVEL AND INDICATOR CHOICE
The Indicator Should Balance Proximity to Objective, Ease of Measurement and Appropriate Incentive Effect

KEY FINDINGS

- Choosing the correct indicator for a RBA program is a key design decision, as bad indicators can undermine the effectiveness of the program.
- The indicator chosen should be closely linked to the programs objective, relatively easy to measure, and provide an appropriate incentive effect.
- The process of choosing an indicator can be broken down into three steps.
  1 | Define the objective of the program. For many energy sector programs, the objective will be to improve access to energy, promote renewable energy or increase energy efficiency. Having a clear objective in mind helps select an indicator with proximity to impact.
  2 | Choose the results level that the indicator will measure. Output indicators are likely to be easier to measure but harder to link to impact. Impacts are generally very hard to measure (or at least attribute to the program) but are closely linked to the objectives of the program. Outcomes balance each of the three requirements, and are often the best choice for RBA.
  3 | Select an indicator for the chosen results level. This indicator should perform well against the three criteria identified: proximity to impact, ease of measurement, and appropriate incentive effect.

If it appears that a country would be suitable for a results-based approach, it is then necessary to identify how a results-based scheme could be designed and, in particular, the result and indicator against which payment might be made. If it is not possible to find an appropriate result, satisfying various key criteria, and appropriate indicators for that result, then it will not be possible to proceed with an RBA scheme.

This chapter provides an analytical framework for this process. This design can be split into three steps, shown in Figure 4.1:

1 | Defining the objective of the intervention

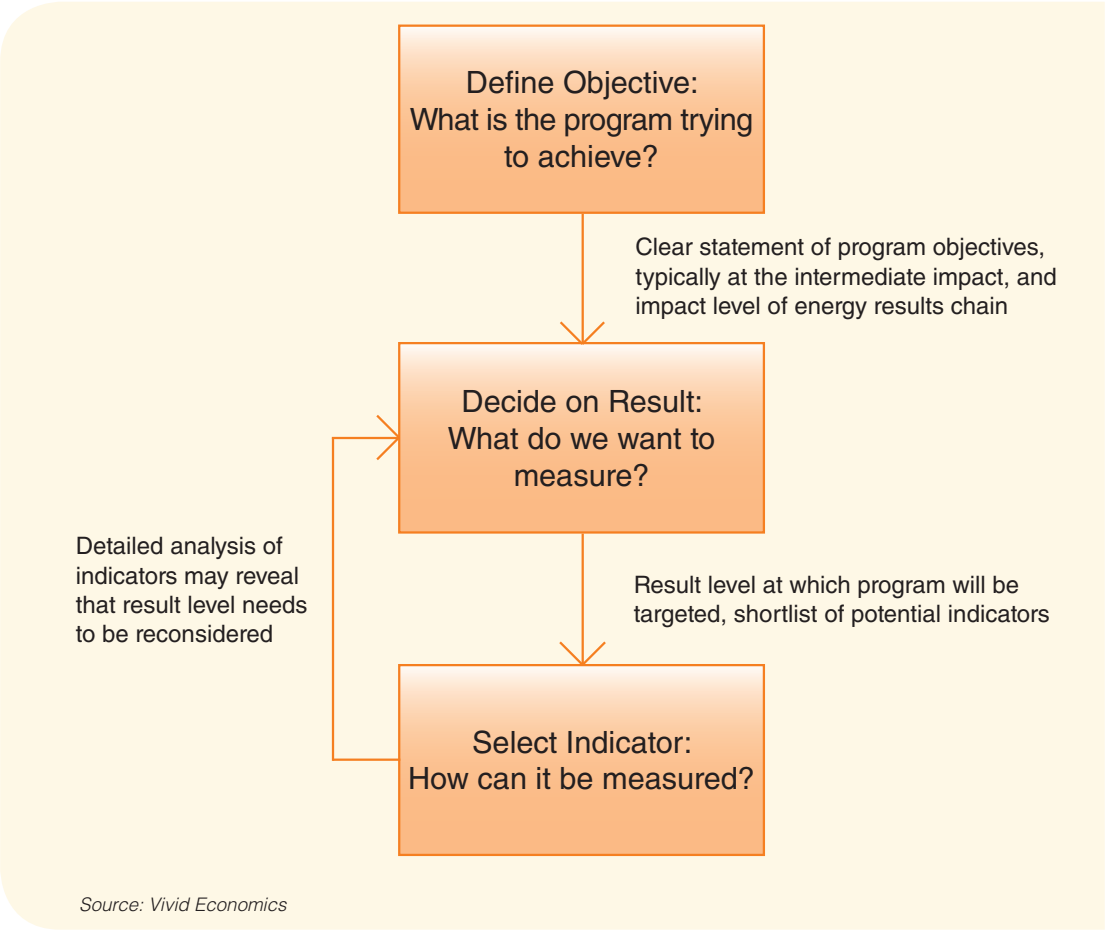
2 | Deciding on the results level and associated result to measure

3 | Selecting an indicator for this result level

The choice of result and indicator cannot be undertaken independently of a consideration of the likely recipient. One result and indicator may be very suitable in an agreement for one recipient but inappropriate for another recipient. Given this, this chapter proceeds with a presumption that

FIGURE 4.1

Three-Stage Process for Choosing Result and Indicator



Source: Vivid Economics

those designing a possible RBA scheme have a particular recipient in mind and that this recipient is a national, regional, or local government. It then applies the framework of this section to that particular recipient. Chapter 5, discusses some of the factors that might want to be taken into account to check that the choice of recipient is appropriate.

DEFINING OBJECTIVE

First, it is necessary to clearly define the objective of the intervention. To structure a results-based intervention, it is first necessary to determine what the intervention is trying to achieve. For these purposes, it can be helpful to refer to the three objectives outlined in Chapter 3—promoting energy access, increasing use of renewable energy, and improving energy efficiency. Many interventions

within the energy sector will have objectives that fall into one of these three categories, though they often have a tighter sectoral focus and may feature other objectives. For instance, an intervention may be seeking to improve residential energy efficiency, reduce the carbon intensity of grid electricity, or improve household access to energy for cooking.

It can be helpful to locate the objective on the results chain; is the intervention targeting outputs, outcomes, or impacts? While intervention objectives are often stated in terms of outputs (e.g., installation of capacity, provision of appliances), the ultimate objective of an intervention should ideally be to create changes at the impact and intermediate impact level. For instance, an intervention may have the stated goal of distributing 10,000 solar home systems or installing 30,000 biodigesters, but ultimately, in most cases, the intention is to improve access to low carbon energy—which requires both installation and use by households.

If the objective of the program appears to be at the output level (changes in the energy ecosystem), it is worth considering whether these goals are relying on the assumption that changes in output will lead directly to development impacts. Changes in the energy ecosystem do not automatically lead to increased use of energy services, and the increased use of energy services does not necessarily create the envisaged development benefits. Often, there will be other factors that will intervene. Installed cookstoves could be unused because of lack of fuel availability or because they do not match user needs, for instance. Increased capacity on the grid may be prevented from affecting user experiences by poor transmission or distribution infrastructure. Programs that state objectives in output terms are likely to actually intend to change the use of energy and, therefore, create development benefits, but may be relying on a set of implicit assumptions regarding how changing outputs will lead to impacts. It is worth explicitly confronting these assumptions before thinking about how a results-based scheme would be designed, and questioning whether the ultimate objective of the program is really at the output level. In most cases, it will be appropriate for objectives of the program to be specified at the higher levels of the energy results chain.

DECIDING ON RESULTS LEVEL

Once the objective has been clearly defined, it is possible to consider which result, and at which level of the results chain, should be targeted. Payments could be tied to results at the output, outcome, or impact level. The advantages of choosing results at one level or another will depend on the specific context, and each will have different implications for the advantages of the scheme over conventional modalities.

There are three main factors that determine the appropriate result level to target:

1 | The proximity of the result to the desired objective

2 | The effectiveness of an incentive based on the result

3 | The costs of monitoring the result

Each will be explored in more detail in this section. Broadly, targeting a result level closer to impact is beneficial because it ensures that measured progress implies that the objective

(and broader development goals) of the intervention are likely to be achieved. Effectiveness of incentive requires that the result provides a clear and well-articulated goal that the recipient has the ability to influence, which enables better promotion of the result throughout government. Finally, costs of monitoring should be modest to ensure that regular and accurate monitoring is possible and both parties trust the independently verified data to represent the reality on the ground.

When choosing a result, it is necessary to consider the indicators available. The result and result level cannot be chosen entirely in the abstract, without considering the indicators by which that result may be measured. To make concrete trade-offs between different design options, it will be necessary to think about the indicators that could be used for results at different levels of the results chain, and the costs of monitoring each of them. This report, therefore, suggests an iterative process; first, consider the advantages of disadvantages between targeting different results at different levels of the results chain, then consider the indicators available for that result. The information gathered on indicators may influence result level choice.

Nonetheless, there are some considerations that are likely to hold regardless of indicator. Table 4.1 outlines the advantages and disadvantages of targeting different results levels. It also discusses potential remedies to some of the problems identified in relation to the criteria at different stages of the results chain.

Proximity to Objective

The tighter the causal link between the results and the objective of the intervention, the more effective an RBA is likely to be. Considering this factor alone, an RBA would be linked directly to the desired objective and that objective would be at the impact level of the energy results chain. This would ensure that payments would only be made as a result of improvements in the wellbeing of the targeted population. However, as explored below, impacts are often difficult to observe, unsuitable for quantification, and changes are often impossible to attribute to a particular program. At the impact level of the results chain, it may also be more difficult to be confident that the process followed to achieve the results is appropriate. Given these disadvantages, it may be necessary to compromise on a result level further from the objective. Nonetheless, the proximity to objective remains important and, when linking an RBA to a result level, it is essential to consider how tightly linked the result and objective are.

For example, the most commonly used and widely available indicator for measuring electricity access is the rate of household connection to electricity, but this output indicator fails to capture many important aspects of electricity access and non-electricity energy needs. It is rare that either funders or recipients would consider that an appropriate objective of a program is to increase the number of household electricity connections per se. Instead, they tend to care about increasing the availability of electricity services for households. The use of electricity services depends on a wide variety of additional factors beyond the presence of a connection, including:

- Capacity (peak power) and quality (voltage) of electricity supply

- Daily duration of electricity supply, especially during evening hours
- Reliability of electricity (frequency and duration of unscheduled outages)
- Affordability of electricity
- Legality of electricity supply
- Ownership of electrical appliances
- Ongoing consumption of electricity

To determine the proximity to objective, it is useful to explicitly identify the assumptions linking the result level to the objective. In many cases the chain of assumptions linking an output result and the intended objective of the intervention will be relatively long, and certainly longer than the assumptions linking outcome and intended impact. These assumptions should be tested for plausibility to assess proximity to impact. If it seems plausible that one or more of these assumptions may fail to hold, RBA using a result from this level of the results chain is much less attractive.

When there is significant distance between the result level and desired objective, the use of results-based approaches will also be more likely to create misaligned incentives. For instance, in incentivizing the maintenance and repair of infrastructure is important for many interventions within the energy sector. However, directly targeting and rewarding repair work (an intermediate input) is likely to lead to misaligned incentives. Recipients may seek to perform unnecessary maintenance on operational equipment, allowing them to collect the payment but not achieving development impacts.

Measurement

Results that are close to the desired objective can be prohibitively expensive to monitor. There are three main elements to take into account: ease of measurement, detectable changes over time, and attribution. Often when considering each of these elements, it will be important to consider not only the result itself but also the indicator(s) that could be used to measure that result.

First, the ease of measurement for different results levels will vary, with outcome level often the most problematic. Output-level results and associated indicators can often be collected on the supply side, but outcome-level results and indicators may need to be measured through end-user surveys. This requires more resources, in terms of time and money, and can also introduce additional possible sources of inaccuracy and potential for manipulation of results or gaming.[8] Fortunately, measurement of many impact-level indicators is already in place but, in the energy sector, outcome-level measurement is rare. Tying to outcome may be more costly.

In addition to directly reducing the costs of monitoring, easily observable results and associated indicators are particularly important in results-based approaches as they will increase the cost-effectiveness of a result through increased recipient confidence. There is a renewed emphasis in the monitoring of results across all forms of development co-operation (The Paris Declaration on Aid Effectiveness and the Accra Agenda for Action, 2008), which might imply that the costs of monitoring are of no greater importance in designing results-based approaches than in other aid modalities.

TABLE 4.1

Relationship between Result Level and Three Criteria for Choice of Results Level for an Energy Access or Renewable Energy Program

RESULT LEVEL	OUTPUT	INTERMEDIATE OUTCOME	OUTCOME	INTERMEDIATE IMPACT	IMPACT
Result Parameter	**Elements of Energy Ecosystem:** Assets delivered or policies implemented	**Attribute(s) of Energy Supplied:** Characteristics of energy supply that affect the user	**Usability of Energy Supply:** Aggregate effect of all attributes that make energy usable for desired applications	**Actual Use of Energy Services:** Consumption of energy to avail the desired applications	**Development Goals:** Achievement of developmental goals through enhanced use of energy for applications
Indicator Examples	Capacity installed, connections provided, solar lanterns delivered, laws/ policies/regulations	Connection capacity, reliability, hours of supply, quality, legality, health & safety, convenience, environmental sustainability	Increase in number of households at a given tier of the multi-tier energy access metric, increase in energy access indices	Consumption of electricity, ownership/use of appliances, indices based on consumption or appliance ownership (MEPI or EDI)	GDP growth, increase in household income, reduction in GHG emissions, HDI
Proximity to Objective					
Effectiveness[6]	**Low**	**Low**	**Medium**	**Medium**	**High**
	Difficulty: Outputs may not automatically lead to developmental goals if causal linkages are weak	**Difficulty:** Identified attributes may not be the only constraint to increased use of energy for developmental goals	**Difficulty:** Targeted level of usability of energy may be inadequate, appliances may be out of reach, causal chain to developmental impacts may be weak	**Difficulty:** Delivery agencies may not be able to identify downstream causal chain and may lack funds for investment. Developmental goals may not be achieved despite increased energy use	**Difficulty:** Causal linkages with energy may be difficult to establish upfront. In most cases, energy may not be the only deficient factor. Funds are not tied to energy initiatives
	Remedy: Apply output based RBA only where the causal chain from outputs to developmental goals is assured either by nature of energy supply technology or by program design	**Remedy:** Apply attribute based RBA only where deficient attribute has been clearly identified, and further links in the causal chain are well established.	**Remedy:** Careful design to target the correct level of usability of energy and confirm access to appliances. Upfront establish evidence of causal chain to developmental goals	**Remedy:** Careful analysis of expected interventions by delivery agencies and availability of funds for investments. Upfront establish evidence of causal chain to developmental goals	**Remedy:** Use Development Goal based RBA in rare cases where energy is the sole barrier to the achievement of developmental goal. Programs may not be classified as energy interventions

TABLE 4.1 *Continued*

Measurement					
Ease of Measurement (Time, effort, & resources + Risk of inaccuracy + Risk of gaming)	**High** **Difficulty:** Although supply side data from the delivery agency is easy to collect, it may be prone to errors and manipulations. Also, supply side data may not correctly reflect the continued use of equipment or the use of multiple solutions (such as two solar lanterns in the same house)	**Medium-High** **Difficulty:** Attributes can be measured on the supply side (delivery agency data or product design specifications) and/or consumer side (user surveys). Supply side data is prone to inaccuracies and manipulations, demand side data requires expensive consumer surveys	**Low** **Difficulty:** Measurement of usability of energy usually requires extensive demand side data, particularly so in the absence of energy supply utilities. Energy surveys are difficult and expensive	**Medium-Low** **Difficulty:** Apart from electricity consumption data (which is measured by utilities), most other information–such as range of energy services availed–has to be obtained through surveys	**Medium-High** **Difficulty:** Although data on developmental goals is routinely available, further refinements to data may be necessary to relate such data to energy availability in terms of causal links
	Remedy: Design the output-based RBA to have a third-party verification mechanism. Identify and plug any possibility of gaming in program design itself. Supply side data may be supplemented with small user surveys to confirm user behavior	**Remedy:** (i) Invest in robust supply side measurement that minimizes errors or manipulation; (ii) Develop low cost survey mechanisms that use modern IT: (iii) Use a combination of short-term updates using supply side and longer term true-ups using periodic demand side surveys	**Remedy:** Use standardized survey approaches for accurate data gathering using consumer surveys. Use modern IT to reduce cost of data gathering. Use a combination of short-term updates using supply side and longer term true-ups using periodic demand side surveys	**Remedy:** Use standardized survey approaches for accurate data gathering using consumer surveys. Use modern IT to reduce cost of data gathering. Use a combination of short-term updates using supply side and longer term true-ups using periodic demand side surveys	**Remedy:** Use data available from omnibus surveys to measure indicators on developmental goals and correlate them with aspects of energy use to derive RBA payment metrics
Detectable Changes Over Time	**High** Result observable as soon as output is delivered	**Medium** Attributes pertaining to product performance are observable as soon as the product is delivered. Attributes pertaining to user behavior may require a longer period	**Medium** Tier 0 to Tier 1 movement is observable as soon as the adequate product is delivered. Movements to higher tiers may require a longer period. Hybrid approaches[7] should be explored	**Low** Even if usability of energy supply is high, actual use usually increases gradually over time, depending on external factors such as income, spending preferences, appliance availability, economic activity	**Low** Result will be observable over a longer period of time. Hybrid approaches should be explored

Attribution	**High-Medium** Often possible to directly link changes in output to investment by program (e.g., increased investment in renewables resulting in increased capacity)	**Medium** Changes in attributes can sometimes be linked directly to program investment, but can also depend on other factors	**Medium** Depends on all attributes affecting energy use; some of which are likely to change as a result of the program, but others of which will be determined by other developments	**Low** Determined by many factors, difficult to determine whether changes are due to the program or other developments	**Low** Determined by many factors, difficult to determine whether changes are due to the program or other developments
Appropriate Incentive Effect					
Ease of Explaining Result to Stakeholders	**High** Output is a tangible result that can be easily explained	**High** A single attribute is a one-dimensional transparent result that can be easily explained	**Medium** Need to explain the rationale behind each tier profile as well as the index calculation. This may be mitigated if multi-tier metrics are widely adopted	**Medium** Result may be easy or somewhat difficult to explain depending on the indicator chosen	**Medium** Result may be easy or somewhat difficult to explain depending on the indicator chosen
Agent Influence on Result	**High** The **energy provider** will usually be in control of the output delivery, while **local** or **national government bodies** may or may not be in control depending on the output	**High-Medium** The **energy provider** will be in control of attributes pertaining to specification. **Local government** may be in control of certain attributes such as availability of fuels. **National government** may be in control of most related attributes through utility	**Medium** Depending on the geographical area in question **local or regional government** bodies may be in control of the usability of energy supply. **National government** may be in control of the national utility and of the outcome of grid based projects	**Low** Very difficult for any agent to control actual use of energy services, as it depends on multiple factors such as income, spending preferences, economic activity, and traditional practices. Only **national government** bodies likely to be willing to enter RBA agreement	**Low** It is very difficult for any agent to control most impact-level results, as multiple external factors come into play. Only **national government** bodies may be willing to enter RBA agreement, if payment level is sufficiently attractive
Autonomy	**Low** Necessary to determine which outputs to incentivize. This will limit autonomy and is likely not to be technology neutral	**Medium** Necessary to determine which attributes to incentivize, but can be technology neutral	**High** Technology neutral and allows recipients autonomy over how to achieve improve usability of energy supply	**High** Technology neutral and allows recipients autonomy over how to achieve improve usability of energy supply	**High** Technology neutral and allows recipients autonomy over how to achieve improved usability of energy supply

Source: ESMAP, Vivid Economics

However, if monitoring and verification are costly and unreliable, the recipient may be concerned that the improvements that they make will not be sufficiently rewarded. They will be less willing to invest for any given financial incentive. In an extreme case, verification of results may become a point of conflict and acrimony between the recipient and funder. Indicators with clear line of sight should be preferred not only for the direct savings from cheaper measurement, but for the indirect benefit of transparency of disbursal and increased recipient confidence. For the same reason, the definitions discussed in Chapter 1 emphasize the importance of *independent* verification as a necessary feature of results-based approaches, even if this necessitates higher costs.

Second, even for very successful programs, at higher levels of the results chain there can be a significant time lag before measurable improvements are achieved. This is most obvious by comparing the output and impact level; while a program may expect to achieve increased investment in energy infrastructure within one or two years, benefits in terms of decreased emissions, better health, and so on may only become apparent after several years. For instance, the health benefits of widespread use of improved cookstoves may only begin to be identifiable in local health statistics after a few years. Timely disbursement is important because it provides a stronger incentive and facilitates further scale-up if a program is yielding improvements, but this may be impossible if payments are linked to changes in impact that are only observable after a lengthy lag.

Finally, attribution of changes to the program becomes more difficult at higher levels of the results chain. Changes in outputs can often be traced directly to investment conducted through a particular program, but many different factors will be interacting to produce changes at the impact level. As a result, setting the baseline against which progress is measured may be very difficult. It may be impossible to distinguish the program effects from wider changes that are happening within a country or region, and this could undermine confidence in a results-based framework for both funders and recipients.

Appropriate Incentive Effect

The result (and its associated indicator) will form a more effective incentive if it is easy to explain. One of the advantages of an RBA scheme is promoting the pursuit of a shared objective within the recipient government, by drawing attention to a specific and well-articulated goal. This effect will be stronger if the result is relatively easy to explain and reflects a concrete improvement in conditions within the recipient country. In addition, complex or composite results render it more difficult for the agents involved in the program to understand the consequences of their actions in terms of disbursements. In general, explanation will be easy at the output level, where the relevant results will largely be concerned with concrete changes in energy infrastructure. At the outcome or impact level, although there will be significant differences between results (and indicators), explanation may be harder. For instance, an output level result such as the capacity of the electricity network arguably sends a clearer message than an impact-level result tying to increases in the education index of the human development index (HDI), even if progress against the former may result in progress against the latter. As a result, the goal implicit in the former result will be easier to disseminate and agents will have a clearer idea of the effect of their actions on payments. Chapter 7 discusses an example of this issue within the energy sector, looking at

how the use of different approaches to measuring energy access may be better or worse at incentivizing improvements in energy access due to their ease of understanding.

In addition, for the result to be cost-effective, the recipient should have significant influence over the result that is being incentivized. If the result depends on a wide number of contingencies beyond the control of the recipient, the incentive effect will be weak. In the most extreme case, the recipient government may feel that the policies that they can put in place will have little or no effect on the result, and the introduction of an RBA will give them no incentive to affect change. For example, incentivizing units of energy saved (in watts or equivalent) in an efficiency program is likely to be more effective than incentivizing the monetary value of energy saved. Whereas the energy savings are likely to be under the control of the recipient, the monetary value of savings will depend on a wide number of macroeconomic and international factors beyond the recipient's control. On the one hand, they may be able to benefit from windfall profits due to rising energy prices, despite expending little effort on improving efficiency. On the other hand, they may receive little benefit from aid despite significant improvements in efficiency, due to unexpected decreases in price. The incentives they face will therefore be much weaker.

Broadly speaking, ease of influence will decline at higher levels of the results chain. For example, at the outcome level, it will be difficult for any agent to control actual use of energy services, as such a result depends on multiple external factors such as income, spending preferences, education, traditional practices, and overall economic activity. This is closely related to the above point regarding ease of attribution.

However, the extent of influence over the result will depend significantly on the recipient. As stressed at the beginning of this section, the choice of result cannot be conducted entirely independently from the choice of recipient. Different recipients will have influence over different results to varying extents, making some results suitable for use with particular recipients but not others. In particular, while implementers and service providers may have little ability to influence the wider factors that may impede or support the relationship between success in their project and the objective of the intervention (e.g., grid weaknesses may preclude renewable power capacity translating into renewable power consumed), government may have considerably more control over these factors.

The level of autonomy provided to recipients will also have an important bearing on the extent of incentive. A number of studies point to the importance of providing recipient autonomy as being crucial in the successful delivery of results (Honig, 2014), including by allowing exploitation of the greater knowledge that recipients have regarding barriers that may be holding back delivery of results. The level of autonomy is likely to increase at higher levels of the result chain: while there are only a small number of ways that a recipient can increase the number of solar lanterns distributed to rural households in a specific region, there are many more ways in which the same recipient can bring about a sustained increase in electricity consumption in rural households across a country.[9]

Combining the Three Factors

Each of the issues outlined above demonstrates some general considerations to take into account when choosing a result at a particular results level to target. Although the specific context

will have implications for which results level should be chosen, the advantages and disadvantages outlined in this section are likely to hold across most programs in the energy sector.

Outcome-level measurement will often be the most desirable, provided that resources are available for measurement. Outcome-level results will often balance the competing requirements for a desirable result. They maintain a decent proximity to impact, with increased usability of energy (taking into account the characteristics of energy supply, such as carbon intensity) generally implying increased use of energy and developmental impacts, at least when programs are well designed to meet local needs. The incentive effect is also well judged, with results being within the control of national governments while also providing significant autonomy. The only major disadvantage is that measurement can be costly, potentially requiring extensive end-user surveys.

If outcome-level measurement is not possible, it is important to be aware of the potential advantages and disadvantages of measurement at other levels in the results chain. Output and intermediate outcome levels are relatively easy to measure and provide a transparent incentive, but are far from impact and limit recipient autonomy. If applying a results-based scheme at this level, it is worth being particularly wary of the potential for the results to be achieved without development impacts; it is essential to think through the theory of change and assumptions implicit within the program to ensure that they are robust. On the other hand, intermediate impact and impact level changes can be difficult to attribute to the program and changes will often be apparent only after a significant delay, but they do ensure that incentives are well aligned with program objectives. Programs targeted at this level are likely to encounter problems with the design of an appropriate baseline and with ensuring that disbursements occur regularly enough to fund further progress against the objectives.

SELECTING AN INDICATOR

To make concrete trade-offs between different design options requires consideration of specific indicators available. The discussion above highlighted the trade-off that will typically exist when choosing between results at different levels of the energy results chain. However, ultimately, these trade-offs will need to be informed by the context surrounding a specific intervention and the indicators that are consequently available for different results: existing monitoring frameworks may or may not exist within a particular country, for instance, which may make targeting outcome results more or less viable in some contexts than in others. Similarly, the choice between an output and outcome result may be finely balanced, and depend on the specifics of particular indicators available for each type of result. There will also be some cases where there is more than one indicator available to measure a particular result; for instance, changes in the usability of energy can be tracked at different levels of disaggregation or with emphasis provided to improvements for particular groups, as will be discussed in more detail in Chapter 7.

To determine the best indicator to apply in a results-based approach, each of the available options can be rated according to the same three factors identified above. Any candidate indicator can be graded and compared according to the criteria. The best indicator to apply will combine

proximity to impact, appropriate incentive effect, and ease of monitoring. A traffic light analysis, as shown in Figure 4.2, which was prepared in relation to the possibility of using RBA to support a cookstove intervention in Ethiopia, may be helpful when choosing between multiple options. As far as possible, comparable empirical evidence should be used when determining the optimal result to apply. In the early stages of applying RBA within the energy sector, such evidence may be limited and, to a certain extent, practitioners will have to rely on judgment.

Any indicator that is seriously deficient in any of the three factors should be dismissed as an option. Strong performance in any two of the three determinants should not be expected to compensate for poor performance in the third. For instance, a result with very strong performance on incentive effect and ease of monitoring may still be undesirable for use within an RBA if it has a very weak link to the desired objective.

To determine the proximity to objective, it is useful to explicitly identify the assumptions linking the result level to the objective. In many cases the chain of assumptions linking an output result

FIGURE 4.2
Traffic Light Analysis Can Help Select Between Candidate Results

	Output Cookstove Distribution	*Outcome* Multi-tier Framework	*Output/Outcome* Hybrid Indicator
Proximity to Objective	Significant risk of improvements not leading to impact	Improved performance as measured by indicator likely to lead to impact	Maintains incentive to encourage use, but some disbursement still based on distribution
Appropriate Incentive Effect	Easy to disseminate, government can influence performance	Attributes intuitive though indicator can be difficult to explain, government can influence performance	More complex incentive scheme, may be hard to communicate priorities
Ease of Monitoring	Requires limited extension of planned MRV framework	Under current MRV arrangements would require additional annual household surveys, likely to be expensive	Requires semi-regular household surveys, but largely based on supply side data

Note: This diagram is taken from the case study analysis of a cookstove program in Ethiopia. The suitable indicator for use depends on the availability of budget for monitoring and verification. The multi-tier framework is most suitable in terms of proximity to impact, but also requires increased expenditure on monitoring.

Source: Vivid Economics

and the intended objective of the intervention will be relatively long, and certainly longer than the assumptions linking outcome and intended impact. These assumptions should be tested for plausibility to assess proximity to impact. If it seems plausible that one or more of these assumptions may fail to hold, RBA using a result from this level of the results chain is much less attractive.

IMPLICATIONS FOR MODALITY CHOICE

If it is not possible to find a result–and associated indicator–that performs well against all of these criteria for a particular recipient, an RBA approach is unlikely to be viable. For instance, incentivizing some aspects of good governance is likely to be difficult through RBA. Indicators in this area are highly contested and lack a clear line of sight. This makes it unlikely that recipients will be willing to accept the risk of receiving payment against these indicators and potentially makes monitoring the obtainment of these results too expensive. From this perspective, other sectors, such as health, education, energy, and transport, are likely to be better suited to RBA.

By contrast, if, for a particular government, a result and associated indicator is available that performs well on all three criteria, there could be a strong case for using RBA. Whether RBA is preferable to a conventional approach also depends upon other factors, such as recipient and funder capacities, discussed in Chapter 2. However, the availability of a strong indicator that is closely tied to the program objective, easily monitored, and provides an appropriate incentive should greatly increases the benefits of applying RBA.

TARGET CHOICE
Project and Recipient Characteristics Should Both be Considered When Determining Whether to Use an RBA

KEY FINDINGS

- For some situations, RBF, which provides financial support directly to service providers or implementers, may be more appropriate than RBA.
- RBA will be more appropriate when the government is in a much better position to influence performance than service providers. For instance, if the barriers preventing improvements are regulatory, RBA is likely to perform better than RBF. Often, this implies that service providers or implementers may better manage output-level results, while governments may better manage outcome-level results.
- RBA will be more appropriate when there are large returns to scale from managing the program at a national scale. RBF may be better if problems are localized or returns to scale are limited.
- RBA will be more appropriate when the government is relatively high capacity compared to other actors. RBF may be better if the private sector is higher capacity and has better access to finance.

This section considers the factors that might influence the appropriate recipient in a results-based agreement. The analysis in Section 4 considered whether a result–and associated indicator–might be appropriate for a given indicative recipient. However, there are also a number of factors which can also influence whether a particular recipient may or may not be appropriate, and hence, if a results-based approach is to be used, whether RBA or RBF may be preferable. The text in this section assumes that this analysis in Section 4 has been undertaken and discusses how an RBA designer may check whether government is indeed the best recipient, or whether alternative recipients should be considered. In practice, both the result and the target would have to be considered in combination: designers of a potential RBA need to consider whether there is an internally consistent set of results, indicators, and recipients that lead to a robustly designed RBA.

There are three main factors that will influence the appropriate target choice:

1 | The recipient should be in the best position to influence the chosen result

2 | Intervention characteristics may imply that interventions will be more or less cost-effective if pursued at the national level

3 | Recipients may have different capacities, incentives to deliver, or access to pre-finance

RECIPIENT INFLUENCE ON RESULT

For a given result and indicator, the most appropriate recipient will be the body that has the greatest influence over the result. Each potential recipient will have control over different sets of interventions. Private companies may be able to use existing networks to increase market penetration or marketing activities for solar lanterns, but unable to change regulatory or tax regimes that could encourage increased adoption. Governments may be able to improve the regulatory environment for large-scale renewables but could be unable to access the capital necessary to directly increase investment. Either party could lack the relevant expertise necessary to design and implement effective energy efficiency projects.

This requires consideration of the specific barriers that are relevant within the context of the intervention, as well as the capacity of different agents to deal with these barriers. Low energy efficiency in industry, for instance, could result from any number of different barriers. Firms could be unaware of or underestimate the returns from energy saving investments. Returns could be too low as a result of fossil fuel subsidies, limited carbon pricing, or high interest rates. Credit may be unavailable. Some recipients will be well placed to deal with some of these barriers but not others; choosing the right recipient depends on which barriers are most salient in the context in question.

Typically, output level results may be better influenced by service providers/implementers while outcome results may be better managed by government. Outputs in the energy sector, such as renewables generation capacity or provision of improved cookstoves, could, in many contexts, be expected to be heavily influenced by private investors, companies, or NGOs. These organizations that will typically have the greatest ability to influence delivery of generation, transmission, and distribution infrastructure. However, the corresponding outcomes, such as increased usability of low carbon energy or improved cooking solutions, will depend on a wide array of factors, many of which may be external to any service provider. For example, a company could reasonably be expected to identify and invest in renewables projects but ensuring that the energy produced is usable may depend on broader factors such as electricity distribution and regulation. Governments, on the other hand, are not well positioned to become directly involved in renewable power provision or capacity investment, but will have a wide array of potential tools at their disposal to influence the usability of energy, including the regulatory regime, pricing policy, and so on.

In a situation in which a particular result and indicator have been identified using the analysis described in Section 4, a series of questions can be used to cross-check whether the combination of recipient and result/indicator is robust. The analysis described in Section 4 already requires assessment of whether the relevant level of government has influence over the result. Given this, the following additional questions might be considered at this stage:

- In the event that the prior analysis suggests targeting an outcome, it should be checked that there are no other levels of government or public agencies that may be in a better position to influence the result. This might involve considering whether regional governments or rural energy agencies, rather than the national government, have more influence over the result.

- In the event that the analysis in Section 4 suggests that the only viable results and indicators are output-level results, it may be considered whether one or a series of RBF schemes may be preferable to an RBA scheme.

INTERVENTION CHARACTERISTICS

Some interventions can be more effectively managed through a national program. A classic example in this category is vaccinations, where widespread immunity has the potential to impart benefits far beyond local immunization. National programs have the capacity to create 'herd immunity,' greatly decreasing or even eliminating the incidence of a disease. Local projects will never be able to realize these benefits, and this should be taken into account when determining scale. Renewable power production in some countries may fall into this category: there are likely to be some regions in a country that have no or limited resource availability. It may be necessary to exploit more abundant opportunities in other regions and export this power to the regions with fewer natural resources. However, in other cases, national and even international benefits can be created through the results generated by a single service provider: for example, renewable resource mapping exercises undertaken by one provider can yield benefits across a country.

Certain interventions may be more cost effective at the national rather than local scale. This can be the case due to increasing returns to scale. Returns to scale will be present if there are large fixed costs, which will be paid regardless of the size of the intervention. In particular, monitoring and verification costs may greatly decrease with wider scope. If a separate system of monitoring will need to be developed for each project, the costs will often be prohibitive. Setting the baseline for results-based approaches, discussed in Chapter 6 under "Baseline," may be expensive at the project level but affordable for the country as a whole. There will be similar duplication of costs if a large number of separate contracts with individual service providers need to be negotiated to deliver the desired results. On the other hand, if there are no substantial economies of scale then the structures that would need to be created to organize an RBA within a country could lead to unnecessary bureaucracy.

A wider geographic scope can help manage the risk associated with results. For example, a renewables project limited to a particular technology or location may be more or less successful in any given year due to favorable or inclement weather. An individual service provider may have very little ability to mitigate this risk. By contrast, national governments managing a nationwide renewables program and applying multiple technologies may be much less exposed to weather risks. This is due to both natural diversification in weather across the country, and because the array of legal, regulatory, and other (soft) powers that national governments have access to can be used to influence the location of plants. In these cases, moving to a governmental recipient creates significant improvements in cost effectiveness. However, in cases where there is limited diversification gained by scale, then adopting an RBA scheme would, as above, create unnecessary bureaucracy.

RECIPIENT CHARACTERISTICS

The agents at the regional and national levels may be in very different positions with regard to their capacities. Competencies will vary between local, regional, and national governments, NGOs, private companies and other potential partners to a results-based agreement. This will affect the ability of different recipients to influence the result.

National governments may have weak control or oversight over the implementers that they need to work through to deliver results. In highly federal or decentralized systems, higher levels of government may actually have significantly less control over results than lower levels of government or implementers working at the local level. The extent to which this is the case will vary from country to country, and can be mitigated by budgetary control systems that replicate the incentive facing the national government at the local level.

Some actors may have far superior access to capital. On the one hand, national governments have access to tax revenues and will in most situations, be able to raise at least some quantity of pre-finance. On the other hand, access to international capital markets among developing countries is often limited. For example, large companies operating within a developing country may well have superior access to international capital, especially in the context of a partnership with a reliable funder. Since access to pre-finance will affect the overall cost-effectiveness of the project or program, these factors should be taken into account when choosing scale.

Some actors may have stronger internal incentives to deliver results. A results-based program will only be effective if the targeted recipients have an incentive to deliver. For private sector actors, this incentive may arise due to the results-based financing scheme itself and the opportunity for profit that it creates. As discussed in Chapter 2, it should not be assumed that a RBA program aimed at government will incentivize actors in the same way; individuals within government will already be embedded in a complex system of incentives that responds to many factors other than opportunities to increase revenue. If it seems unlikely that internal incentives within government will support the targeted result, it may be better to consider non-state actors.

Designers of results-based approaches should ensure that the initial recipient continues to be the most appropriate. Following the logic of this report, an initial assessment of the government in terms of institutional capacity and access to pre-finance will have been undertaken prior to proceeding with any RBA design work. The analysis of this section suggests that these issues should be revisited in light of an initial choice over result and indicator to check that there is not an alternative recipient who could deliver the chosen result more effectively

CONCLUSION

Three main factors need to be considered when deciding between RBF and RBA. The appropriate target level depends on recipient influence over the result, the general cost characteristics of the program, and the characteristics of the recipients.

The energy results chain illustrates the general principle that higher level results may be better targeted through governments. Governments are likely to be able to influence a wider array of factors than individual energy providers. This means they will often be better placed to influence higher level results, such as outcomes or intermediate impacts, which are often beyond the control of implementers. Conversely, they will typically be less effective at influencing output-level results.

The underlying characteristics of the intervention should be taken into account. For particular interventions, a wider geographic scope can help manage risk and achieve program goals, or help to achieve return to scale through reduced contracting or monitoring costs. However, where these benefits do not exist, there is a risk that RBA agreements will generate unnecessary bureaucratic structures that do not support result delivery.

Finally, the differing capacities of potential recipients should also be considered. This includes both their ability to access pre-finance and their institutional capacity.

INSTRUMENT DESIGN
RBAs Can Vary by Payment Type, Payment Structure, Baseline, and Strategies Adopted to Avoid Time Inconsistency

KEY FINDINGS

- Payment type for RBA can be structured as either grants or concessional loans, although this choice is often dictated by the nature of the funding organization. Concessional loans allow refinancing of debt at lower interest rates, but are likely to offer a less transparent incentive than grants.
- Payment levels should be set by negotiation between government recipients and funders. Both are equal parties to the contract, and have opposing interests. Nonetheless, the payment should not be set above the value that the funder places on results, nor below the level that recipient requires to cover the marginal cost of the intervention (broadly understood), taking into account additional sources of income and the financial resources that the recipient can provide.
- When targeting multiple indicators, relative payment amounts should be set with reference to the relative value of achieving improvements in each. In the energy sector, this can often be approximated using consumer surplus.
- The baseline should isolate the effects of the program. This typically requires studying the historical trend. RBA is easiest to apply in areas where progress against the indicator is relatively static, and progress is likely to result from the additional support provided through the program.

Once it has been determined that an RBA would be the best instrument to apply, there are a number of design questions that need to be addressed. The RBA agreement should be designed by negotiation between the recipient and the funder. There are at least four elements to the agreement that could have a significant effect on the effectiveness of RBA:

- **Payment type.** Should aid be provided in the form of a grant or a loan?
- **Payment structure.** At what level should the payment be set? How often should it be updated?
- **Baseline.** Against what baseline should progress be measured? How often should the baseline be updated?
- **Commitment.** How should the agreement be designed to assure both parties that it will be honored?

PAYMENT TYPE

There are four fundamental types of funding: grants, loans, guarantees (or other risk instruments), and equity. Each could, in theory, be tied to results. There is some reflection of this potential in existing results-based schemes; Cash on Delivery Aid, for instance, has focused on grants, whereas the World Bank's Program-for-Results (P4R) has focused on loans.

Guarantees and equity purchases make little sense in the context of an RBA. Both equity purchases and guarantees are typically required at the early stages of a project or program. Since a results-based approach requires that the reward be distributed following the successful implementation of the project, neither equity purchases nor guarantees would be suitable.

Concessionary loans, on the other hand, could be used within an RBA. Concessionary loans implicitly combine a grant with some degree of capital recovery for the funder and could provide a significant incentive to the recipient by offering the prospect of refinancing an initial high cost loan at lower interest rate. Loans could be more appropriate than grants, particularly in circumstances where the activity being financed will generate additional economic activity or revenues that can be used to repay the loan. This would allow the loan capital to be redeployed and generate additional development gains.

The choice of financial instrument will often be determined by donor type. Many agencies considering the use of RBA will have already made decisions about which financial instruments they can provide. For example, many bilateral organizations only directly provide grant resources, though they often also contribute to organizations that provide loans, such as the International Development Association. By contrast, most support[10] from multilateral development banks is only provided in the form of loans with varying degree of concessionality, depending upon the country's income.

The use of loans may be particularly relevant in the energy sector given its capital intensity, especially in relation to on-grid renewables, increasing energy efficiency, and improving and expanding grid infrastructure. Some, but not all, of the effective interventions in these areas may require access to capital. These investments have the capacity to generate relatively predictable additional revenues but are highly capital intensive. The sums involved may be too great to make grants sustainable, whereas an RBA using concessional lending may be more attractive.

PAYMENT STRUCTURE

In theory, a wide variety of different payment structures could be used within an RBA. An RBA could offer a predetermined incentive per-unit of result, a lump-sum payment once a particular threshold has been reached, or increasing or decreasing payments as the level of achievement varies.

However, to maintain transparency and provide clear incentives, a predetermined per-unit payment will often be the most appropriate. Other structures may prove problematic. For instance, a threshold design will create perverse incentives; once the threshold has been achieved there is no

reason for the recipient to continue to invest in the outcome. Any change in situation which makes reaching the threshold appear unviable to the recipient will also have the effect of completely removing financial support. Increasing or decreasing the per-unit payments based on the quantity of the result achieved may be possible but would also increase the complexity of design, reduce transparency, and complicate the negotiation process. This is a separate issue from reducing the quantity of the payment gradually as a means of phasing out the program, which could be achieved over the longer term as agreements are renegotiated.

Payment per result should be determined in negotiation between the recipient and the funder. An RBA program is an agreement between two parties. Both stand to gain from the arrangement; the result represents a shared goal that both parties wish to achieve, but are unable to achieve independently. As a result, there is no 'right' price for a result, even theoretically, and the payment level needs to be set through negotiation. This contrasts with an RBF scheme where concerns over the potential distortion to competition imply that the minimum amount of concessional funding should be used that is consistent with realizing the result. In addition, with a variety of potential implementing partners, 'price finding' tools such as reverse auctions can be used to determine payment levels within RBF, whereas they are not applicable within RBA. The sections below explore the factors that should influence the RBA payment negotiation in more detail, looking first at the (relatively) simple case in which a payment only needs to be set for one results indicator, before considering how to set payment levels when an RBA agreement includes multiple results.[11]

Boundaries to Payment

The highest amount that the funder should be willing to pay is the value that they place on the result. The value they place on the result will, in turn, depend on the development benefits that the result implies and the broader objectives of the funder. They will be willing to pay up to this amount. Any higher, and the expense of the program will fail to justify the benefits. Any lower, and they will be achieving a net gain through the program; the development benefits will outweigh the fiscal costs.

The lowest amount that the recipient should be willing to accept is the marginal cost (MC) of achieving the result, minus the value that they place on achieving the result and the other sources of funding that are available. The marginal cost to the recipient is not necessarily the technical cost of the intervention. They may be in a position to, for instance, achieve the result by changing the regulatory regime or changing the priorities of government staff. In addition, given that they value the result, they may also be willing to dedicate some of their own resources to the program, which may enhance the attractiveness of the intervention for the funder as their resources would be 'leveraged'. This will decrease the total necessary compensation to below the full marginal cost. Finally, the RBA scheme may not be the only source of revenue. Some of the cost may be recoverable from consumers or some may be provided by other funders. As such, they generally are willing to accept below the marginal cost of the most efficient intervention. The minimum willingness to accept of the recipient is therefore given by:

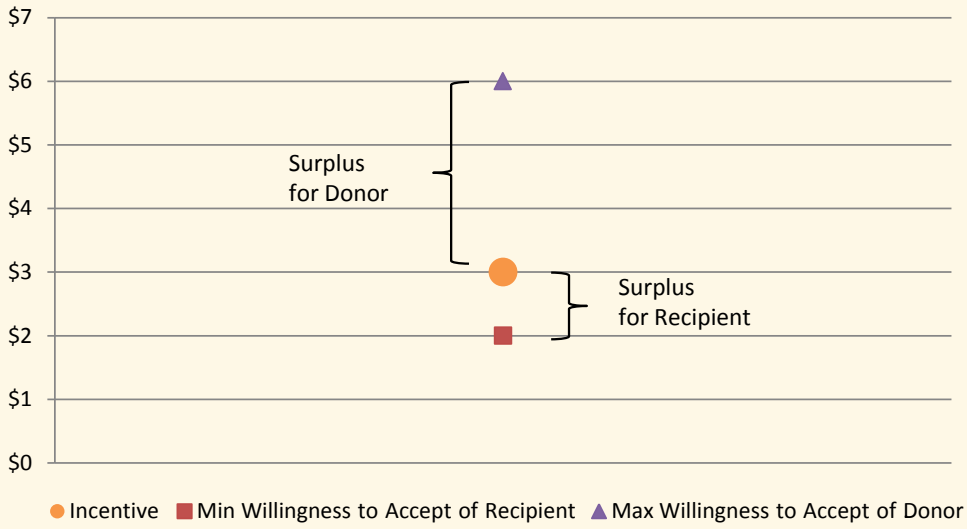
minimum willingness to accept of recipient = MC of intervention – other sources of funding

– value of result to recipient

Any payment level between these two values represents a different split of surplus between the two parties to the agreement. The difference between the funder's maximum willingness to pay and the recipient's minimum willingness to accept is the total surplus created by the program per result. The position of the payment level between these two values divides up this surplus. In the extreme, if payment is set at the funder's maximum willingness to pay, then the recipient receives the entire surplus and the funder has no net benefit (though also no loss) from participating in the program. If the payment is set at the recipient's minimum willingness to pay then the opposite is true; the funder receives the entire surplus and the recipient has no net benefit from the program. Figure 6.1 provides a graphical representation. In this example, the funder is willing to pay $6 per result achieved, corresponding to the value that they assign to the result. The recipient's minimum willingness to accept is $2. This could be lower than the full marginal cost of the intervention. For instance, if the marginal cost is $9 per result, but the recipient can also access other funding worth $4 per result and values achievement of the result at $3, then their minimum willingness to pay would be $9 – $4 – $3 = $2, as shown in the diagram. Any payment level between $6 and $2 is, in principle, acceptable and each represents a different division of the $4 surplus per result between the recipient and the funder.

FIGURE 6.1

The Position of the Incentive Level Determines the Division of Surplus Between the Two Parties



Source: Vivid Economics

Other considerations beyond the value and costs of the intervention may be relevant due to institutional constraints facing the funder or the recipient. Aid budgets are rarely set exclusively with reference to the available interventions within a particular country or the value of the results that these interventions could generate. Instead, high level allocation decisions, for political or other reasons, can shape the total quantity of aid available for programs in a particular country. Disbursement requirements can then affect whether individual programs are funded. Macroeconomic conditions and absorption constraints can also limit the quantity of aid that can prudently be allocated to a particular country. The factors determining aid allocation vary across funding institutions and recipient governments, and cannot be treated at depth within this report. Nonetheless, the factors discussed above, such as the cost of the intervention and value of the results generated, should be the primary considerations when allocating a set budget between different competing programs.

Paying Near Minimum Willingness to Accept

The funder will generally want the incentive to be as close to the recipient's minimum willingness to accept as possible. If they paid any less, the recipient would be unwilling or unable to achieve results and there would be no development impact. On the other hand, paying higher achieves the same development impact but costs the funder more. Paying exactly the minimum willingness to accept is sufficient to incentivize the recipient to achieve the result, but (from the funder's perspective) requires minimum possible resources to do so. This will allow them to realize the highest possible surplus per result. This surplus can then be reinvested in achieving additional development goals in other programs.

There may be a case for paying above the minimum to mitigate the risk that the result is not achieved. A number of developments may increase the recipient's minimum willingness to accept:

- Fluctuations in raw material or labor costs may increase the marginal costs of the most efficient intervention
- Other aid flows may be retracted, decreasing the resources available to the recipient to achieve the result
- Political changes may decrease the value that the recipient places on the result

If, as a result, the minimum willingness to accept exceeds the incentive payment, the recipient may be unwilling or unable to pursue the result at all. Setting the incentive higher than the minimum mitigates this risk at the cost of additional resources.

Payment levels that are tied to raw materials or labor costs may help mitigate this problem, but may be hard to apply in many contexts. If there is an independently verifiable source for the cost of raw materials, the incentive payment can be tied explicitly to their market price. This may help ensure that the incentive payment moved with changes in willingness to accept. However, as explained above, it is not necessarily the case that the relevant marginal cost will correspond to the technical cost. Instead, the marginal costs may represent the costs of compensating for necessary

regulatory changes or instituting a performance-based pay scheme which aligns the incentives of workers with the objectives of the program. Neither need be affected by changes in the cost of raw materials. Even when this is the case, raw material costs may vary significantly across a country and may move independently in different regions, particularly if they are locally produced and transport infrastructure is poor. Finding a reliable adjustment method will be difficult in these situations.

Paying Near Maximum Willingness to Pay

Symmetrically to the funder, the recipient will want to receive as close to the maximum willingness to pay of the funder as possible. This allows them to maximize surplus gains per unit of the result delivered. This surplus can then be reinvested in achieving domestic or other development objectives.

Paying a higher amount may draw more attention to the program, improving performance. As emphasized in Chapter 4, part of the benefit of RBA is likely to derive from providing a well-defined target and focal point for the attention of diverse actors. It is hard to capture this effect within the formulation of minimum willingness to accept provided above, which implicitly assumes a single well-defined 'value' placed on the result by the recipient. This is a useful framework for thinking about payment levels, but it should be emphasized that RBA is intended to affect the behavior of many agents operating at different levels within government and civil society. Attaching a greater payment amount to the result may generate more focus and attention on the target, leading to more effort being exerted to achieve the goals of the program and, as a result, better performance.

Uncertainty in Estimating Maximum Willingness to Pay and Minimum Willingness to Accept

All of the determinants of willingness to pay and willingness to accept will be highly uncertain. The value of achieving a result, for both the funder and the recipient, is not clearly defined. Either may discover that they are willing to pay more than they expected over the course of negotiations or implementation. For instance, new evidence may become apparent that demonstrates greater benefits to the intervention than expected. Marginal cost is also highly uncertain and subject to change; as discussed above, it is not necessarily the technical cost, given that governments will often have a wide array of levers that they can use to support a result. It also will often not be known in advance, by either the funder or recipient, which of a number of potential interventions will be the most cost effective. Finally, other sources of revenue can be uncertain. Other aid flows, for instance, can be quickly and unexpectedly withdrawn and, as a result, it is very hard to know how much aid money will be supporting a program one or two years in the future.

The result level being measured will determine the information that is available to negotiators. For lower level results, such as outputs, there is likely to be better information on the technical costs of achieving the incentivized improvements. For instance, there may be reasonably good information on how much it would cost to build additional solar capacity or to purchase and distribute solar home systems. On the other hand, it will be much less clear what the value of these output measures are

unless there is strong evidence on how they affect development outcomes. As such, it may be relatively easier for an external party to estimate recipient minimum willingness to accept and harder to estimate funder maximum willingness to pay. The opposite will be true for higher level results; there may be better information available on how much the results are valued but, because of the diversity of interventions that could be used to achieve them, much more uncertainty over how much they would cost to achieve.

Negotiations

A starting point for negotiations could take into account current aid flows and a benchmark for the marginal cost of the more cost-effective interventions. First, approximate answers could be found to each of the following questions:

- How much do current aid programs pay to achieve this result per household?
- What are the marginal and average costs of the more cost-effective interventions likely to be?
- What are the current foreign aid disbursements to the country, both as a whole and to the energy sector?
- How has the country performed historically against the chosen indicator, and how is it performing currently?
- Taking into account past experience and domestic capacity, what range of achievement against the indicator is possible?

Using this information, three scenarios could be constructed to correspond to high, expected, and low achievement against the result, given the resources available. Total payment amounts could then be calculated under each scenario and compared to other flows of finance and the funding necessary to cover the costs of planned interventions.

A reasonable payment level should meet at least three criteria:

1 | The total RBA flows under realistic achievement scenarios should be comparable to other relevant aid flows and domestic investments, to ensure that the RBA will have traction with policymakers while also being consistent with the absorptive capacity of the recipient.

2 | RBA funding should be sufficient, in combination with recipient resources and other sources of funding, to cover the marginal cost of the more cost-effective interventions to achieve the target.

3 | The payment per result should be low enough to provide reasonable value for money for the funder.

The modeling results, combined with the information described above, should be sufficient to gain an approximate idea of what payment levels are likely to fit these criteria, and form a reasonable starting point for negotiations.

It may be necessary to set a total payment cap to prevent large budget overruns. While overachievement would be welcome, the funder may need to control the outlays that could result from success far exceeding expectations. This could take the form of an annual expenditure cap or an expenditure cap for the entire program, and should also be discussed as part of the negotiations.

BOX 6.1

Results-Based Aid Applied with an Open Contract

The analysis above assumes that an RBA agreement is negotiated between a predetermined pairing of funder and recipient, but other arrangements are possible. Earlier work on RBA (Birdsall & Savedoff, 2011) also identifies the possibility of an 'open contract' offered by either one or multiple funders to a set of qualifying recipients.

Funders could propose a standard contract, including a fixed payment per result, and allow any government from a set of eligible recipients to participate. For instance, a contract could be drawn up that includes a flat offer of $100 each year per household with improved access to cooking solutions. This contract could then be offered to a number of countries, each of which would be free to take or leave the offer.

Birdsall and Savedoff (2011) identify three major advantages of this approach. First, an open contract could reduce administrative costs; one off negotiations between the potential recipients and the funders determine the structure of the contract and payment level, and no further negotiations are then necessary. By contrast, negotiating separate contracts between each recipient and funder could be expensive and time-consuming. Second, an open contract would be transparent and uniform, increasing accountability and efficiency by ensuring that the same level of resources are dedicated to achieving a result across a number of countries. Finally, an open contract would encourage self-selection of countries for whom the terms are most attractive, either because the value of the returns are higher than average or because the results could be achieved at a lower than average cost.

It would, however, still require a baseline to be identified for each eligible country; this may limit the reduction in administrative costs. Payments for results within an RBA are disbursed only for improvements above a baseline (see Chapter 6 under "Baseline"). The baseline is intended to capture what would happen in the absence of the program and, if it does so perfectly, it ensures that payments are only made for results that are truly additional. In many contexts, setting an accurate baseline is difficult. An RBA aiming to decrease the carbon intensity of power supply, for instance, would have to approximate how much renewable and fossil fuel capacity would be constructed in the absence of an RBA, through the private and public sectors, in each of the eligible countries and over the full course of the program. If there are significant differences between countries, baselines would have to be different for the open contract to function well. Setting a low baseline (e.g., rewarding for all improvements over the levels in the first year of the program) would encourage the governments that were likely to make good progress in the absence of the program to apply even if the RBA would make little difference, whereas a high baseline (e.g., assuming that in the absence of the program results would improve at a multiple of their average rate over the last five years) could discourage applications from countries whose performance was deteriorating but would greatly benefit from an RBA. Where countries differ substantially, negotiations on the baseline would still, presumably, have to be conducted bilaterally and would continue to require significant resources.

Given uncertainty about the baseline, self-selection may also bring some disadvantages. As highlighted above, setting an accurate baseline might be difficult. At best, the final baseline represents a broad approximation based on the limited available information. If national governments are privy to information that is unavailable to those negotiating over the baseline, self-selection into an open contract may have perverse effects. For instance, governments may be aware of upcoming sources of funding that will allow them to decrease the carbon intensity of their electricity generation. Selecting into the RBA may then allow them to generate additional revenue from the RBA scheme, even if the program itself does not encourage improved performance.

Nonetheless, in situations where the baseline is transparent across a number of countries, an open contract could be a better long-term solution than bilateral negotiations. Given the complications with baseline, open contracts are most suitable where a reasonable number of countries can be identified with historically stable performance against the result. For instance, it may be suitable when a number of countries could be identified who have historically made very little progress on provision of improved cooking solutions. It is likely to be more difficult when dealing with a variable that is likely to behave unpredictably in different countries, such as the carbon intensity of electricity generation.

Multiple Results Indicators

Some schemes will incentivize multiple results simultaneously. For instance, it may be sensible to offer different payment amounts for different levels of access to electricity or cooking solutions. Likewise, stakeholders may wish to set different payments for results achieved in rural areas versus outcomes achieved in urban areas or it may be desirable to reward different renewable energy technology types at different levels due to their differing co-benefits. Any scheme like this will require setting multiple payment amounts and will create a menu of options for the recipient concerning which result to target.

When multiple results are incentivized simultaneously, the recipient will choose to invest in the result which maximizes their surplus. Interventions aimed at promoting different results will generate different levels of surplus. The recipient will target the result which maximizes:

$$\text{recipient surplus per \$ spent} = \frac{\text{(value of result to recipient)}}{\text{(MC of intervention} - \text{incentive} - \text{other sources of funding)}}$$

In the context of a fixed budget, targeting this result maximizes the total surplus that the recipient receives. This is the inevitable consequence of offering multiple results to the recipient; upon consideration of their options, they will focus on pursuing the result that offers them best value for money.

The value of different results within the energy sector depends on the benefit offered to the consumer and society. The value of an intervention may be approximated by the consumer willingness to pay for the benefits that they receive. This could be estimated either directly, through consumer surveys, or by reconstructing a demand curve. The latter approach uses data, such as market prices, quantities, and other observed factors likely to change the amount that a consumer is willing to buy or a firm is willing to supply, to statistically estimate the relationship between the price of a good and its demand.

Adjustments may be necessary to take into account additional external benefits that are not captured by consumers or benefits of which the consumer is unaware. A classic example is the external benefit of greenhouse gas emission reductions. Consumer surplus may be misleading due to the extent that these external benefits are not included. In addition, in some situations consumers may not be fully aware of some of the benefits that they will receive from the intervention; for instance, consumers may not fully appreciate the health benefits of improved cookstoves. Where available, consumer surplus estimates could be supplemented by evidence from rigorous studies that attempt to measure the value of these benefits. The value of emissions reductions from energy efficiency improvements, for instance, can be estimated by combining existing evidence on decreased fuel use with more efficient stoves with an estimate for the social cost of carbon.

Payments for different results should be set in line with the amount of the value they generate. For instance, if providing a grid connection is expected to generate double the value of basic solar electrification, then the payment for a grid connection should be around double the payment for basic solar. This ensures that the incentives facing the recipient are aligned with the developmental

benefits of results. A recipient, acting to maximize their own surplus, will then target a result if, and only if, it generates the most individual and social benefit per dollar spent.

This provides an objective basis for determining relative payment amounts. Both the recipient and the funder can refer to the existing evidence base on development benefits of interventions when negotiating the relative payment amounts for different results.

A similar process of iterative modeling of different payment amounts and scenarios can, again, form a useful starting point for negotiation. As when setting a single payment amount, information should be collected on expected marginal costs, other sources of funding, foreign aid disbursements, and the range of potential performance against each of the results. A value for the different results can be approximated in consultation with the recipient by looking at how much they currently spend on programs with similar objectives and the evidence base on consumer surplus. Putting this information together, a simple model could be constructed which has two main functions:

1 | As before, to calculate total disbursements under different achievement scenarios; and

2 | Based on broad approximation of marginal cost, other funding, and recipient value, to determine which results are likely to be targeted by recipients.

The payment schedule chosen should result in disbursements that are of reasonable value compared to existing aid flows. In combination with other available resources, they should be sufficient to cover the cost of at least the more efficient interventions that could be used to achieve results. The relative payments should be expected to provoke a cost-efficient response from the recipient, with their resources used to promote results that offer significant development benefits per dollar spent. Finally, as before, the scheme should be expected to also provide reasonable value for money for the funder.

BASELINE

In any RBA, it is necessary to determine a baseline against which progress can be measured. Payments are made for results over and above the baseline, but not for results that fail to exceed the baseline. To take an energy example, payments seeking to increase annual renewable generation could be linked to:

- total renewable generation (in MWh);
- the extent to which renewable generation exceeds 10 percent of total electricity generated; and
- increases in renewable generation over the previous year.

The first measure sets the baseline at zero. The second uses a baseline of 10 percent, and only rewards improvements above this level. The third sets a moving baseline, which updates each year. All RBAs implicitly make a decision about baseline, even if it is not explicitly considered during negotiation.

Ideally, payments should only be made when improvements in results would not have been achieved in the absence of the RBA. Rewarding improvements that would have occurred in the

absence of the program is in effect a pure transfer of income to the recipient country, and will not create a change in behavior. As a result, setting the baseline too low restricts the amount of funds available for other development uses without creating any gain in development goals. Setting the baseline too high, on the other hand, may greatly reduce the incentive effect of the RBA.

Constructing this counterfactual can be problematic. Attempting to predict what would have happened in the absence of a policy is invariably a substantial challenge. While historical trends are instructive, projecting forward requires a thorough knowledge of the key drivers of results and an accurate understanding of likely changes in these drivers in the short to medium term.

Guidance can be drawn from the literature on programs aiming to reduce emissions from deforestation and forest degradation (REDD). Establishing an accurate baseline is a long standing problem with incentive programs to reduce deforestation. REDD is intended to reward countries for deforestation rates underneath an established baseline. As with RBA, the intention is to reward all additional decreases in deforestation. In theory, this should also include slower than expected increases in deforestation.

Chagas et al. (2013) identify three main methodologies for establishing a baseline within REDD programs:

- **Historical.** The baseline is constructed using some average of historical rates of achievement within the sector.
- **Business-as-usual (BAU) projected.** The baseline without the RBA is projected using statistical techniques to take into account the key drivers of the result, where the drivers may include the impact of government policy and targets.
- **Historical adjusted.** The baseline is constructed from historical rates, with adjustment upward or downward based on evidence that the historical rate is likely to be inaccurate.

The 'historical adjusted' methodology is essentially a middle ground between 'historical' and 'BAU projected'; integrating some of the key drivers of results without undertaking a full projection exercise.

They suggest that using a historical approach is transparent and low cost, but may systematically over- or under-estimate the effect of the RBA. A historical baseline may be, for instance, an average over the last three years. Before applying a historical baseline, it is necessary to consider whether there is good reason to expect the future to be different from the past. If the sector within which the RBA is being applied is undergoing significant changes, the historical baseline may be inappropriate. For instance, in the context of rapidly improving results, applying a historical baseline is likely to transfer windfall profits to the recipient. An RBA to increase renewable energy use in Cambodia would need to take account of the fact that renewable energy has been growing at a compound annual growth rate of more than 15 percent in the period between 1990 to 2010 (IEA, 2013). At the same time, applying a simple historical average has the advantage of transparency, which may strengthen the incentive effect.

BAU projections can integrate future trends, but are likely to be more contentious and data intensive than historical rates. There will often be insufficient data available to conduct a full

projection. With no widely established methodology, BAU projections are likely to include a number of methodological decisions that will affect aid disbursals, which may be problematic during negotiations. Transparency is also limited, and costs are likely to be high.

Historical adjusted baselines provide a middle ground between the other two approaches, but need to be carefully justified. For example, a three year historical average could be adjusted using a growth rate to take into account recent and expected improvements. This approach is flexible and reasonably transparent, but, without investing in a sound statistical understanding of key drivers, determining the adjustment is likely to be contentious.

The regularity with which baselines are updated will affect dynamic incentives. For recipients, more regular updating reduces the risk that a country may be left with an unachievable baseline, and that the incentive effect will disappear. Setting an absolute level without updating creates the danger that, for some unforeseen reason, a dramatic decrease in the measured result occurs early in the course of the RBA, causing the recipient to face no reward for improvements until the high baseline is attained. There is a similar issue for funders in the opposite direction: infrequent updating could render an RBA program very expensive if the recipient proves able to more easily generate results than anticipated. On the other hand, regular updating undermines some of the dynamic incentive to make significant progress. If the recipient knows that improving results now will imply a tougher baseline in every subsequent year, they may have less of a reason to push for quick and dramatic improvements.

Updating reduces risk exposure for the recipient, but also reduces the incentive effect. When designing the RBA, the recipient and funder will need to consider this trade-off, and determine how regularly to update the baseline based on how much risk and incentive they want to shift to the recipient.

AVOIDING TIME INCONSISTENCY

The effect of an RBA program will be weakened if the aid recipient becomes convinced that the funder may deviate from his side of the agreement. This risk, called *time inconsistency* (promising to do A at time X, but really doing B when time X comes) is highlighted in two different forms by Birdsall and Savedoff (2013) and Öhler et al. (2012).

Focusing on the risk of non-disbursal, Öhler, Nunnenkamp, and Dreher (2012) found that the incentive effect from the Millennium Challenge Corporation (MCC) was noticeably weakened by vacillation and uncertainty about budget allocation on the funder side. The MCC has been found to have been effective in promoting better control of corruption. However, it was unclear at various times whether the US Congress would approve sufficient funds for the program, which undermined the strength of the incentive.

The risk of non-disbursal can be partly mitigated by the use of independent verification. The definition of RBA discussed in Chapter 1 emphasizes the role of independent third-party verifiers as one of three key elements that define RBA approaches. The use of independent verification can help build trust within recipients that delivery of the results will generate payments, and provide a clear

audit trail in cases where there are disputes. In order to ensure that these benefits are realized, both funder and recipient should jointly agree on which parties should act as third-party verifiers.

Birdsall and Savedoff (2011) focus on the risk of disbursal *in cases where disbursal is not warranted according to the RBA agreement.* Where a recipient government performs badly, a funder may be motivated by good intentions (Samaritan's Dilemma, Gibson et al., 2005) or by non-transferable budgets (leading to very low opportunity costs for disbursal, Svensson, 2003) to disburse the projected level of funds regardless of the level of performance, either via a one-off waiver or via changing the definition of results. However, doing so when limited achievement has resulted from poor government performance would materially weaken the incentive effect of RBA (Klingebiel, 2012a, p.18), both for the current RBA agreement and any future arrangements entered into by the same funder, thereby defeating the purpose for choosing this aid modality in the first place.

When underperformance has resulted from an exogenous factor, such as natural disaster, there is a stronger case for renegotiation. Exogenous factors are developments that are beyond the control of the recipient. Since exogenous developments are often unpredictable, disbursing in these cases does not undermine the incentive facing governments. However, distinguishing between exogenous and endogenous factors can be extremely challenging. There is no clear-cut method for determining what proportion of performance was driven by exogenous versus endogenous factors—a natural disaster is clearly beyond the control of government, but the impact a natural disaster has on access to energy depends on the quality and resilience of energy infrastructure, which is within governmental control. Equally, a particular natural disaster may be difficult to predict, but there is generally good knowledge about which areas are more or less exposed to natural disasters. Arguing for disbursal following a disaster is effectively arguing for increasing untied budget support for the recipient. This may well be important but whether it is justified or not largely depends on humanitarian or other concerns that are independent of the RBA program and may best be dealt with separately. As with any contract, renegotiation will be possible if it is in the interests of both parties, but it is important to bear in mind that persistent renegotiation undermines the transfer of accountability and ownership to the recipient and should be avoided unless necessary. A parallel problem is created by lack of coordination among funders, as discussed in Box 6.2.

Birdsall and Savedoff give four specific design recommendations for avoiding time-inconsistency. If possible, the first-best solution is to directly tackle the pressures that push funders to override an agreement, for example, the requirement to spend money in the year it was budgeted. Where this is impossible, the following three alternatives are available:

- **Operating with realistic expectations.** In particular, Birdsall and Savedoff recommend making a 10 to 20 percent downward disbursement adjustment for optimism bias.
- **Diversifying the RBA portfolio.** A funder can reduce the variability in their overall disbursement by including programs with different start dates, in different geographies, and in different sectors.

BOX 6.2

Risk transfer may only be possible when funders coordinate.

Achieving a risk transfer from funders to the recipient requires a coordinated approach between funders. For example, a country like Uganda, which has a combination of significant hydro-electricity resources but also other renewable potential, could be incentivized through an RBA to increase the usability of renewable power. To deal with uncertainty over rainfall (which is likely to be exacerbated by climate change) a suitably designed RBA agreement may provide an incentive over the medium term to develop a balanced portfolio of renewable power resources. However, in a specific year in which rainfall is particularly low, and RBA payments fall below that expected, a second funder may provide additional resources to allow the country to respond to the challenges the country faces. While an understandable response, this behavior may reduce the incentive on the country to develop a more balanced portfolio of renewable capacities; either in advance, if the country is able to predict the behavior of the second funder, or from that point onwards.

Whereas an RBA is intended to shift the financial risk of non-delivery from the funder to the recipient, in the absence of coordination it can instead transfer financial risk from one funder to another. This materially weakens the incentive of the RBA (Klingebiel, 2012a). To avoid these effects, funders should coordinate.

- **Making contingency plans.** If funds are legally required to be spent in the year they are budgeted for, then a funder could create room for under- or over-performance on an RBA by having a contingency plan. This could involve a 'back-stop', possibly complementary, destination for unspent money, for example, a climate change fund, a research pool, or a proven vaccination program; or the creation of a buffer stock-type trust fund, which could absorb under-spends in one year, and hold them over for potential over-spends in other years. The former approach would take resources away from the program, but, perhaps, divert them to a complementary area, whereas the latter approach would allow for variable performance trends to be smoothed out.

Another potential solution is to commit finance to an independent body or fund. By establishing an independent body, or administering the RBA in partnership with a third party, the funder may be able to credibly tie up funds in a manner that prevents their future withdrawal. An independent third party could serve as an intermediary, allowing the funder to simultaneously provide finance 'upfront' and limit disbursal to the recipient until after results have been achieved.

Lastly, there are circumstances where payments should not be made, even though the formal criteria for disbursement may have been met. A particular such situation would be a deterioration in the wider political circumstances, that does not directly impair the delivery of the RBA-linked results, for example a burst of violence surrounding an election. Under such circumstances, it may be dangerous to continue disbursing funds, since their untied nature (once disbursed) creates the danger of misuse (Klingebiel, 2012b). In order to reduce the challenges that this might create, a political instability clause could be discussed between funder and recipient, so that both parties know in advance the impact of wider political developments on disbursements.

RESULTS-BASED AID WITH THE MULTI-TIER FRAMEWORK

KEY FINDINGS

- The multi-tier framework for measuring access to electricity, developed by ESMAP, offers a unique opportunity to tie aid payments to energy outcomes.
- It measures the usability of electricity supply along multiple dimensions through representative household surveys. These different dimensions are aggregated into a tier score for each household that ranges between 0 (no electricity access) to 5 (high quality accessibility of electricity).
- Payments could be tied to changes in the average tier score. This links well with how the indicator may be used in the SE4ALL initiative, but it may not be flexible enough to allow the reward structure to target particular user groups. For instance, it would require the same payment to be disbursed for a single tier improvement experienced by a user who already had good electricity access as for a single tier improvement experienced by a user with no access to electricity.
- Payments could be tied to individual tier improvements. For instance, $200 could be disbursed for any household moving from Tier 1 to Tier 2, $150 for any household moving from Tier 2 to Tier 3, and so on. This allows the reward structure to be more carefully targeted, but would somewhat complicate presentation and explanation of the program.
- Payments could be tied to individual dimensions of electricity access. For instance, a payment of $100 could be made for each household that experiences improved quality or affordability of supply. This approach provides a simple, easy-to-explain incentive that could be used to target particularly salient barriers.

The discussion in the previous sections of the report could be applied to RBA design in any part of the energy sector, and also has relevance outside of the sector. In Chapter 3, three overarching objectives for energy sector interventions were identified: promoting access to energy for households, communities, and productive enterprises; increasing adoption of renewable energy; and improving energy efficiency. The framework described in Chapters 2 to 6 could equally be applied to any of them, and potentially to interventions in other sectors.

This section specifically examines the potential application of RBA to household electricity access in more depth. This is a particularly important objective within the energy sector, and is

receiving significant attention at the moment due to a number of international initiatives, such as SE4ALL. There is a growing evidence base suggesting that electricity access can have large impacts on the education and wellbeing of households (e.g., IEA, UNDP, & UNIDO, 2010). By focusing on one particular objective within the energy sector, this chapter aims to provide a practical example of an energy sector RBA.

There are a wide range of possible interventions to improve electricity access. Grid extension or the construction of mini-grids can expand electricity access to unconnected households. Off-grid technologies, such as solar home systems, solar lanterns, or diesel generators, can provide electricity services where grids are infeasible. Additional grid capacity, improvements to transmissions and distribution infrastructure to reduce losses, better metering, and a whole range of other interventions can support the development of higher quality electricity access, both for those with and without connections.

The effectiveness of interventions varies widely depending on context and time. The feasibility of providing grid connections versus mini-grid or off-grid solutions depends on a large range of factors, which will vary within countries and over time. The cost of grid extensions will depend on raw materials, labor costs, current grid capacity, the geographic locations to which the grid is to be extended, and so on. Similarly, the cost of providing off-grid solutions will vary extensively; for instance, the costs of solar photovoltaics (PV) have dropped significantly over the last five years. As a result, there is no single, invariant answer to how household electricity access should be promoted–it requires a careful weighing of the advantages and disadvantages of different options within the national and regional context. Ultimately, a variety of approaches are likely to be adopted, even within countries.

Governments will also have different priorities for how electricity access should be promoted. A government may, for instance, wish to support domestic supply chains in renewable energy or improve energy security by reducing fossil fuel dependency. As a result, they may legitimately prefer some energy access solutions to others, even if they are higher cost.

This implies that there would be great value in providing recipients autonomy. As argued in Chapter 4 under "Appropriate incentive effect" and throughout the report, one of the benefits to RBA is the greater autonomy and ownership it allows the recipient. This is especially true in an area such as energy access where there are a broad range of possible interventions, and in which cohesion with the government strategy is essential for long-term sustainability.

Until recently, this has been hard to do within a results-based framework because measurement at the outcome level has been extremely difficult. Electricity access is usually measured using a simple binary indicator; either households have a connection to a grid or they do not. As discussed in more detail below under "Measuring Household Access to Electricity," this measure fails to capture many of the important qualitative aspects of an electricity connection. Two households, both with connections to the grid, can have varying levels of access to electricity depending on the quality of their connection, frequency of unscheduled outages, and so on. It also excludes off-grid solutions that might offer equally good or better service. Given this, a results-based program would have needed to be based on limited intermediate

outcome measure (e.g., electricity connection rate) or disbursed based on more detailed output measures that limit government autonomy (e.g., distribution of solar lanterns that meet minimum quality criteria).

The multi-tier framework developed by ESMAP offers an opportunity to directly measure and target the usability of electricity supply. The multi-tier framework directly monitors, through surveys, the qualitative experience of end users. It is technology neutral and takes into account the wide variety of attributes that affect user experience. As a result, it allows disbursement in a results-based scheme to be tied to outcomes, rather than outputs, and can capture all the relevant aspects of energy usability while still affording recipients autonomy over how to pursue these goals. As identified in Chapter 4 under "Combining the three factors," where indicators capturing outcome level results are available, this is likely to allow for a more robust RBA design.

This chapter will therefore consider how an RBA scheme could be structured using the multi-tier framework. The multi-tier framework offers a unique opportunity for outcome-based RBA in electricity access, but, as a relatively complex and multifaceted indicator, it also poses some unique challenges. The section begins with a broad discussion of both traditional measures of electricity access and the multi-tier framework, before discussing three alternative ways in which the framework could be applied to RBA. The three alternative methodologies considered are:

1 | Tying payments to the average tier score

2 | Tying payments to individual improvements in tiers

3 | Tying payments to improvements in attributes

After considering the advantages and disadvantages of each of these methodologies, a method for payment setting is outlined for each of these three RBA structures.

MEASURING HOUSEHOLD ACCESS TO ELECTRICITY

Historically, electricity access has been monitored with binary measures. These measure whether a household has access, often taken to be synonymous with being connected to the grid, or not. Overall access to electricity within a country or region is then calculated as the proportion of households with grid connections.

However, these measures do not recognize the wide range of factors that determine the user experience of that energy. For example, two households, both with electricity access, can experience extremely different service levels. Limited duration or frequent unscheduled outages can severely detract from the true access to electricity enjoyed by a household, as can low capacity levels. Quality of the electricity access, in terms of recurrent voltage problems, has a significant effect on the user but is not picked up in a simple binary metric.

As a result, binary measures fail to capture important differences in access to energy. All of the variation detailed above is hidden within binary metrics. Understanding access to energy requires a richer set of data.

In response to these challenges, a new multi-tier framework has been developed for measuring users' experiences of energy; one of the first applications of this multi-tier framework has been

in relation to household electricity access. This explicitly recognizes that a number of use attributes may impact a user's energy access experience and so integrates a much broader range of information than simple binary indicators of access. The development of the multi-tier framework has partly been driven by the pledge of the international community, under the SE4ALL initiative, to achieve universal energy access by 2030. To understand what this means and to track achievement against this goal requires a richer measure of energy access than existing binary indicators. The multi-tier framework has been adopted to be used in the SE4ALL Global Tracking Framework.

The framework allocates each household to a particular tier for a number of separate use attributes; the overall household score is the minimum tier score across all use attributes. Figure 7.1 shows the structure of the multi-tier framework for household electricity access. Instead of simply measuring whether or not a household has an electricity connection, the electricity access framework instead combines into a single measure all of the relevant information on the user experience, including the capacity and duration of the connection, frequency of outages, quality of voltage, and affordability. This information is then integrated into an overall tier ranking for each household. More details on the framework can be found in ESMAP (2014) and SE4ALL (2013).

This framework maintains the focus on the user experience. Aspects of access that are irrelevant to the quality of service provided to the user, such as the technology used, are excluded. In theory, any technology could be used to reach Tier 5. It recognizes that, in principle, it is not necessary to be connected to a grid to enjoy a reliable and (relatively) high capacity electricity access.

FIGURE 7.1

The Multi-Tier Framework Measures Access as the Minimum Tier Score across Seven Attributes

Access to Electricity

	Tier 0	Tier 1	Tier 2	Tier 3	Tier 4	Tier 5
Capacity	No electricity	1-50W	50-500W	500-2000W	>2000W	
Duration	<4hrs	4-8hrs		8-16hrs	16-22hrs	>22hrs
Reliability	Unscheduled outages				No unscheduled outages	
Quality	Low quality			Good quality		
Affordability	Not affordable		Affordable			
Legality	Not legal			Legal		
Health & Safety	Not convenient				Convenient	

Source: ESMAP, 2014

An index score can be constructed as the average tier score within a geographical area, and can be used as an overall indicator of access to electricity. The index score is calculated as:

$$\sum_{(k=0)}^{5} (P_k * k)$$

where k is the tier number and P_k is the proportion of households at the k^{th} tier. This index can be compared between different geographical regions for a high level overview of which region has better access to electricity.

APPROACHES FOR SETTING RESULTS-BASED AID AGREEMENTS USING THE MULTI-TIER FRAMEWORK

Application of the multi-tier framework for household electricity access provides three different results indicators that could be used within an RBA agreement. Payments can be tied to:

- the overall index score, which is one indicator measuring outcome-level results on the usability of energy supply;
- the number of households moving from tier to tier, which is an alternative indicator for measuring the same outcome-level results; and
- improvements in a particular attribute, which is an indicator measuring intermediate outcomes.

As discussed in Chapter 4, there are three requirements for a good RBA indicator. The indicator should perform well in terms of:

- proximity to objective;
- ease of measurement; and
- appropriate incentive effect.

More detail on these criteria is provided in Chapter 4.

This subsection considers the relative merits of the different indicators that can be captured by application of the multi-tier framework for household electricity access against these three criteria. As discussed in Chapter 4, it jointly considers both the appropriateness of these results (at differing levels of the energy results chain) and the particular indicator measuring that result level for an RBA agreement.

Overall Indicator

Payments could be tied to the overall index score for a country or region. For instance, a payment of $10 million could be made to the recipient for each 0.1 increase in index score above the baseline.

Proximity to objective

The index score integrates all the relevant household level information on household access to electricity. As described in the previous section, the overall index is an average of individual

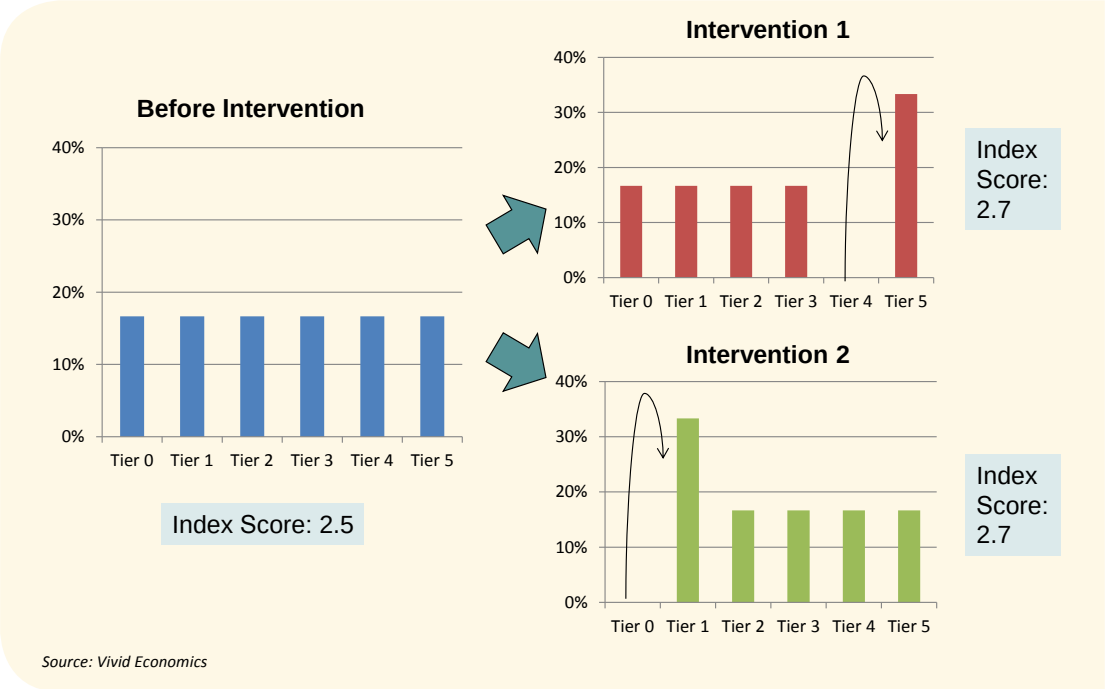
household tier scores. The household-level tier scores themselves are robust and provide a much better measure of access to electricity than simple binary metrics. As an indicator measuring results at the outcome level of the results chain, this approach ensures reasonable proximity to likely objectives; any improvement in overall index score is likely imply real development benefits.

However, aggregating household-level tier scores using a simple average may obscure important information about how an intervention has changed the distribution of household electricity access. An average is just one of many potential methods of aggregating household-level tier scores, with particular consequences for how the overall score changes in response to different interventions.

One consequence of the aggregation method is that shifting a household from Tier 0 to Tier 1 has the same consequence on the overall index score—and payment within RBA—as shifting a household from Tier 4 to Tier 5. For instance, Figure 7.2 shows that two possible interventions, labelled Intervention 1 and Intervention 2, have the same impact on the overall tier score despite having very different impacts on the distribution of households across the tiers. Intervention 1 benefits only those who already had good access to energy, shifting all Tier 4 households to Tier 5. Intervention 2, on the other hand, benefits only those who had poor access to energy beforehand, shifting all Tier

FIGURE 7.2

The Index Score Increases by the Same Amount if a Household is Moved from Tier 4 to Tier 5 as when a Household is moved from Tier 0 to Tier 1



Source: Vivid Economics

0 households to Tier 1. If payments were tied to overall index score, both interventions would imply the same increase in disbursals.

This may or may not be desirable depending on the particular context. In some circumstances, funders and recipient may want to adopt this approach. However, on other occasions, both funder and recipient may wish to incentivize improvements in household electricity access at particular tier levels. For instance, there is significant evidence that the initial stages of electrification provide disproportionately large benefits. Some studies estimate that basic electrification can increase household income by up to 30 percent (Khandker, Barnes, & Samad, 2009), others find a doubling in children's study time (Millinger, Marland, & Ahlgren, 2012). Given this evidence, stakeholders within an RBA focused on household access to electricity may wish to remain focused on outcome-level results on the usability of electricity but use an alternative indicator that better reflects distributional concerns.

Ease of measurement

The index score has been adopted by SE4ALL for use within their Global Tracking Framework (SE4ALL, 2013). Consistency with the metrics of success adopted by international initiatives offers some clear advantages for an RBA. First, there will be synergies in monitoring. Measurement is likely to take place for the purposes of SE4ALL and this may reduce measurement costs for an RBA. Second, it will contribute to the index becoming increasingly well publicized and understood, which may increase its efficacy as an indicator.

As with all measures based on the multi-tier framework, regular household surveys will be required. Under an RBA agreement, the more infrequently that an indicator is measured, the less responsive income for the recipient will be to performance in the program. If measurement occurs infrequently, the recipient may be unwilling to dedicate significant resources to the program. Therefore, use of the index within RBA will require regular (ideally, at least annual) household surveys, which may be expensive. Costs could potentially be reduced by combining with other initiatives to create a survey that had broader applicability than measuring energy access, but this would have to be addressed on a case-by-case basis.

However, the additional information provided can be invaluable in project design and implementation. The benefits of regular surveys go beyond their potential for use in RBA. They allow those involved in the project to track progress, identifying any barriers preventing success as the program is implemented and adjusting design accordingly. The quantification of results allows for an improved process of learning and experimentation which, over the long term, is likely to greatly increase the development impacts achieved.

Appropriate incentive effect

Control over the indicator depends on the agent and the context. This indicator will only be suitable if the agent has good control over it; if the agent's actions have no effect on performance against the indicator, then it will not be able to incentivize them to act. However, whether this is the case cannot be determined in the abstract, but requires consideration of a concrete program and context. It must be looked at on a case-by-case basis.

A **payment scheme based on index score may be difficult to communicate.** The index score indicator combines a huge amount of information, aggregating performance on up to seven attributes for all households within a geographical region into a single number. As a result, the relationship between the index score and improvements in energy access for particular households may not be transparent, at least until use of the index is widespread.

There is a risk that this may weaken the incentive effect. Improving the multi-tier index score from 2.5 to 3.5 is more abstract target than, for instance, ensuring all households have basic access to electricity. One of the major benefits of RBA is coordinating different actors around a single, well-defined goal and this effect may be weakened if it is unclear what, in practical terms, success would look like.

The recipient may lack a clear line of sight between their action and their effect on the overall index score. Because the various attributes that the recipient directly controls are aggregated in a non-linear manner when constructing the household tier score, the effect of a particular intervention on the index (and therefore, on disbursements) can be unclear. On the one hand, aggregating information on multiple attributes reflects one of the strengths of the index; the overall tier score will only increase if households actually experience an increase in access to electricity, so improvements on one use attribute does not translate if performance on another attribute is holding back the usability of electricity. On the other hand, it requires a reasonably detailed knowledge of the construction and measurement of the indicator before agents within the program can understand the effect of their actions on disbursement, which may discourage action.

Conclusion

Performance against each of the three requirements is summarized in Table 7.1.

The index score could be a suitable indicator in some contexts, but rewarding all shifts in tier may be distortionary. It provides decent proximity to likely objectives and allows recipients flexibility

TABLE 7.1

Advantages and Disadvantages of the Index Score as an Indicator for Results-Based Aid

CRITERIA	ADVANTAGES	DISADVANTAGES
Proximity to Objective	Aggregate measure of energy access outcomes within the relevant geography	Rewards all tier improvements equally; this may not always reflect objectives of funder or recipient
Ease of Measurement	Metric has been adopted by SE4All and is expected to be increasingly used for measuring energy access globally	Requires regular household surveys
Appropriate Incentive Effect	Metric gives flexibility for recipients to respond in a way that takes account of local circumstances	The complexity of the indicator may mean that stakeholders find it difficult to understand how interventions will lead to changes in the index score

Source: Vivid Economics

in choosing how to respond to the RBA. While regular household surveys are required, the costs of monitoring may be reduced due to synergies with the SE4ALL Global Tracking Framework, and the additional information provided through the surveys can be used to improve project design and implementation. However, applying the index will mean that all tier shifts are rewarded equally. If this does not reflect the underlying values of funder and recipient, this may distort recipient decisions.

Rewarding all tier improvements equally will not be a concern if distribution of energy access within a country is concentrated within a particular tier. For instance, if most households have little or no access to electricity, interventions will inevitably be aimed at Tier 0 and Tier 1 consumers regardless of the different payment amounts chosen for shifts in tier. But, if access to energy in the baseline survey appears to be evenly distributed across the tiers, then an alternative may need to be considered. The recipient will face choices between which service levels to target, and unless the payment amounts reflect the underlying value of improving access for Tier 0 households versus Tier 5 households, the scheme may encourage resources to be used in a way reflecting the values of the funder and recipient.

Improvements in Individual Tiers

An alternative indicator assigns payment amounts to each tier separately. The approach still combines all of the information provided by the application of the multi-tier framework and so captures the outcome-level result on the usability of energy services. It is simply an alternative (set of) indicators for measuring this result. Table 7.2 provides an example of this methodology. Payment for each tier is calculated as payment per household multiplied by change in number of households at the relevant

TABLE 7.2

Setting a Different Payment Amount for Each Tier Allows the Incentive Scheme to be Tailored Towards Improvements in Particular Tiers

	PAYMENT PER HOUSEHOLD	NUMBER OF HOUSEHOLDS IN TIER BEFORE INTERVENTION	NUMBER OF HOUSEHOLDS IN TIER AFTER INTERVENTION	DIFFERENCE	PAYMENT
Tier 0	$0	30	10	-20	$0
Tier 1	$100	20	15	-5	-$500
Tier 2	$180	15	30	15	$2,700
Tier 3	$240	15	20	5	$1,200
Tier 4	$280	10	20	10	$2,800
Tier 5	$300	10	5	-5	-$1,500
Total					**$4,700**

Note: Payment is calculated for each tier by multiplying change in the number of households by payment per household. Total payment is then calculated as the sum of tier payments, including negative values.

Source: Vivid Economics

tier (including negative values). For instance, if payment at Tier 3 is set at $240 per household and the number of households at Tier 3 increases by five, the total payment for Tier 3 is $1,200.

The total transfer is the sum of payment amounts for individual tier levels. Allowing negative values ensures that the total payment is reduced if backsliding occurs. For example, if the intervention reallocates resources from Tier 5 households to Tier 2 households and, as a result, the number of Tier 5 households decreases, then the payment amount will be reduced in line with the stated payment amounts.

Payment per household should be increasing for higher tiers. For instance, the payment for Tier 2 should be at least as high as the payment for Tier 1. This ensures that improvements in energy access never lead to a decrease in the total payment amount. A useful heuristic for understanding the effects of the payment scheme is to note that the change in total amount received due to a household moving from one tier to another is equal to the difference between the payment amounts for those two tiers. The total increase in disbursals for shifting a household from Tier 1 to Tier 2 in the example above is $180 minus $100, which equals $80. This implies that if two tiers have their payment set at the same level, there will be no change in the total amount paid for movements between the two tiers. If Tier 4 and Tier 5 both have a payment level of $300, for instance, moving a household from Tier 4 to Tier 5 will have no effect on total payment received. On the other hand, if Tier 4 has a payment level of $300 and Tier 5 has a payment level of $100, shifting a household from Tier 4 to Tier 5 will decrease total payments by $200. This should be avoided as it disincentivizes improvements in energy access.

It is possible for the *total* payment amount to be negative if energy access deteriorates. Equally, it is possible for the change in the overall multi-tier index indicator to be negative if energy access deteriorates. It should be determined beforehand what will happen if this occurs. One option is to ignore negative results, only making disbursements for improvements in energy access and taking no action if energy access becomes worse. Another option is to deduct negative amounts from future payments. This ensures that there is no incentive to suppress access for one year so as to create conditions for large improvements (and therefore significant disbursements) the following year. This second option slightly reduces the risk of gaming, and should be preferred unless there are context-specific reasons to think it would not be effective.

Proximity to objective

As with the overall index score, this method integrates all relevant information on access to energy at the household level. The focus is on the outcome level of the results chain, and it integrates a wide array of the relevant information, providing a rich picture of the user experience.

This approach is more flexible than applying the index; it allows payment levels to be adjusted to target particular tiers. If it is believed that consumer (and social) value is concentrated in the initial shift from Tier 0 to Tier 1, the payment levels can be tilted to encourage interventions aimed at the lower tiers. This is the case in the example provided in Table 7.2 where moving a household from Tier 0 to Tier 1 increases disbursement by $100 versus $20 for a shift from Tier 4 to Tier 5. Equally,

if consumer (and social) value is concentrated at the higher tiers, payments can be constructed to encourage interventions aimed at providing Tier 5 access.

Unlike using the index, it would require stakeholders to come to an explicit agreement on priorities. It is necessary for additional processes to be introduced to determine relative payment levels for different tiers. A general approach to achieve this is described in Chapter 6 under "Multiple results indicators" with more detail on how this might work in the specific context of household electricity access discussed further below under "Payment Amount." Nonetheless, this may complicate the negotiation process.

Ease of measurement

The information required is identical to the information required to construct the index score. Regular household surveys are required to determine performance across the six or seven attributes measured within each multi-tier framework.

It, thus, shares the advantages and disadvantages, in terms of ease of measurement, with the index score. The data required dovetails with the SE4ALL Global Tracking Framework, which may reduce costs. The monitoring requirements are more extensive than for binary indicators, but provide additional feedback on the failures and successes of the program which could be valuable for intervention design and implementation.

Appropriate incentive effect

Again, control over the indicator depends on the agent and the context. The indicator's suitability in these terms is extremely important, but cannot be determined independently of the context.

An explicit payment amount per tier may provide a more intuitive hook into the RBA scheme. The presentation in terms of a payment schedule for each tier level may provide a more intuitive hook for potential recipients than the overall index score. Paying $100 per household provided with Tier 1 energy access may be easier to explain and understand than providing $1 million per 0.1 increase in the index score. There is a clearer link between the action of the recipient and the change in disbursements, and the goal may be easier to disseminate across government staff and regional offices.

Conclusion

Performance against each of the three requirements is summarized in Table 7.3.

Tying payments to improvements within a tier is a more flexible approach and will be suitable for use in most cases. Tailoring payment levels to match the value attached to movements across tier levels will typically allow for a greater proximity to objective than the index score, while ease of measurement is identical. There are still potential problems that may need to be overcome; for instance, the aggregation process from use characteristics to household tier remains relatively complex and the negotiation process will have to agree on five payment values rather than one. Nonetheless, this indicator can be applied in a wider variety of settings than the index score and, in particular, will be useful in situations where the index could create distortionary incentives.

TABLE 7.3

Advantages and Disadvantages of Improvements in Individual Tiers as an Indicator for Results-Based Aid

CRITERIA	ADVANTAGES	DISADVANTAGES
Proximity to Objective	Integrates rich measure of energy access outcomes within the relevant geography Flexibility of application allows transparent and explicit targeting of particular tiers	May be difficult for stakeholders to agree on relative priorities
Ease of Measurement	Grounded within the multi-tier framework that has been adopted by SE4ALL and is expected to be increasingly used for measuring energy access globally	Requires regular household surveys
Appropriate Incentive Effect	Indicator gives flexibility for recipients to respond in a way that takes account of local circumstances Potentially easier to explain than the overall index score	Even though it may be easier to explain than the overall index score, the complexity of the indicator may mean that stakeholders find it difficult to understand how interventions will lead to changes in tier rankings

Source: Vivid Economics

Improvements in Use Characteristics

Finally, payments could be tied to performance in specific attributes. For instance, capacity is ranked as Tier 2 if the household has access to electrical capacity of between 50 and 500 Watts and Tier 3 if they have access to between 500 and 2,000 Watts. $100 could be disbursed for every household whose capacity attribute is improved from Tier 2 to Tier 3.

This indicator measures intermediate outcome results within the energy results chain. In contrast to the above two indicators, which are measures of outcome-level results, this indicator is measuring improvements in individual use attributes. These use attributes are intermediate outcomes within the energy results chain.

Proximity to objective

Most programs will be aiming to improve household electricity access as a whole, rather than one aspect of it; this implies that proximity to objective will be weaker than for the other two indicators. Improvements in the nominal electricity capacity to which a household has access, for instance, will not improve the usability of energy if the associated electricity output from that capacity remains unaffordable. This weakness reflects that this indicator is only measuring an intermediate outcome within the energy results chain.

It may, however, be sensible to use this indicator to target one specific use attribute if that attribute is identified as the most important barrier to improved access in baseline surveys. For instance, suppose that access to electricity is found to be poor due to limited affordability, but performance against the other attributes is generally strong. Directly targeting affordability will then be likely to lead to improvements in access, at least in the short term.

Even in this case, there may be unintended consequences if the scheme is long term. Using the multi-tier framework means that if the barriers to access change, the incentive changes with them. This is not the case with targeting attributes. For instance, suppose that in the example given above, affordability is improved for much of the population by capping electricity prices but, as a result, returns for electricity providers decrease and unscheduled outages become more common. When payments are tied to the multi-tier framework, continued strong performance will require that the new barrier be tackled. This is not the case if affordability is targeted directly.

Ease of measurement

Regular surveys will still be required, so the cost of monitoring may be comparable to the alternative approaches. Tying to a specific attribute does not require as much data to be collected as the full multi-tier approach but it will still require regular household surveys. Costs may be lower, since a more streamlined survey design could be adopted. However, attributes are still monitored at the household level and cannot be verified without directly surveying households. Given that regular surveys are required, additional questions covering other attributes of energy access may be added at low cost and will greatly increase the information available for the recipient.

Appropriate incentive effect

Tying payments to an attribute is easy to understand and therefore this incentive effect may be strong. Attributes have a clear interpretation. Improving capacity of a household's electricity connection directly corresponds to observable changes in access.

This could provide a strong coordinating effect. The goal of the program will be transparent and easy to disseminate across government. Agents operating throughout the program are more likely to have a good understanding of how their actions affect disbursements.

Conclusion

Performance against each of the three requirements is summarized in Table 7.4.

Tying directly to an attribute will typically only be suitable when there is one specific barrier restricting energy access which has been identified within a baseline survey. In other cases, proximity to objective will be poor. Ease of measurement is also not greatly improved over the other two approaches, since regular household surveys will still be required. However, targeting a particular attribute does provide a clear and strong incentive effect, so the result may be effective in the context of a short-term program to tackle a specific barrier.

TABLE 7.4

Advantages and Disadvantages of Attributes as an Indicator for Results-Based Aid

CRITERIA	ADVANTAGES	DISADVANTAGES
Proximity to Objective	Local circumstances may justify concentrating on a particular attribute when it is the key barrier to improved access	Loses emphasis on holistic measure of energy access Larger risk of unintended consequences or gaming
Ease of Measurement	May be easier to verify than index score, depending on attribute	Requires regular household surveys
Appropriate Incentive Effect	Simple, powerful incentive that can be used to tackle specific barriers	

Source: Vivid Economics

PAYMENT AMOUNT

Depending on the approach chosen, either a single or multiple payment amounts will need to be set. For RBA schemes based on the index score and improvements in attribute, only a single payment level will need to be set. For RBA schemes using a payment schedule for improvement in tiers, payment will need to be set for each tier.

As emphasized in Chapter 6 under "Payment Structure," the payment amount needs to be determined through negotiation between the recipient and funder. Higher payment amounts will allow the recipient to receive more surplus for every result that they deliver, whereas lower payment amounts will allow the funder to receive more surplus. There is no 'correct' division of the surplus gains from a program. Instead, the proper division should be the outcome of a negotiation that both parties are happy with.

A reasonable payment level should:

- be sufficient to cover the costs of achieving improvements in the index, taking into account recipient contribution and other sources of funding;
- be comparable to existing aid flows; and
- offer reasonable value for money for the funder.

Payment Level for Improvements in the Multi-Tier Index Score

As suggested in Chapter 6 under "Payment Structure," a reasonable starting point for negotiations could be found by running some modeling exercises across a number of scenarios. Each scenario would draw on historical and cross-country evidence to construct a reasonable trajectory for the index over time. For instance, a starting point for a scenario could look at the currently planned interventions in energy access and the impact that they are expected to achieve. A scenario with performance above the expected level and one below the expected level would also be

constructed. The impacts of each of these scenarios would then need to be translated into projected changes in the index score, which would require a careful consideration of the benefits that they are likely to provide and how these benefits fit into the multi-tier framework. The end product would be three or more scenarios, covering a range of potential achievements, each of which corresponds to a change in the index score.

Next, the resources required to attain each of these scenarios could be approximated. This would draw on the evidence on marginal and average costs of the relevant interventions. Often the costs of interventions that have been planned will have been estimated, though such estimates may be optimistic and the scenarios should include underperformance relative to expected costs.

Finally, total payment amount could be calculated in each of the scenarios under different payment levels. Since each scenario is linked to changes in the multi-tier index score, it is straightforward to calculate the total payment amount that will be disbursed in each scenario under different payment levels.

The modeling exercise described above should provide enough information to roughly assess the suitability of a payment level. Combined with information on other flows of finance, the cost estimates will be sufficient to assess to what extent the results-based payment is likely to fully fund the intervention, and how much will be required in contributions from the recipient. The projected full payment amount can be compared to aid flows to determine whether it is large enough to provide a reasonable incentive effect. Finally, costs for different scenarios will allow an approximate value for money calculation to be made by the funder. For instance, they could calculate the amount of their resources required per result and compare this to benchmarks from comparable programs that they have funded previously. They could also compare cost per result to a rough estimate of the development value (drawing on the consumer surplus literature) that they expect the result to create. This process is illustrated in Figure 7.3.

Payment Levels for Improvements in Tiers

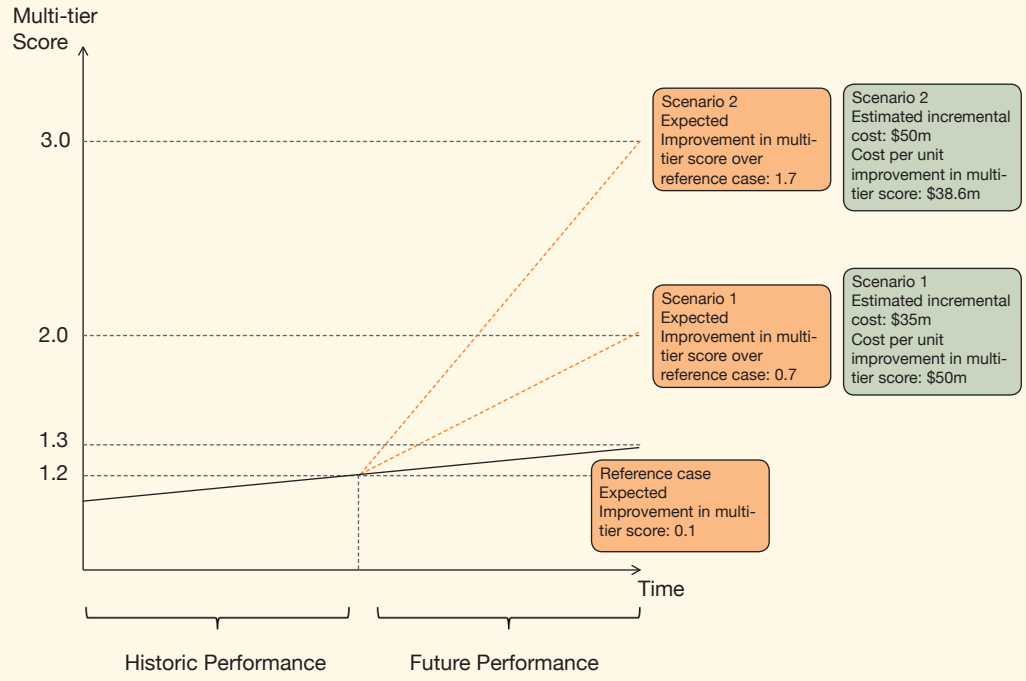
Under this approach different payment amounts need to be set for changes in each tier; as argued in Chapter 6 under "Multiple results indicators," these payment levels should be set based on the relative value of the consumer and wider social benefits associated with each tier. This ensures that the incentive scheme encourages the recipient to invest in results that provide high development benefits per dollar spent. It also provides an objective basis to support negotiations on payment amounts.

In electricity, the costs of improvements in access are likely to increase quickly as each tier is reached. Differ (2013) estimate the costs associated with attaining different tier levels with a (slightly older) version of the multi-tier framework. Providing Tier 1 access off-grid is expected to cost around $70, whereas providing Tier 5 access through grid extension can cost up to $8,250.

The consumer surplus, however, seems likely to be concentrated at the lower tiers. Basic electrification has been estimated to increase income by up to 30 percent (Khandker et al., 2009). Children's study time has been found to increase by more than a factor of two following solar electrification (Millinger et al., 2012). These benefits accrue at the lower access tiers, probably

FIGURE 7.3

The Payment Level for Improvements in the Multi-tier Score can be Constructed through Looking at Multiple Scenarios



Source: Vivid Economics

corresponding to a shift from Tier 0 to Tier 1 or 2. It seems unlikely that higher tiers will provide commensurate benefits.

Unfortunately, there is currently little robust evidence directly linking consumer surplus and wider social benefits to the multi-tier framework. There is extensive evidence on the development benefits of access to electricity and on consumer surplus associated with, for instance, electricity access. However, most rely on more simple measures of access to energy than the multi-tier framework. For instance, the Independent Evaluation Group (IEG, 2008) estimates consumer surplus per kWh of electricity supplied, but this can be linked only approximately with the tiers within the multi-tier framework.

Nonetheless, based on the existing evidence base, some initial rough approximations can be constructed. An illustrative example is provided in Table 7.5. This is constructed by assigning approximate kWh consumption levels to different tiers of energy access within the access to electricity framework. An estimate of $0.05 of consumer surplus per kWh, taken from IEG (2008), is then used

TABLE 7.5

Approximate Consumer Surplus Estimates Using Existing Evidence Base

	CAPACITY (KW)	HOURS/ DAY	DAYS/ YEAR	TOTAL (KWH)	VALUE PER YEAR ($)	PLUS EDUCATION BENEFITS	VALUE RELATIVE TO TIER 1
Tier 1	0.05	4	100	20	$1	$145	1.0
Tier 2	0.50	6	200	600	$30	$174	1.2
Tier 3	2.00	8	300	4,800	$240	$384	2.6
Tier 4	2.00	16	365	11,680	$584	$728	5.0
Tier 5	2.00	24	365	17,520	$876	$1,020	7.0

Note: Access to a kWh of electricity is assumed to generate $0.05 in consumer surplus. In addition, $144 of education benefits is assumed to accrue from access to basic electrification of Tier 1 and above.

Source: Vivid Economics

to translate the approximate kWh measures into estimated consumer surplus. In addition, $144 is added at Tier 1 and above, based on the estimated value of the education effect (IEG, 2008). A more extensive literature review could form the basis of a more robust set of calculations of this type.

Building on these results, modeling could further refine the values used. As discussed in Chapter 6 under "Multiple results indicators," information on marginal costs and other sources of funding could be combined with an estimate for the value of the result to the recipient to model their expected response to a particular incentive scheme. This would help test intuitions about development value where the evidence base is lacking. For instance, if the results suggested that the likely response would be for the recipient to dedicate significant resources to interventions which do not seem to offer good value for money, it suggests that the relative values assigned to the different tiers may need to be altered.

As before, it will be necessary to ensure that the payment levels are sufficient in comparison with existing aid flows and cover the additional costs of the program. Again, scenarios could be constructed based on high, expected, and low attainment against the results. Information on marginal costs and total costs of planned interventions could then be used to produce a rough estimate of the resources required. Flows due to the RBA within each of the scenarios could be compared with other aid flows to ensure that they are sufficiently high to provide a significant incentive.

Payment Levels for Improvements in Attributes

Setting the payment level associated with improvements in an attribute will follow a similar process to setting payments for the index score. Ultimately, the payment level should be set through negotiation, but a starting point can be reached through some basic modeling exercises. Scenarios could again be constructed, each based around performance against the attribute in question rather than the index as a whole. Using these scenarios, different payment levels can then be modelled to determine the resources required for each scenario, the size of payments from the RBA in comparison to other flows, and the cost

efficiency for the funder. Gathering data for analysis may be easier—there are a more limited range of interventions possible to improve performance against an attribute than against access as a whole.

CONCLUSION

The index score may be suitable as a basis for RBA if a baseline survey indicates that energy access is concentrated within one or two tiers. The index score has a number of advantages. Based wholly on the multi-tier framework, it aggregates a wide variety of household data into a single indicator of household electricity access at the outcome results level. As an indicator of outcome-level results, it provides relatively good proximity to objective. Along with the multi-tier framework in general, it is likely to become more widely understood due to its use within the SE4ALL initiative. However, by construction, an RBA based on the index will reward all shifts in tiers equally. This may be inappropriate given that certain tiers are likely to create more value for consumers and society than others. If the baseline survey reveals that the vast majority of households are concentrated in one or two tiers, this may not be much of a concern. Otherwise, a more flexible approach may be more appropriate.

A full payment schedule based on improvements in tier enjoys many of the same advantages, while also being suitable for use in a wider variety of situations. A payment schedule shares many of the advantages and disadvantages of the index score. It measures energy access at the outcome level, and should hence be able to achieve close proximity to objective. Ease of measurement is no more or less complicated than with the index score, and the presentation in the form of payments for tiers may be easier to understand and disseminate throughout the program. It also offers increased flexibility versus the index score; payments can be tweaked to correspond with consumer surplus at different tiers. This process complicates the negotiation process somewhat, but also means that the result is more robust to changing circumstances or to situations with significant heterogeneity in access.

A particular attribute could be targeted if the baseline survey indicates that it forms the main barrier to improved access, but this approach should be used with caution. Targeting an individual attribute provides a strong message that is easily disseminated throughout the program, but as an indicator of an intermediate outcome only, it is further from the likely objectives of any intervention which increase the risk of gaming. It should only be used if there is one clearly identifiable barrier to access, and even then may still be problematic—changing circumstances or unforeseen approaches may still mean that progress against the indicator does not result in improved access to energy.

Simple modeling can provide some initial insights into the range of values at which the incentive could be set. For each of the results, constructing a range of plausible scenarios would allow for some simple calculations of the total flows of RBA payments. This can then be used to assess whether they are comparable to existing aid flows, whether they cover enough of the costs of achieving results and whether they offer value for money for the funder.

When negotiating over a full payment schedule, evidence on consumer surplus and development benefits of different tiers should be used to inform the relative incentive payments assigned to each tier. The payment for Tier 1 versus Tier 5 should reflect the difference in consumer surplus and

wider societal gains between Tier 1 and Tier 5, at least approximately. This ensures that incentives are aligned with underlying development objectives, and the RBA scheme does not unwittingly provide the recipient with incentives to invest in less efficient interventions. While the current evidence base on development impacts does not link directly to the multi-tier framework, a detailed review would be useful in informing their design.

Iterative modeling can help clarify intuitions. By collecting data on the marginal cost of different interventions and the approximate value placed on tier improvements by the recipient, their response to a payment structure can be approximated. This can be used iteratively to check the consequences of a payment schedule derived from data on development impacts against intuitions about which interventions are likely to be cost effective.

ENDNOTES

[1] One notable exception is the Energy+ initiative led by the Norwegian government, which intends to use RBA instruments in pursuit of objectives in the energy sector. However, the associated country programs are still in a preparatory phase of capacity building.

[2] P4R is a World Bank lending instrument that links the disbursement of funds directly to the delivery of defined results with a special focus on strengthening institutions.

[3] In the MCC program, countries are eligible for an aid grant if they score higher than the median score for their peer group on 17 independently compiled indicators intended to demonstrate good governance, investment in human capital, and economic freedom. Countries that are eligible can then apply for a grant for a specific project. RBA, by contrast, would specifically tie disbursal to the achievement of results, rather than using results to construct eligibility criteria.

[4] This definition is consistent with Klingebiel (2012a) and DFID (2012), but slightly broader. In particular, it does not insist on a per-unit payment schedule.

[5] The energy results chain being developed by ESMAP and used in this report, although already a useful framework and tool, remains a work-in-progress. In particular, further work is likely to be needed in relation to energy efficiency and the energy results chain.

[6] An effective result is a result that, by design, when achieved, automatically implies that the objective is achieved; an effective result is impossible to achieve without achieving the objective.

[7] Hybrid approaches refer to cases where a certain share of the payment amount is made upon the delivery of the output or the attribute and the rest upon the delivery of the outcome or impact.

[8] A 'gamable' result is a result that can be manipulated by the recipient often in an unfair or unscrupulous manner and, as a consequence, the objective is not achieved.

[9] One possible downside of this, as noted above, is that the more flexibility that there is around achieving a particular result, the more difficult it may be to identify whether the processes by which the result has been achieved is appropriate (e.g., consistency with social and environmental safeguards, following of due process).

[10] Multilateral development banks do run trust funds and other initiatives that disburse grant resources.

[11] The payment amount, in the context of an RBA applied to a loan, would be the grant equivalent amount of that loan.

REFERENCES

Birdsall, N., & Savedoff, W. D. (2011). *Cash On Delivery—A new approach to foreign aid*. Washington D.C.: Center for Global Development.

Birdsall, N., & Savedoff, W. D. (2013). A Critical Moment for COD Aid or "How to Be Patient When It Matters ". *Global Development: Views from the Center*.

Booth, D. (2011). *Aid effectiveness: bringing country ownership (and politics) back in*. London: Overseas Development Institute

Caravani, A., Nakhooda, S., Watson, C., McFarland, W., & Schalatek, L. (2013). Climate Finance Thematic Briefing: REDD+ Finance. *Climate Finance Fundamentals 5*.

Chagas, T., Costenbader, J., Streck, C., & Roe, S. (2013). *Reference Levels: Concepts, Functions, and Application in REDD+ and Forest Carbon*. Climate Focus.

CIFOR [Center for Intenational Forestry Research]. (2012). *Analysing REDD+: Challenges and choices*. (A. Angelson, M. Brockhaus, W. Sunderlin, & L. Verchot, Eds.).

DFID [Department for International Development]. (2012). *DFID pilots on payment by results*. London: Department for International Development.

Differ. (2013). *The way towards universal access—Putting value on electricity services*.

Dollar, D., & Svensson, J. (2000). What Explains the Success or Failure of Structural Adjustment Programs? *The Economic Journal*, *110*(466), 894–917.

ESMAP [Energy Sector Management Assistance Program]. (2014). *A New Multi-Tier Approach to Measuring Energy Access*.

Gibson, C., Ostrom, E., & Shivakumar, S. (2005). *The Samaritan's Dilemma: The Political Economy of Development Aid*. Oxford: Oxford University Press.

Honig, D. (2014). *Navigation by Judgement: Organizational Autonomy and Country Context in the Delivery of Foreign Aid*.

IEA [International Energy Agency]. (2013). *Tracking Clean Energy Progress 2013*.

IEA, UNDP [United Nations Development Programme], & UNIDO [United Nations Industrial Development Organization]. (2010). *Energy Poverty: How to make modern energy access universal?*

IEG [Independent Evaluation Group]. (2008). *The Welfare Impact of Rural Electrification: A Reassessment of the Costs and Benefits*.

Khandker, S., Barnes, D., & Samad, H. (2009). *Welare impacts of rural electrification: a case study from Bangladesh*.

Klingebiel, S. (2012a). *Results-Based Aid (RBA)—New aid approaches, limitations and the application to promote good governance*. Bonn.

Klingebiel, S. (2012b). Results Based Aid–a New Aid Label or Potential for More Impact. *NORRAG NEWS*, *No. 47*, 76–78.

Millinger, M., Marland, T., & Ahlgren, E. O. (2012). *Evaluation of Indian rural solar electricification: A case study in Chattisgarh*.

Perakis, R., & Savedoff, W. D. (2014). *An Introduction to Cash on Delivery Aid for Funders* (Vol. 20036).

Rasul, I., & Rogger, D. (2013). *Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service*.

SE4ALL [Sustainable Energy for All]. (2013). *Global Tracking Framework*.

Svensson, J. (2003). Why conditional aid does not work and what can be done about it? *Journal of Development Economics*, *70*(2), 381–402. doi:10.1016/S0304-3878(02)00102-5

The Paris Declaration on Aid Effectiveness and the Accra Agenda for Action. (2008; 2005).

Vivid Economics. (2013). *Results-based financing in the energy sector, an analytical guide*.

Öhler, H., Nunnenkamp, P., & Dreher, A. (2012). Does conditionality work? A test for an innovative US aid scheme. *European Economic Review*, *56*(1), 138–153. doi:10.1016/j.euroecorev.2011.05.003


Vestas
Vestas
Vestas

ESMAP
Energy Sector Management Assistance Program

THE WORLD BANK

ENERGY SECTOR MANAGEMENT ASSISTANCE PROGRAM
THE WORLD BANK
1818 H STREET, NW
WASHINGTON, DC 20433 USA
EMAIL: ESMAP@WORLDBANK.ORG
WEB: WWW.ESMAP.ORG

PHOTOGRAPHY CREDITS
COVER AND INSIDE BACK COVER | CABEOLICA / © ASHDEN
INSIDE FRONT COVER | WINDAID / © ASHDEN
TABLE OF CONTENTS | INFOSYS / © ASHDEN
EXECUTIVE SUMMARY | GREENWAY / © ASHDEN
P. 9 | IBEKA / © ASHDEN
P. 15 | GIZ / © ASHDEN
P. 26 | RURAL ENERGY FOUNDATION / © ASHDEN
P. 33, 53 | PEPUKAYE BARDOUILLE / © INTERNATIONAL FINANCE CORPORATION
P. 46 | TECNOSOL / © ASHDEN

PRODUCTION CREDITS
PRODUCTION EDITOR | HEATHER AUSTIN
DESIGN | AUTOMATED GRAPHIC SYSTEMS, INC.