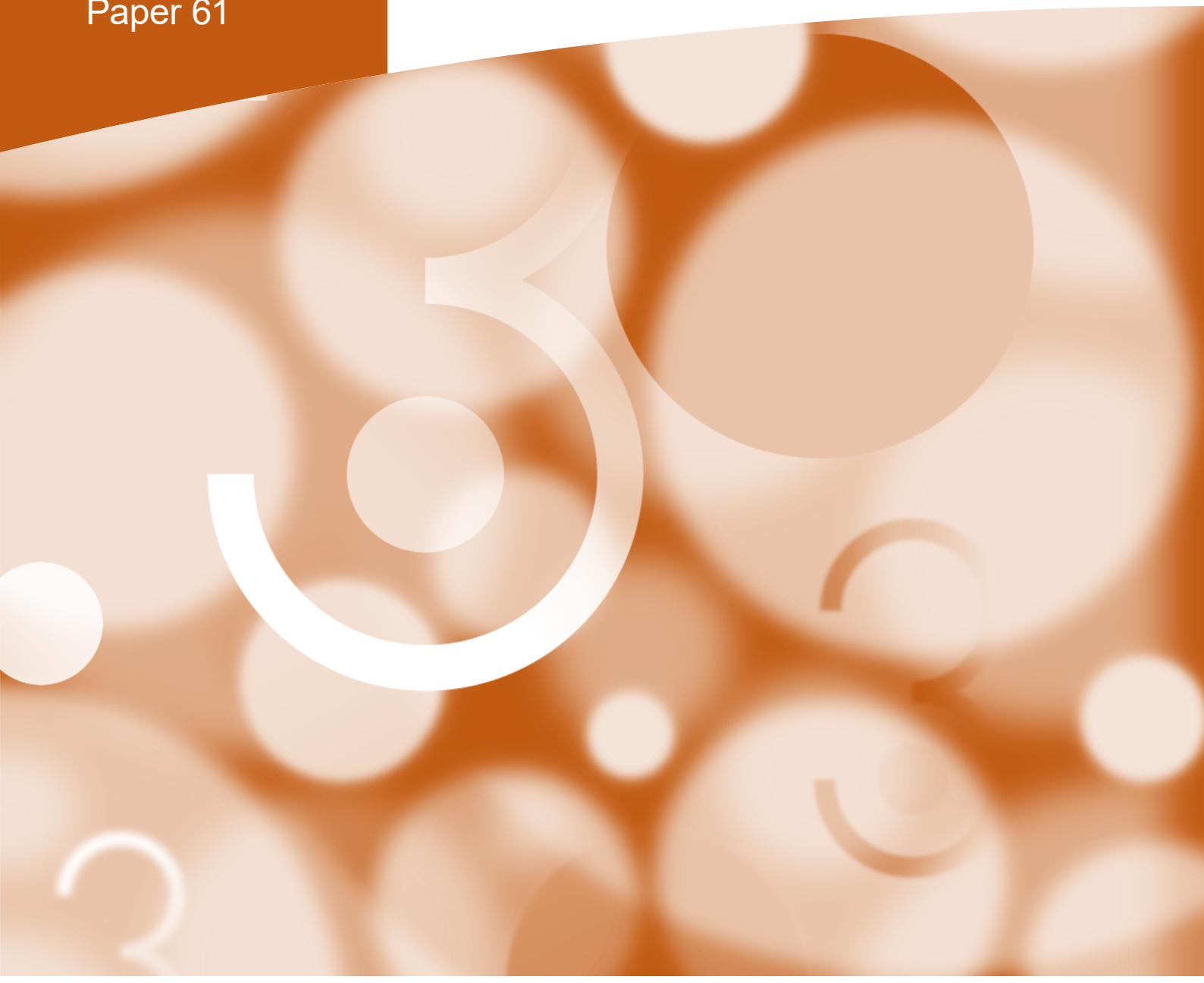
Arnab Acharya
Jane Hammaker
Douglas Glandon

Measuring cost-effectiveness in impact evaluation

May 2024

Working Paper 61



3ie International Initiative for Impact Evaluation

About 3ie

The International Initiative for Impact Evaluation (3ie) develops evidence on how to effectively transform the lives of the poor in low- and middle-income countries. Established in 2008, we offer comprehensive support and a diversity of approaches to achieve development goals by producing, synthesizing and promoting the uptake of impact evaluation evidence. We work closely with governments, foundations, NGOs, development institutions and research organizations to address their decision-making needs. With offices in Washington DC, New Delhi and London and a global network of leading researchers, we offer deep expertise across our extensive menu of evaluation services.

3ie working papers

3ie working papers cover diverse content, ranging from methodological guidelines, discussion of findings and lessons learned from 3ie's research and policy initiatives to address the enduring challenges faced by development policymakers, program managers, practitioners, and researchers. Papers focusing on methods and technical guidance help advance understanding, design, and use of rigorous and appropriate evaluations and reviews. Papers in this series may also synthesize or draw on relevant findings from mixed-method impact evaluations and systematic reviews funded by 3ie, as well as other rigorous evidence to offer new analyses, findings, insights, and recommendations. Please direct any comments or queries to the corresponding author, Jane Hammaker, at jhammaker@3ieimpact.org.

About this handbook

Measuring cost-effectiveness in impact evaluation offers practical guidance on how to design, implement, and report metrics of cost-effectiveness in experimental and quasi-experimental impact evaluation designs. The content of this handbook is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors, or its board members. Any errors and omissions are also the sole responsibility of the authors. All affiliations of the authors listed on the title page are those that were in effect at the time the handbook was published. This handbook has been copyedited and formatted as per 3ie style.

Suggested citation: Acharya, Arnab, Jane Hammaker and Douglas Glandon. *Measuring cost-effectiveness in impact evaluation*, Working Paper 61, New Delhi: International Initiative for Impact Evaluation (3ie). Available at: DOI http://doi.org/10.23846/WP0061

Technical editor: Sebastian Martinez and Thomas Kelly
Managing editor: Tanvi Lal
Publications and web design: Akarsh Gupta
Production manager: Mallika Rao

Measuring cost-effectiveness in impact evaluation

Arnab Acharya
International Initiative for Impact Evaluation (3ie)

Jane Hammaker
3ie

Douglas Glandon
3ie

Working Paper 61

May 2024

Acknowledgments

The authors are grateful to Elizabeth Brown, Sebastian Martinez, Alessandra Garbero, Grayson Sakos, Isabelle Stordeur, Sridevi Prasad, Monica Jain, Charlotte Lane, Shreya Banerjee, Carolyn Huang, Daniel Handel, Tanvi, Tanvi Lal, and Akarsh Gupta for their support in shaping this handbook.

Note from the Authors

We are pleased to announce the publication of 3ie's *Measuring cost effectiveness in impact evaluation*. This handbook provides a comprehensive field guide for implementing cost-effectiveness analysis (CEA) in impact evaluation. It aims to add value to the existing body of CEA literature by providing 1) practical guidance on integrating CEA into evaluation, including during the evaluation design, implementation, and reporting phases; 2) a case study to demonstrate empirical applications of CEA, with illustrative calculations to support intuition behind key methodological steps; and 3) generalizable guidance that can be applied to multiple sectors.

This handbook remains a work in progress, and its utility is contingent on meeting the evidence needs of our partners. We will continue to refine and iterate these guidelines based on piloting and user feedback. To this end, we are currently expanding our scope to include other common CEA applications, such as scale-up decision-support, using incremental cost-effectiveness ratio (ICER) thresholds, and budget impact analysis. We welcome all inquiries, critiques, questions, and suggestions. Please write to us at info@3ieimpact.org.

Content

List of boxes, figures, and tables

Glossary

Defining the intervention and study scope

Activities: Activities are the key components of a project and the organizing principle for a project's inputs. For example, enterprise development training and vaccination camps.

Budget*:* The budget is the monetary amount allocated for a project. The costs of resources used for project implementation may differ from the budget.

Cost Capture: The actual collection of data on the costs and prices of ingredients or resources used in the implementation of a project.

Cost Center: Cost centers are the major parts or components of projects or activities that have costs attached to them. These are typically reflected in key administrative structures in the implementation process, such as mobilizing participants and implementing training and refresher courses.

Input: This term refers to components of a project, such as goods, labor, and services. Inputs could include farming equipment, vaccines, community mobilizers, and awareness brochures.

Impact: The effect of a project on its outcomes, measured using methods that establish causal attribution of an observed change in an outcome to the project under study. Note: This handbook often uses 'outcome' and 'impact' interchangeably.,

Log-frame*:* A logical framework (logframe) outlines the steps of project implementation, which is sometimes also called logic model.

Output: The quantity of goods or services that a project delivers to participants (USAID 2021), for example, the number of training sessions conducted and the number of vaccinations delivered.

Outcome: This is the pre-stated metric that may change as a result of participation in a project. For example, changes in factors such as women's incomes and the prevalence of measles after the intervention of a project.

Perspective: Perspective asks who is incurring the project costs, and at what price. Different stakeholders may incur different prices for identical project costs. For example, an NGO may face different prices than a government agency to procure project inputs. Perspective should be stated and justified during the phase in which the project is defined.

Developing the Cost Collection Template

Activity-Based Costing: This is an approach to cost-capture that collects the costs of the inputs used in carrying out project activities, as defined by the project theory of change or logical framework (log frame).

Cost*:* Costs are the economic value of project inputs, which may include the market value (price) of inputs and the opportunity cost.

Costing: Costing is the collection and reporting of information on project costs.

Costing Template: A data collection instrument for organizing and capturing a project's costs.

Ingredient Approach: The ingredient approach (IA) estimates the cost of an intervention based on listing and valuing the cost of all ingredients required to implement the intervention (Levin and McEwan, 2001).

Prospective Cost Capture: This is data capture that is planned before the project is implemented. Data collection occurs as the project progresses, and is meant for the purpose of using actual cost data for an ex-post analysis. For example, the cost-effectiveness of the project once it is completed.

Retrospective Cost Capture: This is data capture that occurs after the project is implemented. Data sources may use project data, secondary sources, or historical data on similar projects.

Collecting and Adjusting Costs

Amortization: A method of deriving yearly opportunity costs of capital.

Capital or Fixed Costs: These are project inputs that are unaffected by changes in project outputs. For example, buildings that are rented for farmer training. The rental cost will not change even as participation in the project varies.

Currency Adjustments: Many goods in a development project are imported. To express the value in terms of domestic opportunity costs, it is sometimes advised that the official value of foreign currency through the domestic currency be adjusted to reflect the true opportunity costs of the currency. This is because the currency value of many low- and middle-income countries is not set through a market mechanism.

Cost Sharing: In the production of a good, inputs may be shared from the production of some other good or goods. Thus, any joint production from an input must be acknowledged. The costs of the input use will not be identical to the purchase price of that good. Adjustments will need to be made to reflect that the input has shared use and, therefore, the costs are shared.

Discounting: A cost-adjustment method that accounts for the valuation of money in the future compared to the value of money today. As economies grow, one can save today to increase earnings in a later period. Monetary value in the future is typically discounted by the economy's growth rate.

Exchange Rate: The currency of one country will have a certain value in exchange for another country's currency. This value is called the exchange rate, which can be determined through market mechanisms, or be set by a country with respect to the value of its currency.

Economic Costs: Economic costs of goods and services are the opportunity costs to the society of their usage. When the term economic evaluation is used, it should be understood as an assessment that takes into account societal costs and benefits. It is

possible that the price someone pays for a good does not reflect the true opportunity costs to that person. The term economic cost can apply to private costs as well.

Indirect Costs (Induced Costs): These are any costs that are imposed on society due to the project being in place, but are not the direct costs. It is the negative financial impact that the project generates, as well as the cost undertaken to minimize any negative costs. For example, wells may be unsafe for livestock, necessitating some precautions that impose costs.

Inflation Adjustment: CEAs report all costs valued at a point in time, while activities incurring costs take place at different times. Since prices fluctuate across time, one must adjust prices so that they can be reported as valued at a single point in time.

Opportunity Cost: The opportunity cost of a resource is defined as the best alternative usage that can be made of the resource by the user (Levin and McEwan, 2002), for example, the value of a farmer's wages that are forgone while participating in a training. The purpose of accounting for opportunity costs in costing is that oftentimes, the market price of an input or resource does not reflect the 'true cost' of that resource. For example, for an NGO that rents office space at a subsidized rate, the price of the rental is less than the value of the 'next best alternative'. This difference should be documented and accounted for when costing.

Price: This is the amount of money required to purchase a particular good in the market.

Shadow Price: This is the underlying value that reflects the true cost of an ingredient used in the production of a good or service. Many goods are subsidized, so their actual costs may differ. Adjustments will have to be made in estimating the cost of an item to reflect the true costs.

Variable Costs*:* These are project inputs that change as project outputs change, for example, the costs of staffing trainers to provide training for farmers. The costs of hiring additional trainers will increase as participation in the project increases.

Reporting Costs and Impact

Cost-benefit Analysis (CBA): CBA compares an intervention's costs to monetized outcomes.

Cost-utility Analysis: CUA compares an intervention's costs to utility – a measure of value based on stakeholder preferences.

Cost-effectiveness Analysis (CEA)*: CEA compares an intervention's costs to the (non-monetized) outcomes produced.

Cost-effectiveness Ratio (CER): CER is calculated as project costs divided by impact.

Cost efficiency Analysis: Compares the costs per output produced by an intervention.

Ex-ante Analysis: This is analysis of a project that has not yet been implemented and includes modeling and forecasting.

Ex-post Analysis: This is analysis of a project that has been implemented.

Incremental Cost-effectiveness Ratio (ICER): ICER measures the ratio of extra expenditure or costs incurred due to a project being put in place, and the corresponding additional unit of impact.

Sensitivity Analysis: This helps to generate confidence interval around CER or ICER, undertaken together with uncertainty analysis. Sensitivity analyses are based on certain key assumptions. They can be conducted by varying the discount rate, particularly in the case of environmental projects, exchange rates or other price indices.

Uncertainty Analysis: This helps to generate confidence interval around CER or ICER, undertaken together with sensitivity analysis. Often, values of effectiveness should be understood as a point estimate with a standard deviation. If cost data comes from multiple sites, the point estimate is also stochastic. These values lead one to offer a confidence interval as ICER or CER.

1. Introduction and Motivation

1.1 What is cost-effectiveness analysis?

(CEA) brings together a project's costs and impact. It reports the value-for-money of a project by calculating the costs of achieving a specified output or outcome.

CEA can be conducted ***ex-ante*** (before the project is implemented) or ***ex-post*** (after the project is implemented). *Ex-ante* analyses may be used for modeling to inform investment decisions, or for a project that will yield outcomes in the future, to estimate future impact. These assessments rely heavily on secondary cost benefit data sources and assumptions of analysts. *Ex-post* analyses use primary data on project costs and impact that can be causally attributed to the project of study to assess alternatives.

Conducted as ex-post analysis, CEA can – and in many cases, should – be integrated into impact evaluation. It is a natural fit for evaluation because cost capture can be integrated prospectively into the evaluation design, and subsequently combined with the change in effect sizes that the evaluation estimates.

This handbook focuses on conducting *ex-post* CEA as a component of experimental or quasi-experimental impact evaluation.

1.2 What are costs?

In economic evaluation, costs are the opportunity costs of inputs or resources for project implementation.

Intervention inputs may include direct costs, such as personnel time, facility rent, and material costs. Often, only some of an intervention's costs can be estimated through the direct market value of inputs. There may be some inputs that incur indirect costs, or do not incur direct spending by the intervention; for example, an intervention that requires 50 hours of participant time spent in training. The cost of this input could be estimated using an opportunity cost of participants' time, such as local wages in the region of implementation. Other examples include donated resources; the cost of these inputs could be, but not always, estimated using the market price of the resources.[1]

1.3 Why include cost-effectiveness analysis in development programming?

Public resources have competing usage, so the efficiency of resource allocation for development projects in producing social outcomes matters. In contexts where need is high and financial resources are scarce, it is essential

[1] Of course, the opportunity cost depends on whose costs we are considering. The opportunity cost of an action is the highest value that a person would assign to an action among the set of all actions that the person forgoes in order to carry out that action. Thus, the cost of working at a job is the valuation of whichever among the alternatives (spending time with friends and family, volunteering, or doing nothing) has the highest value. When someone purchases dinner from a restaurant, the valuation of the alternative is the cost of ingredients of a home cooked meal and the time spent in cooking, assuming the person has to eat. When costing from the societal perspective, we follow the same logic by valuing costs as the 'next best alternative usage' of public funds.

to provide policymakers with information on both impact and implementation costs. This allows them to compare results of alternative interventions when deciding how to allocate resources.

Cost-effectiveness analysis is used to answer questions such as:

1. What is the marginal cost per unit of impact of a project, intervention, or policy?
2. Is the project 'worth' implementing?
3. Is the project more cost-effective than 'business as usual'? Or, what is the extra cost per additional unit of impact?

The policy implications of bringing cost and effectiveness together are that we can choose among project options and be confident that implementing a project is financially justified. With accurate cost data, cost analysis can also inform decisions on resource planning, budgeting for future projects, and scaling up.

1.4 What are other methods for economic evaluation?

CEA is not the only method to analyze cost and impacts. Cost utility analysis (CUA) is used if the outcome is expressed as a measure of utility (for example, Disability- or Quality-Adjusted Life Years). Cost efficiency analysis is used to compare the costs per project output and cost-benefit analysis (CBA) is often conducted to compare cost to monetized project return on investment (ROI) or monetized project impact (Glandon *et al.*, 2023).[2] CBAs are generally used as an *ex-ante* cost analysis for large, multi-component programs with multiple and different outcomes, often to determine whether a planned investment is expected to meet or exceed a threshold rate of return. For these analyses, future costs and benefits are usually based on information from secondary sources, such as budget data from similar projects.

1.5 Why measure cost-effectiveness?

Economic evaluation methods vary by sector. When comparing the costs and impacts of social programs, many outcomes cannot easily be monetized, such as women's agency, resilience to climate change, or government accountability. Given the difficulty of monetizing the benefits of development projects, in some cases it is more appropriate to assess project performance through cost-effectiveness. CEA is also useful for comparing programs within comparable contexts that use project outcomes with the same unit of measure. As compared to CBA, CEA is

[2] Some evaluations use outcomes in which monetary value can be estimated or measured, such as changes in crop yield, household income, sales, access to markets, or household wealth. Other outcomes such as productivity, maintenance of the ecosystem, forestry management, household decision-making, resilience to changing weather patterns, and women's empowerment are generally challenging or not feasible to monetize. It could even be the case that the monetary value of an outcome may underestimate benefits, for example, a positive change in a monetary benefit such as women's income may not account for or underestimate changes in women's agency or influence in decision-making. CEA is a practical way to analyze ratios of costs to benefits for interventions in which monetary benefits are not easily ascertained, or even accrue far in the future. When it is possible to monetize project outcomes, CBA should be conducted. This is elaborated in Section 1.2.

relatively more useful for specific projects that can be replicated or need to be scaled up.[3]

1.6 How is cost-effectiveness analysis conducted?

The 'costing' component of cost-effectiveness analysis is conducted by collecting and measuring the cost of project implementation. These are the economic costs of the processes that allow the project to be in place before the project results.[4] In **prospective cost capture**, costs are captured and documented as the project is implemented.

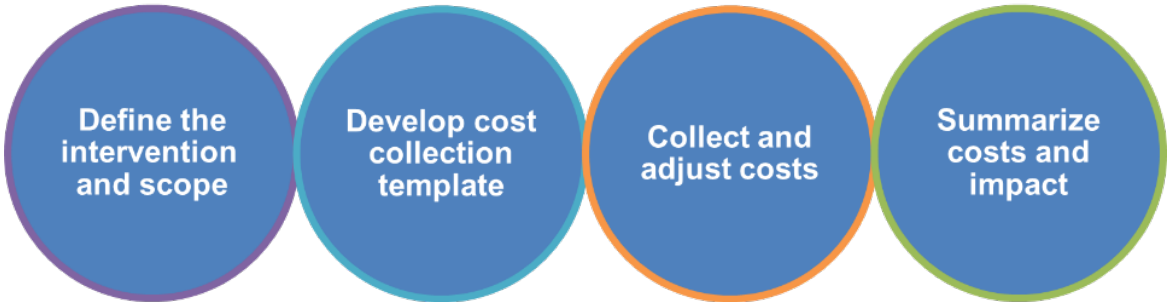
The other component of CEA is effectiveness. In the past decade, considerable efforts have been made to rigorously measure the impact of interventions using methods that establish causal attribution. Since there is extensive literature on impact evaluation methods, this handbook does not discuss evaluation methods for establishing causality.

In this handbook, CEA is broken down into four key steps, which are referred to as the 'CEA framework': 1) defining the intervention and study scope; 2) designing a cost collection template to capture costs incurred throughout the intervention lifecycle, 3) collecting and adjusting cost data as the intervention is implemented; and 4) reporting costs and impacts (Figure 1).

[3] Technical note on CBA vs CEA: CBA usually calculates an internal rate of return on the investment to weigh the cost of a project to its effects, presented as a 'net benefit'. It is well suited for a large investment, usually infrastructure, undertaken at time 0, the current time, yielding a stream of impacts in the future. For large infrastructure projects, analysts often assess impact at some time after implementation. This analysis would inform future decision making processes. Cost-effectiveness analysis (CEA) is often preferable when there is reluctance to assign a monetary value to the project impact. CEAs allow comparability within the specified impact. For example, a project to empower women with certain skills may yield monetized gains; however, it may also lead to reduction in physical violence against women. It is likely that policy makers may not want to monetize such an outcome.

[4] Technical note on implementation costs: The discussion in this handbook will be limited to examining the costing of project implementation. Generally, going beyond costs for projects, there may be costs of implementation not accounted for, such as extra costs associated with implementing environmental regulations or anti-discrimination laws. It is important to be clear about distinguishing costs of project implementation from negative or positive effects (for example, externalities) of the project once it has been implemented. For example, although a CEA of environmental regulation will assess the direct costs of implementing the regulation, it will also need to include losses in certain sectors. Environmental regulation in a town can result in some job losses while improving the health of many people. Although costing of negative consequences would follow the same economic principles that will be used here, the methodology used to assess those costs are specific to the project. It is convenient to separate out resources spent for the implementation of the project from the distributive consequences of the project once implemented. The implication of this focus for the document is that methods for extracting project results are not highlighted, and we indicate that we offer an overarching method for costing which can be adapted for different types of projects that most development organizations engage in. In the later sections, where implementation costs are juxtaposed with impact, we clearly distinguish between CBA and analyses that assess costs and impacts that are not monetized.

Figure 1*:* A Framework for CEA



1.7 Learning objectives for this handbook

This handbook illustrates how cost and effectiveness information can be collected and interpreted, with a primary emphasis on planning and developing tools to capture intervention costs.[5] It focuses on CEA conducted as ***ex-post* analysis using prospective cost capture**, in which cost data is collected as the project progresses. The authors emphasize prospective cost capture because project costs may differ substantially from budgeted costs, expected costs, and secondary cost data sources.

This handbook offers an overarching method for costing which can be adapted for different types of development projects. The goals of the handbook are:

- To acquaint researchers, implementation teams, monitoring and evaluation (M&E) officers, and funders, such as government and donors, with CEA so that they can determine the need for including CEA in impact evaluation.
- Provide sector-agnostic general guidance for impact evaluators and project M&E teams to design, collect, and use CEA.
- Provide methods that are scale-independent and easy to integrate with experimental or quasi-experimental impact evaluations that are designed to measure effectiveness (or replicability of previously successful interventions).
- Account for opportunity cost of public resources, for example, estimating costs beyond the budgetary costs, to incorporate the value of the 'next best' or alternative use of resources into CEA.
- Apply CEA theory to practice through case study.

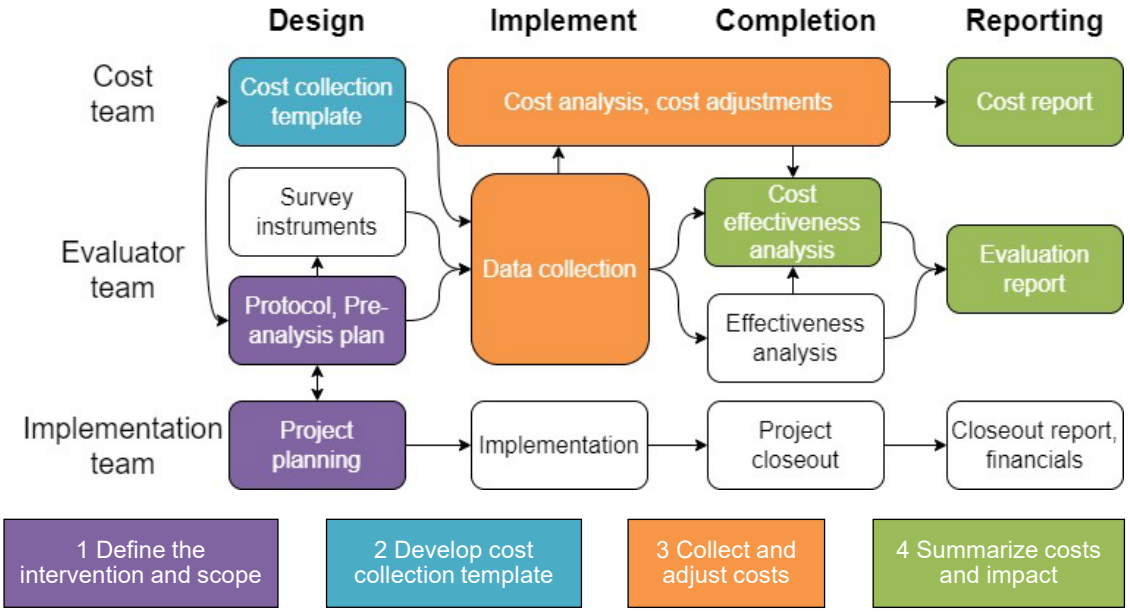
1.8 Planning for CEA in the project lifecycle

Integrating CEA with impact evaluation requires collaboration between three key project roles - implementation, evaluation, and costing teams. The cost team, which may include a costing subject matter expert, cost analyst, or designated team members responsible for CEA, will need to work closely with the implementation team and evaluators to coordinate cost data collection and analysis activities.

[5] There is abundant, freely available technical guidance on the estimation of project effects; this document responds to a lack of practical guidance on accounting of costs of projects.

Figure 2 outlines a workflow in which the cost team, evaluation team, and implementation teams collaborate throughout the four simplified[6] phases of a typical project life cycle: Design, Implementation, Completion and Reporting. For example, in the Design phase, while the implementation team plans project activities, the cost team might contribute cost research questions to the impact evaluation pre-analysis plan or protocol documents. These documents will inform the design of the cost collection template. During project implementation, the evaluation team typically leads data collection activities. The cost team might coordinate data collection activities with the evaluation team, for example, by including a cost data capture tool in baseline surveys. After the project is complete and the endline data collection is finalized, the cost team will conduct the final cost adjustments and calculate the total costs. Total costs, along with the effects estimated by the evaluation team in the evaluation report, will feed into the cost-effectiveness estimates. Although costing activities are separate from activities that bring about changes in the outcomes of a project, it is important that costs are associated with project outputs.

Figure 2: Implementation, evaluation, and cost evidence processes



The shaded boxes in Figure 2 illustrate the four tasks of the CEA framework (Figure 1) that might be conducted during that activity. For example, during project design, the cost team will be responsible for developing the scope of the cost study and developing a cost collection template. Similarly, during project implementation, the cost team will be responsible for collecting cost data which can be analyzed later.

[6] For instructional purposes, the diagram shows a simplified, linear representation of the basic steps involved in collecting and analyzing cost evidence as part of an impact evaluation process. Real-world practice may be more complex. For example, some programs are implemented with an adaptive approach, such that cost information (e.g., a cost-efficiency analysis comparing inputs to outputs) may inform real-time programmatic decision-making during implementation.

Box 1: Guiding case study STARS Impact Evaluation

To connect theory and practice, the authors have applied the CEA framework to an illustrative case study: the impact evaluation of the fictional STARS program. While this case is based on a real intervention, many elements of the case have been simplified for learning purposes.

Stellonia has one of the lowest female labor force participation rates in the world, but many women are closely connected to the informal economy. In rural districts in a western state, for example, women commonly grow raw spices in their home gardens that they sell to 'middlemen,' who process the spices and sell them to urban grocery chains marked upwards of 400% of their purchase price on average.

The Ministry of Rural Development of Stellonia wants to support these female farmers by eliminating the 'middleman' and facilitating women's access to higher-value markets to increase their household income. Over the last year, the Ministry has conducted informal sessions on entrepreneurship through an ongoing initiative. They also conducted a market assessment to estimate the demand for processed spices in urban grocery chains. After determining that the demand is sufficient, in consultation with the farmers and local community development organizations, the Ministry will support the formation of regional collective enterprises. Each enterprise will participate in an 'incubator' and receive start-up capital that includes a centrally located processing center, spice-processing equipment, Enterprise Development and Leadership (EDL) training, and market linkages for processed spices for the first two years.

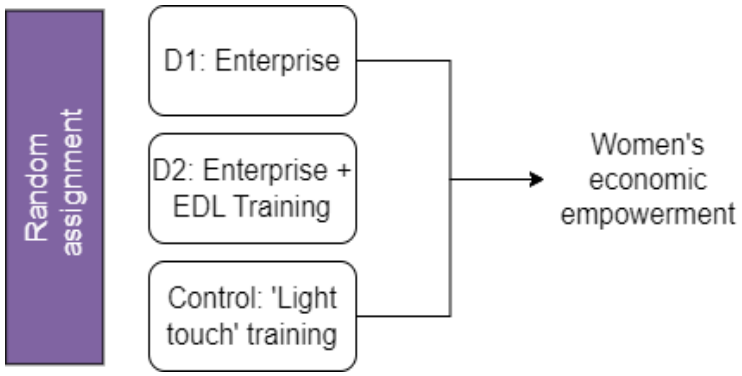
The goals of the initiative are to:

1. increase the productivity and profits of enterprises;
2. increase women's household income; and
3. increase women's economic empowerment such that an additional 40% of women are considered 'empowered' by 2030.

The Ministry needs to know: 1) Is the STARS program effective in increasing women's economic empowerment? 2) Is the training 'worth' the additional investment? 3) What are the costs of the intervention, and how will costs change if the program is scaled up? They have approached 3ie as an evidence partner to help answer these questions.

Due to limited financial resources, the Ministry cannot support EDL training for all 100 district enterprises in the first 3 years of the program. Since there is uncertainty on whether this program will work, the Ministry is interested in conducting a policy experiment. The program will be rolled out as a randomized control trial with two treatment arms: D1) participating in an enterprise, and D2) participating in an enterprise with EDL. These treatment arms will be compared to the 'status quo' or similar communities who will not participate in the program but did receive the 'light touch' entrepreneurship trainings through an ongoing initiative (Figure 3).

Figure 3: STARS Treatment Arms and Outcome



2. Defining the intervention and scope of study

2.1 Developing cost research questions

The first step in the costing process is to work closely with partners to develop a clear and shared understanding of the intervention. This understanding should inform both the impact evaluation research questions and the cost research questions, which guide the development of data collection tools and outcome metrics. Many CEA guidelines suggest thinking about cost analysis as a primary research endeavor, and so it is a key task of the CEA process to integrate cost research questions into the design of the evaluation from the start (Dhaliwal, *et al.*, 2012; Brown and Tanner, 2019; Glandon *et al.*, 2023).

CEA can help implementers, policymakers, and evaluators answer several important questions about an intervention's feasibility, scalability, and impact. Stakeholders often want to know if an intervention or policy is 'worth' implementing. In this case, the cost study may investigate the cost per unit of impact (R1, Table 1). CEA, in this instance, calculates the cost-effectiveness ratio (CER).

In many impact evaluation studies, the primary cost research question could compare two treatment arms of a policy or intervention to assess which is relatively more cost effective. In this case, the CEA will report the incremental cost-effectiveness ratio (ICER; R2). If a new intervention is being tested to replace the 'status-quo' policy or program, the cost study may investigate the cost per change in outcomes of the new intervention, relative to the status quo. In this case, the cost study may assess whether the new intervention produced a more attractive cost per outcome combination than current activities and report an ICER (R3).

CEA is inherently comparative and is interpreted relative to a comparator. The comparator will vary based on the evaluation design. It may be important to interpret the relative cost-effectiveness of an intervention relative to a control group, the 'status quo' policy, another implementation model, treatment arms, or similar programs. The comparator should be clearly defined at the 'define the intervention' stage.

Table 1: Examples of cost-effectiveness analysis research questions

	Theoretical example	Empirical example
R1	What is the cost per unit of impact of the intervention?	A cost-effectiveness analysis of a teacher incentives program found that investing Rs. 10,000 per school produced an estimated 0.27 SD increase in student test scores (Muralidharan and Sundararaman, 2011).
R2	Which treatment arm is relatively more cost effective?	A cost-effective analysis of entrepreneurship training found that, relative to administering an unconditional cash transfer, the Girls Empowered by Micro-franchise program (GEM) cost $568 more per client served. After 14 months, the effects were similar between the two programs. Cash grants (D2) were more cost-effective than GEM (IRC, 2014).
R3	Which policy is more cost-effective relative to the status quo or similar programs?	A cost-effectiveness analysis of a new teacher training program compares the change in test scores of primary school students per $100 spent relative to the *similar education programs* that also aimed to increase student test scores (Dhaliwal *et al*. 2013).

These are questions that can be answered using CEA, but there are other useful cost questions that we can ask when we collect thorough and accurate cost data. Cost research questions should be selected based on the learning priorities of evaluation stakeholders. Box 2 proposes examples from the STARS case study.

Box 2: Developing research questions for the STARS Program

During an evaluation design workshop, Stellonia policymakers, evaluators, and cost analysts brainstormed research questions.

Examples of impact evaluation questions:

- What is the effect of the STARS program on outcomes of interest?
- Can the impact be scaled to other districts?
- Do the impacts last beyond two years?

Examples of cost analysis questions:

- What is the total cost of the intervention?
- What are the cost drivers? How do costs vary by resource usage?
- How do costs vary across comparator and treatment arms?
- Do costs vary by region?

Examples of cost-effectiveness analysis questions:

- What is the cost per unit of impact of an intervention?
- Should it prove effective, is the Ministry of Rural Development willing to invest additional resources in STARS?
- Which STARS intervention modality is relatively more cost effective?

2.2 Defining the intervention

Interventions are implemented to achieve outcomes; we can imagine the project life cycle described in Figure 2 as a production process. The production process mobilizes project inputs (the building blocks of project activities) to implement activities that generate outputs(s), which in turn lead to changes in the well-being of the intended clients or affected population (outcomes) attributed to the intervention. The pathways from inputs to outcomes are often described in the intervention Theory of Change (ToC; Box 3).[7] The ToC also describes the assumptions or risks that influence the effect of the intervention on its expected outcomes.

Costing is the process of assigning monetary values to the activities within the ToC that produce intervention outcomes. Costing starts by taking the ToC and identifying each step of implementation. Eventually, each step will be assigned a monetary value.[8]

This handbook emphasizes that evaluation and costing should be integrated from the start of the project. In case the project has been already evaluated, cost data should be organized by aligning costs with the stages of implementation and the corresponding outputs. The costing team should carefully review the ToC with the implementation team to develop a thorough and shared understanding of the intervention and the implementation process.[9]

2.2.1 Key considerations for planning CEA

The next section identifies and defines key considerations that researchers (for example, evaluators and cost analysts) may need to know during the 'design' phase of the project lifecycle to inform evaluation protocol and pre-analysis planning activities. It also includes illustrative examples from the STARS Case Study (Box 1).

---

[7] A theory of change (ToC) reports how project inputs influence final outputs in a step-by-step manner, and then how the final output influences the project outcome(s) and impacts. ToCs should be accompanied by a written narrative explaining any underlying theory motivating the project design, describing the logical or evidentiary basis for the contributory relationships depicted. ToCs typically include a visual diagram showing detailed theorized contributory relationships between various project components and intended results while also accounting for non-project and contextual factors that may influence the project's implementation and/or results. Other standard elements of ToCs include identifying the recipients and outcomes, duration of the project, the unit of coverage, and assumptions that underpin the theory, ideally including how and when they may be verified or tested. Verifying the steps in a ToC is crucial (White, 2013) for project monitoring (M&E), leading to improved likelihood of achieving project goals. It also reveals what actions are needed to achieve project results. However, the ToC is not fixed, but a living document that may change during the course of project implementation. Referring to the ToC during the process of implementation allows the project implementers to see what modifications are needed. These modifications can change planned activities and costs. ToCs may be accompanied by a (usually tabular) logical framework (log frame) to outline key details about each activity or output, including targets, how and when they will be measured, etc.

[8] Many implementers of intervention develop a logical framework (log frame) for steps that lead to outcome. Each step within the log frame can be thought of as a unit to be costed. Often, it is convenient to find key steps in the implementation process from the log frame and calculate what would be the cost of implementing these steps. Process evaluations look for key features that make implementation successful; a similar effort would identify cost centers in an intervention.

[9] At each step, it important to note that there may be assumptions that affect causal linkages in the ToC.

How can the Theory of Change help in planning for CEA?

A theory of change (ToC) reports how project inputs influence outputs, outcome(s), and impacts (White, 2013). Each step within the ToC can be thought of as a unit to be costed. Often, it is convenient to identify the key steps in the implementation process from the project ToC and calculate the cost of implementing each step, as shown in Box 3.

Participants: Participants are recipients explicitly targeted by the intervention. In some cases, those who were not explicitly targeted but experience costs or benefits through externalities or spillover effects should also be included as participants.

- In the case of the STARS project, participants include farmers (aged 18 and above from two regions), program managers, directors, and field staff involved in implementation.

Intervention activities and inputs: Inputs are resources that supply the intervention's implementation and are central to the costing aspect of CEA. Inputs contribute to the intervention's key activities that are intended to affect change in outputs and outcomes, as stated by the intervention Theory of Change (ToC).

- Examples of inputs from STARS include materials for trainings, transportation for participation, locations for monthly meetings, and spice-processing equipment.

Intervention outputs: Intervention inputs yield outputs that, in turn, produce outcomes. While outputs may not be directly valued, they usually are key to facilitating change in outcomes metrics.

- Examples of intervention outputs from the STARS case study include the number of participants, women farmers trained, machines procured, quantity of household spice production, and the value of farm loans.

Intervention outcomes: The intervention's outcomes are predefined metrics expected to change due to project participation. In many cases, outcomes are difficult to monetize. While the evaluation team will take up the task of measuring changes in outcomes that are causally attributed to the intervention, the costing team will use outcomes to estimate cost-effectiveness. Researchers should agree on the units of measurement to ensure consistency across the cost and impact studies.

- Examples of outcomes from the STARS case study include changes in household decision-making and economic empowerment.
- Since these outcomes are not easily monetizable, the research team will estimate cost-effectiveness of the intervention.

Unit of coverage: The intervention typically operates within regional geographic boundaries, affecting a defined area. Identifying the coverage area helps the cost team estimate the number of participants and the level of cost aggregation required.

- In the STARS case, districts were randomly invited to participate in the program and the district is the unit of analysis in the evaluation. The Ministry is interested in exploring heterogeneity in costs at the regional level, so costs will need to be disaggregated by region.
- There may be overhead costs for monitoring district activities that need to be accounted for in the cost collection template.

Implementation schedule: Awareness of intervention activity timing is crucial for evaluation. Some projects, such as infrastructure, incur costs and have effects far into the future. Implementation teams and researchers should understand how far into the future interventions need to be evaluated and at what points expenditure will occur (O'Mahony, Newall and van Rosmalen, 2015).

- The STARS project involves three days of training in month two of implementation, with refresher courses planned in month six and month nine. The cost team will plan to meet with the program manager in months seven and ten to discuss the breakdown of project expenses by region.

Sample information: Assessing heterogeneity in an intervention's costs by region or other characteristics may necessitate data collection using sampling. Purposeful sampling ensures representation of costs.

- STARS participants who live far away from the enterprise spice-processing facilities may incur higher travel costs. A purposeful sampling strategy will need to be deployed to ensure that cost data is representative of hard-to-reach districts.

Intervention alternatives or comparator: Understanding comparable alternatives is crucial for interpreting CEA results. In some cases, the intervention will replace another intervention; in others, the 'status quo' is no intervention. In cases where the intervention does not replace another program, a comparator will still need to be determined. If it is feasible to assume that the 'status quo' generates zero costs, the team may need to consider conducting a baseline or needs assessment to measure 'status quo' outcome metrics for the analysis. A threshold can also be used as a comparator when it is appropriate to incorporate the decision-makers' valuation of outcomes (see Section 4.4.3).

- The STARS intervention is tested to possibly replace 'light touch' enterprise trainings conducted as part of an existing program.
- The intensity of participation is also of interest, so the STARS evaluation compares two modalities of the program: D1: participation in an enterprise, and D2: participation in an enterprise and EDL training.
- The policymaker is targeting a 40% increase in women's economic empowerment in the next three years.

Perspective: The costing perspective addresses 'who' incurs project costs and at what price (McEwan, 2011). In economic evaluation, it is generally considered best practice to measure costs from the societal perspective and account for societal costs. The argument is that the evaluation should account for the costs and effects of an

intervention on the welfare of the whole of society, and not only on the individuals or implementation organizations directly involved (Byford and Raftery, 1998; Levin and McEwan, 2001). Accounting for societal costs involves adjusting intervention costs for price distortions which are not revealed through prices, and accounting for opportunity costs (see 3.2 'Adjustments').

- In the case of STARS, the cost is estimated from the societal perspective because the project's government partner is accountable to their constituents; therefore, only the opportunity costs of investment are accounted for in the STARS intervention.
- 'Participant time' is a key input of intervention activities that involve attending training. Since the intervention does not compensate participants for lost wages, and there is no direct cost incurred, the cost of participants' time is accounted for using an average local hourly wage. This cost adjustment is an example of costing from a societal perspective.

Whose perspective to cost from?

It may not always be feasible to cost from the perspective of society. For large-scale and complex interventions, for example, the valuation of indirect costs or benefits to society may be uncertain (Cohen, 2020). There could be disagreement on the value of indirect costs of resources used in an intervention, or challenges in data availability for costs incurred by stakeholders outside the scope of study (Levin and McEwan, 2001).

This handbook recommends costing from the societal perspective to: 1) ensure interventions 'do no harm' and account for indirect or unanticipated impacts or costs, and 2) promote the generalizability of CEA findings to other contexts. Since the choice of perspective can have significant implications for the interpretation of CEA, cost perspective used should be documented as part of the CEA methodology (Kim *et al.*, 2020). Multiple costing perspectives can also be considered (for example, by comparing programmatic costs and societal costs) as part of sensitivity analyses.

3. Developing the cost collection template

3.1 Activity-based costing

Once cost analysts have developed a set of cost research questions and thoroughly understand the intervention, they should work together with implementation partners to build a cost collection template to answer the research questions. The purpose of the cost collection template is to identify all of the inputs used in the project activities that produce project outcomes and organize that information in a way that can be readily combined with impact (Brown, 2022).

This handbook recommends the "activity-based costing" (ABC) approach, which organizes costs around the project activities that are usually detailed in the ToC (USAID, 2018; USAID Global Health Supply Chain Program, 2021). The activities-based method

assigns monetary value to project “activities”, which can be any event, discrete unit of work, or task with a specific goal. ABC draws attention to project components or milestones that need to be completed for the intervention to be implemented and helps organize the costing task.

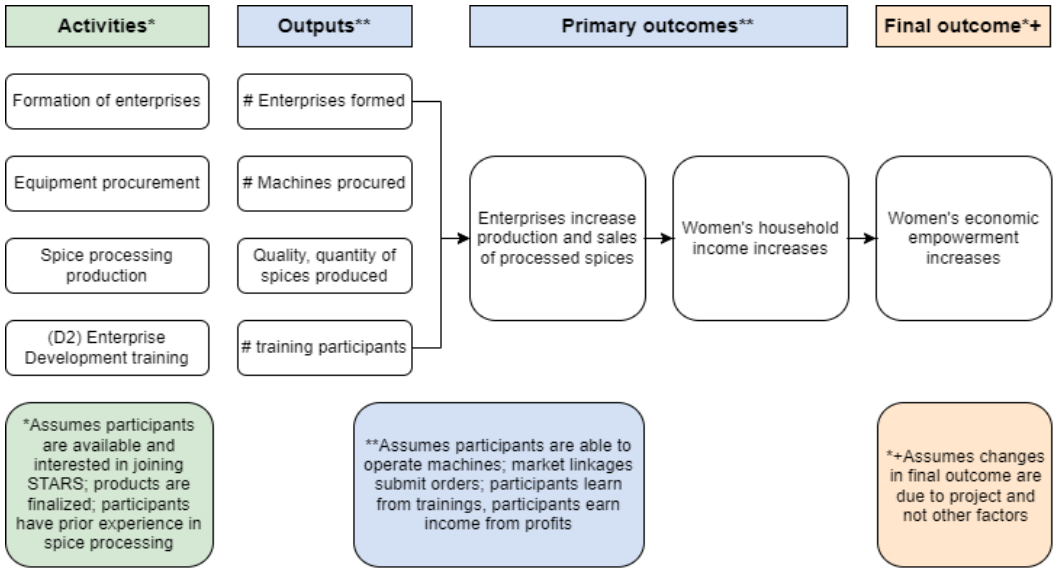
A rule of thumb is to identify a set of activities as a unit that must be implemented. For example, in an intervention that provides sustainable land management training to farmers, key activities might be mobilizing participants, preparing training materials, implementing training, and monitoring or management activities. Activities in the implementation of an intervention can be thought of as cost centers, explained further in Box 3.

Box 3: Identifying Cost Centers for the STARS Program

Key activities from the STARS program are outlined in the project ToC, which has been simplified. Activities include: 1) forming enterprises, 2) procuring spice processing equipment, 3) producing and selling spices, and for the second treatment arm (D2), 4) implementing the EDL training. These activities are assumed to produce outputs that can be verified by several metrics, including the number of enterprises formed or the amount and quality of spices produced. Assuming farmers can use these outputs, the ToC posits that outputs will lead to changes in enterprise productivity and household income and decision-making, which will produce changes in women’s economic empowerment.

These four activities form the cost centers for which the project’s costs will be aggregated. Overarching costs, such as monitoring visits to the enterprise locations to observe production and training activities, are not listed as activities. An additional cost center ‘Project Management’ has been added to capture costs incurred on operations and office equipment for field staff that supported the equipment procurement, spice production, and training cost centers.

Figure 4: Example Theory of Change from the STARS Project



Some activities may be overarching. Many projects expend resources to monitor the overall implementation of the intervention, rather than a specific activity, for example, the office costs of the field manager who oversees field activities, a vehicle used by staff that has multi-usage within the office, as well as field visits to monitor multiple activities. Distinguishing such overarching or cross-cutting activities from the cost centers is an important task prior to data collection.

Once cost centers are identified, each center needs to be linked to resource usage. The next step in the cost collection process is to list all the resources and inputs used to produce each cost center, or the cost 'ingredients' required for implementation (Levin and McEwan, 2001; McEwan, 2011). By treating an intervention's inputs and resources as ingredients, the ingredient approach can make it easier to visualize both the direct and indirect costs of the intervention.

Data collection instruments will aim to list ingredients along with their purchase price of goods, as is seen in the next section.

3.1.1 Listing ingredients

The next section identifies and summarizes five steps to develop the cost collection template.

Step 1: Identify cost centers: Cost centers are the key steps in the ToC that categorize intervention inputs (Box 3). It may also be necessary to include a cost center for overarching activities. Evaluators, cost analysts, and the implementation team should work together to identify the cost centers so that costs are organized in a meaningful and useful way.

Step 2: Identify cost ingredients: For each cost center, the cost analyst will need to list every resource or input used to produce the activity – these are the cost ingredients (Box 4). Examples of cost ingredients include staff time, participant time, transportation, rent of capital goods, training materials, product marketing fees, or official and legal fees. Cost data should be collected on ingredient usage (for example, days, number of trips) and costs incurred (for example, rental payment, salary paid), described in detail in Step 3. Keep in mind that there may be ingredients where the project incurred no price, for example, community or even skilled volunteer labor. It may be challenging to exclusively assign an ingredient to one cost center; perhaps the ingredient is shared by multiple cost centers. These are called "shared costs" and are discussed in Section 3.2.3.

Step 3: Enumerate ingredient usage: Data should be collected on all ingredient usage as well as the prices paid for each cost ingredient within each cost center. These costs should be collected as they are incurred, while the project is being implemented. To enumerate ingredient usage, ingredients are differentiated based on those that incur variable and fixed costs, explained in detail below.

- *Variable Costs* stem from the use of additional intervention ingredients to produce more outcomes, within the current production system. The most common variable input is labor. For example, the number of trainers required could increase as the number of participants trained increases. Material costs may also be variable, for example, training inputs such as booklets or refreshments will also be responsive to changes in the number of participants. Costs will be incurred to use more (or

less) of these ingredients. Some general guidance is suggested on listing labor and other variable costs in Section 3.2.

- *Fixed Costs* are inputs that do not change when the production of an outcome changes. For example, if a program uses large spice-processing machines, most likely, servicing one more farmer will not require the purchase of additional equipment. Adding an additional village to an agricultural extension program would not require the purchase of another vehicle. However, adding 20 more villages might require the purchase of another vehicle. Often, fixed inputs have a time duration for their usage. Thus, they tie up resources which must be adjusted for opportunity costs. Sometimes fixed inputs are rented and at times they are bought to be used for a duration of time longer than the current year. Such fixed inputs may be in the form of buildings, vehicles, machinery, and equipment. Buildings are another common capital good, but if they are rented, then rental prices should be recorded. If they are purchased, their longevity or the duration of usage and depreciation values should be noted at the time of cost enumeration. The purchase price of all fixed inputs should be recorded, although it is sometimes hard to find this type of cost data.

Prices, such as wages, and quantities, such as days worked, should be clearly stated for all fixed and variable ingredients. If relevant, the date of purchase should also be stated. For some ingredients it may be necessary to know the origin and year of the make of the product. Each ingredient has a unit, such as participant hours or the number of machines. For equipment and supplies, brand names and product names should ideally be listed. This process allows for identifying imported goods. Ingredients should always be associated with a unit; see Box 4 for an example.

Step 4: Account for missing data: In a prospective study there still might be missing data, for example, the price paid for a building from years ago, or confidential salary information. Community inputs are also frequently hard to collect. Costing templates should enumerate the need for these types of data. When not found, this data should be listed as missing. If price data is missing, it may be possible to find the prices from other studies or geographical vicinity. If ingredients are missing for a particular activity, one may rely on costing of that activity from other studies. If missing data is substituted through some means it should be described clearly, but the need for imputing data statistically is unlikely.

Step 5: Follow good reporting practice: The collation of collected material should be clear and well documented, with data clearly labeled and variables defined. Although it is a good idea to have a costing template at the beginning of any project, it is possible that revision will be needed, or that there are redundancies in activities and goods purchased. Changes to the template should be documented.

Bringing it all together: When steps one to five are complete, additional adjustments are made to the input and price list, as described in the following sections. After these adjustments are made, the prices of inputs can be added to estimate the cost of each cost center. How the final costs are presented will depend on the output of interest. An application of the cost capture template is provided in Box 4.

Box 4: Listing Cost Ingredients for the STARS Program

Bringing together the activities and cost ingredients, here is an excerpt from the STARS cost collection template using the 'Formation of enterprises' cost center. We recommend gathering at least the following information for each cost ingredient:

1. Unit of time: identifies how the ingredient is measured. For example, the 'staff time' ingredient could be measured in days, while 'travel' is measured in trips per year.
2. Data collection sources: identifies where the cost of the ingredient will come from. For example, the program manager or finance officer will share monthly expense reports that document staff time, travel and administrative costs.
3. Timeline: identifies when cost data is reported or collected. For example, aligning cost data collection with the evaluation baseline and endline.

To report cost summaries by region, cost ingredients will need to be broken down further. For example, to account for differences in STARS staff wages in two regions, an additional column will be added to the sheet to reflect 'Wage Region 1' and 'Wage Region 2.'

Table 2: Excerpt from STARS Cost Collection Template

Cost center	Description	Cost ingredients	Cost description	Unit of time	Data collected from?	Data source	Timeline
Formation of enterprises	Setting up business administration, mobilizing farmers to sign-up	Staff time	Wage	Day	STARS Program Manager (PM)	Expense report	Baseline
		Travel	Cab, Train, Hotel	Year	STARS PM	Expense report	Baseline
		Administrative cost	License fee	Year	STARS PM	Expense report	Baseline, Endline
		Participant time	Wage	Day	Evaluation team	Household listing	Baseline

4. Collecting and adjusting costs

4.1 Planning for cost data collection

Implementation teams and research teams should work together to develop a strategy for collecting data. Collecting cost data can be incorporated into quantitative survey instruments, existing monitoring activities, or as part of qualitative data collection. The research team should identify how and who should collect data during project implementation. The next section answers some frequently asked questions on cost data collection and outlines a few guiding principles for collecting cost data in impact evaluation.

4.1.1 Who provides the data?

Cost data should be collected from the source or the stakeholder who incurs the cost of implementing the project, intervention, or policy. Collecting data during implementation will require close collaboration between the evaluation team, monitoring team (if separate from researchers), cost team, and implementation partners. There is a risk of burdening partners with time-consuming data collection. To mitigate this risk, integrate cost data collection with existing monitoring processes to the extent possible. However, it is important to note that data recorded as costs of project activities may vary based on cost accounting practices in any given context, such as government agencies or NGOs.

It may be necessary for cost data to be collected from multiple sites or through multiple means through which participants access the outcome of the intervention. Sampling may have to reflect heterogeneous characteristics of the recipient group. For example, if an agricultural program is implemented at the village level, multiple villages may make up the sample. It is possible that if villages differ by some characteristic relevant to how the intervention may be delivered, some stratification will be needed.

How can cost data collection fit into impact evaluation?

If implementing a mixed method impact evaluation, one way to collect cost data is as part of qualitative data collection. Key-informant interviews with program staff can provide opportunities to identify cost centers or collaborate on the cost collection template, which may abate the double-burden of data collection.

4.1.2 When will the data be collected?

As discussed, prospective cost capture is recommended as the project is implemented. The timing should be planned according to steps in the project ToC. The evaluation pre-analysis plan should clearly operationalize how data will be collected. The plan should indicate who collects the data, and when to collect the data, and should be mindful that implementation often differs at least somewhat from what was planned.

4.1.3 Is it recommended to collect costs for the comparator or control group?

To interpret the results of CEA, it is essential to identify comparable alternatives to the intervention under evaluation. In some cases, the intervention will replace another intervention. The cost team will need to collect information on these costs to assess the relative change in costs for the ‘new’ intervention. Most impact evaluation studies will measure the effects of the intervention by estimating changes in outcomes in intervention and counterfactual groups. Similarly, costs should be estimated for the counterfactual condition so that relative costs and effects can be compared in the CEA. In other cases, the ‘status quo’ policy is no intervention. When the intervention under evaluation does not replace another program, a comparator will still need to be identified to estimate an ICER; otherwise, a threshold or valuation in terms of currency of the social cost of the extra outcome of interest may need to be considered (see Section 4.4.3).

4.1.4 Can we just use budget data?

In many preliminary or back-of-the-envelope cost analyses, budgets are taken to be the cost of the interventions. The focus on costs as opposed to the budget is important.

Though budgets and spending reports can provide useful data for costing activities, the budget rarely accurately reflects the cost of the implementation of an intervention. Costs focus on the price of resources used in the implementation process. These values will differ once a clear connection is made between resources used to implement the intervention and, subsequently, when the resources are priced appropriately.

How is costing different than budgeting?

There are a few reasons why budgets alone are not enough for accurate cost analysis:

1. Budgets are guess-estimates based on past cost accounting. Unless expected values are replaced with actual costs, actual and expected intervention spending may vary. Actual valuation of resources provides information as to what to do in the future for similar circumstances.
2. Budgets are not based on social cost accounting, but on administrative costs. Relying on the budget may actually under-report the costs of the intervention. For example, budgets may not include information on wages forgone (opportunity cost) to participate in a training or donated resources. Budgets seldom reflect opportunity costs.
3. Budgets typically do not present heterogeneity. In complex interventions with many components, there may be clusters within a budget and unless budgets are being constantly adjusted, costs will not equal the budget. The budget also will not capture variation in spending by location or by participant. Costs may vary by village (for example, remote or hard-to-reach villages) or for some groups of participants.
4. By contrast, cost analyses categorize expenses by the cost center of the production process, rather than by type of spending or by budget head which will not provide the transparency or flexibility that is needed to answer our cost research questions. Cost analyses also add an element of fiduciary accountability by opening the black boxes that budgets tend to be.

4.2 Adjustments to costs

After implementing the cost data collection plan, obtaining non-missing prices or expenditures for each ingredient, enumerating usage, and listing missing data, the cost data collection template may appear complete, but there may still be a likelihood of data missing. For example, is participant time accounted for? Does the price of rented equipment include price distortion? Is the cost of volunteer time included? These are examples of opportunity costs, which are not reflected in prices of ingredients.

Many prices do not reflect the actual cost to the society or the shadow price, the price behind the observed price. For some contexts, there will have to be some adjustments due to price distortion, for example, the prices paid at the exchange of goods and services will have to be adjusted to shadow prices. The purpose of adjusting costs is to better estimate the ‘true costs’, or the economic costs, of all ingredients.

There are two kinds of adjustments to prices that will be needed for the data reported in the costing template: (1) adjustment for time, both in the future (as capital goods are purchased at a particular time and then used for a period of time) and in the past (as costs are reported at a particular time for goods services that may have been used in the past), and (2) non-time price adjustments, such as accounting for shared costs across multiple cost centers (McEwan, 2011).

Time-based adjustments generally apply to all projects, but non-time price adjustments should also be considered, particularly in low- and middle-income contexts. These adjustments use parameter values that describe the economy in which the project is based, with the understanding that these parameters are not always exact.

Note: Sections 3.2 and 3.3 are intended to provide background for evaluators on cost adjustments, and uncertainty and sensitivity analyses. Usually, these analyses are conducted by the cost analyst who supports the research team.

4.2.1 Time-based adjustments

The most common adjustment for calculating costs is adjusting for time. Time adjustments are necessary when intervention costs are incurred over periods in which prices fluctuate. The cost-effectiveness analysis will report costs for goods valued at a single, comparable time. This section describes time-based adjustments for capital goods, inflation (Box 5), and discounting.

Box 5: Time-based adjustment, STARS Case Study

The COVID-19 pandemic affected Stellonia, which experienced changes in the value of money due to supply chain interruptions, stimulus packages and other contextual factors. The nominal costs from each year of project implementation should be adjusted and indexed to a common year. Inflation in Stellonia was 6% in 2019, 8% in 2020, and 7% in 2021, so costs are adjusted to the ‘base year’ which is 2019.

Adjustment for capital goods: Capital goods are usually purchased at one point in time and used for a period of more than a year. In this case, price adjustments for capital goods will be carried out in yearly terms. Alternatively, many interventions rent capital goods. The rental price of goods in low- or middle-income countries is likely to reflect market distortions and should rarely be used in CEA. Most likely, prices will be high due to the presence of inefficient financial markets and constraints to financial accumulation, which limit alternative rental opportunities for capital goods. The cost analyst should determine the opportunity costs of all capital goods that have been purchased or rented to obtain a yearly price and then apportion, as used for the intervention (Box 6).

Box 6: Adjustment for capital goods, STARS Case Study

The STARS team purchased a vacuum-seal packaging machine at a price of $30,000 that is expected to last 20 years. The yearly opportunity cost of $30,000 is the prevailing interest rate, as this amount could have been invested to have a yearly earning; here it is set at 5% per annum. The machine is used for 20 years, so the costs should be distributed across the duration of its usage. For illustrative purposes, we recommend the use of pmt function in EXCEL, = pmt (rate = 0.05, longevity = 20, present value = 30000) = $2,407. In the pmt function, there is allowance for further adjustments that will become clearer when the actual formula is detailed in Annexure 2. For all monetary payments, it is recommended to discount at the prevailing growth rate or the prevailing interest rate of the country.

Adjustment for Inflation: Inflation accounts for the change in the value of money over time. Over time, general prices increase, and money loses value. When there is a general level of rise in price, an economy is said to be experiencing inflation. Often, goods are bought and used at different times during the intervention implementation. However, the cost analyst will report the cost of the intervention as what the goods and services would cost at a particular chosen point in time. To account for inflation, the costs from each year should be adjusted to the price level of a single year. Nearly all governments publish prices of goods and services, which are valued at prices of a single indexed base year. When prices of any given year are adjusted to this price, they are called 'real prices' or 'inflation-adjusted prices.' Unadjusted prices are called 'nominal prices'. The analyst should clearly note the date(s) for which the costs are relevant when reporting costs.

Example 1: Adjustment for inflation

A project was implemented from 2018 to 2021. Goods and services were bought throughout the three years. The costs and impacts of the project will be reported at the end of the project in 2021. For comparison, all project costs need to be reported in 2021 currency value. For example, a good bought for $100 in 2018 would be adjusted using inflation index values in 2018 and 2021; suppose that they are 238 and 250 respectively. The adjustment for $100 is $100 × (250 ÷ 238) = $105.

Example 2: Adjustment for inflation

If the cost L = 200 was incurred at time T and needs to be adjusted to value T-2, then multiple adjustments are needed. Let L stand for labor cost. We start with inflationary adjustment in this box.

Supposing for year T, prices are 2.5 times or 250 when indexed to the base year, and for year T-2 the comparable number is 238. In this example, all costs will be valued at the price level of T-2; that could be the time when the project started. The value of the cost for time T-2 should be expressed as:

$$Cost\ at\ time\ T-2 = L \times \frac{Price\ Index\ at\ T-2}{Price\ Index\ at\ T} = 200 \times \frac{238}{250} = 190.4.$$

Applying a discount rate: Discounting is distinct from inflation. Discounting assumes that you would prefer to have that money now so that you can invest it. There is the additional possibility that each day you wait to receive and spend, that money poses uncertainty and risk that you will not be able to enjoy it in the future, and people are impatient (*Cost-Benefit Discounting*, 2014). The attitudinal part regarding future consumption is called the rate of time preference. Scholars differ as to whether this should be included when we account for social values and costs. Another justification for discounting costs is the opportunity cost of spending today. Had resources been invested by the government, they may have earned interest, or produced a positive return elsewhere in the economy. Thus, this is accounted for when comparing present and future costs using discounting.

Usually, cost analysts use a country's growth rate as the discount rate, but others argue that 5% should be used in most cases. Discounting costs at the rate of growth or at the prevailing interest rate reflects the market opportunity cost in time for consumable goods.

Example 3: Adjustment for Discounting

Building on Example 2, the inflation-adjusted expenditure in time T should be valued when growth has been taken into account in time T-2 as:

$$Discounted\ cost\ at\ T-2 = \frac{inflation\ adjusted\ cost}{(1+r)^2} = \frac{190.4}{(1+0.05)^2} = 172.7$$

Adjusting fixed costs: Although a good is purchased in one period, it may be used over multiple periods. Fixed inputs, sometimes called durable goods or capital goods (for example, buildings, computers, vehicles, and other inputs) do not need to be increased as the production level increases in the short run. In most countries, there is a rental market for capital goods. They can be rented to reflect the regional supply and demand, but the market can be distorted, as the credit markets which foster rental markets for machinery, buildings, or even vehicles are likely to be limited. Thus, there may be friction or imperfection in the local rental market. Even for projects that do not last a long time, the use of rental prices for fixed inputs may thus not be ideal. Amortization should be standard for cost adjustment.

> *Amortization*: To measure the cost of the goods being used each year across time, the method of amortization is used to spread the purchase price across the duration of use. For example, a vehicle may be purchased in one period and used for 10 years. The cost of the use of the vehicle is allocated based on usage after its purchase. This valuation is done through using amortization methods. It allows cost to be ascertained for a single period use given the prevailing discount rate, cost of the item, the intended period of usage, and the depreciation rate.

Example 4: Adjustment for amortization

A building is purchased at a point in time for usage in, for instance, ten years. The amortization method can calculate what it costs to use an item for a single month or a year. A software application such as Microsoft Excel can be used to assist with the calculations. To understand the mechanics, let *C* be the cost of the item when purchased, *δ* be the depreciation rate, *N* be duration of use of the item, and *r* denotes the interest rate.

It is assumed that the item being used can be sold in the future for whatever is left of the good after depreciation. As the sale occurs in the future, it must be discounted. The formula to obtain the current value of the good purchased is the following:

$$V = C - \frac{C\, e^{-\delta N}}{(1+r)^N},$$

where e is the Euler number 2.718. Supposing the item costs $30,000 with a depreciation rate of 10% and the prevailing interest rate (discount rate) is 5% per annum, the yearly cost of the item intended to be used for 10 years is the following: The depreciated present value of the item is $V = \$23{,}225$.

The cost for a particular period *A* is the following:

$$A = V \times \frac{r(1+r)^N}{(1+r)^N - 1} = 23{,}325 \times \frac{0.08144}{0.6288}$$

Using the above formula, one obtains the yearly cost of $3,008. The adjusting with the 'pmt' function in Excel yields the same result.

Calculating the total overall time adjustment*:* Example 5 demonstrates how to combine multiple time-based adjustments to costs that are reported for a single time period. Amortized values use a fixed interest rate, and the value *A* from Example 5 is a constant for the period. Appropriately, it should be both adjusted for inflation and discounted. It is convenient to determine the yearly costs with *A* along with all other costs in the period, then adjust for inflation and the discount rate to obtain a present value.

Example 5: Calculating the overall time adjustment

To illustrate with an example, suppose the fixed cost for a three-year project is 3000 at the amortized value each year and the labor costs are 5000, 5200, and 5300, respectively, for three years. We will determine the present value of the project at year 0. The following illustrates the total cost calculations for an inflation index of 200, 205, 207, and 210 for four years, respectively, with a discount rate of 5%.

The total cost for each of the three years at the yearly prices are the labor costs plus the amortized fixed cost. The inflation adjustments are 0.975, 0.966, and 0.952, respectively. The full cost of the project is the following at the time project decision making process, year 0:

$$Full\ Cost = \frac{0.975\ \times 8000}{(1+.05)^1} + \frac{0.966\ \times 8200}{(1+.05)^2} + \frac{0.952\ \times 8300}{(1+.05)^3}$$
$$= 7433 + 7186 + 6828 = 21{,}447$$

4.2.2 Non-time-based adjustments

There may be ingredients used directly where prices do not reflect either actual costs or opportunity costs. Examples include imported goods that are priced through distorted exchange rates, wages that are under- or over-valued, donated goods that are not priced, transfer payments, and government fees averted for NGOs. Because most CEAs are conducted using the societal perspective, adjustments are needed to capture opportunity costs and report economic costs. In this section, we provide examples of common non-time-based adjustments (price distortions) and illustrative examples.

Pricing labor value: It is essential to appropriately price labor value, especially in countries where unemployment of unskilled workers is common. This can significantly impact the opportunity cost of labor and influence the economic costs. The adjustment for pricing labor value involves considering the opportunity cost of labor, distinguishing between unskilled and semi-skilled workers, and evaluating the wages paid during the intervention.

For instance, in settings where unskilled workers make up a significant portion of the labor force and face high unemployment rates, the opportunity cost of labor may be close to zero. In such cases, wages in these programs might be perceived as transfer payments that are similar to cash transfers. On the other hand, some interventions may employ semi-skilled workers who possess alternative employment options, and thus their wages could impose social costs.

Example 6: Pricing labor value

To address the complexities surrounding labor value pricing, one solution is to cost labor at the intervention's marginal productivity, which could be approximated by the wages paid to the workers.

For example, there is a project that provides guaranteed employment to laborers for 100 days per year. Since there is significant unemployment in this setting, the opportunity costs of labor may be very low. Yet, some notion of efficiency wage may apply. It is recommended that unskilled labor be priced at the marginal productivity, which is likely to be the wages actually paid in the project.

Another example is volunteer labor. Volunteer labor should be priced at actual wages that would be paid if the labor was not voluntarily provided. If a highly qualified person volunteers, the price can be set at the wage of the commensurably skilled person within the context.

However, this approach should be used cautiously. Skilled workers, for example, might be underpaid relative to their productivity, thus skewing the analysis. Hence, researchers should perform sensitivity analysis to explore different scenarios, adjusting labor value from the actual wages paid during the intervention. Conducting sensitivity analysis allows for a more robust evaluation of the intervention's cost-effectiveness and provides insights into the potential impact of varying labor value assumptions (see Section 3.3).

Currency value and import duties: In many low- or middle-income countries, domestic currencies are undervalued. Undervalued or overvalued currencies impose opportunity costs. Organizations such as the International Monetary Fund (IMF) suggest adjustment values. Further, the price paid within the country reflects import duties imposed on the international border price of the good. Prices of imported goods should be adjusted to the international price and adjusted by other shadow costs.

Community inputs and volunteer: Direct inputs to produce the outcome may include community labor and voluntary labor from different sources, including highly skilled labor that may have been funded by sources outside the domestic public sphere. If the implementation process requires community input, it should be priced by the skill level of the activities. For many interventions, it may be required that the recipients are consulted. It may be viewed that community participation is an element of citizenship, therefore, it should be priced at zero. If highly skilled labor was provided for ‘free’ and was essential in delivering the intervention outcome, it should be most likely priced at the labor cost of a similarly skilled-level domestic individual.

Fees to government: Fees to the government, although they may be considered a transfer, should be costed, as governments may incur costs (for example, in issuing licenses). However, in most cases, the fee may simply be a tax payment without any effort. These can be treated as transfer payments.

Exchange rate: In some countries, there are parallel exchange rates – one that is official, and another where informal exchanges take place at a different rate. In such

countries, the official exchange rates should be adjusted. If the analyst is aware of parallel exchange rates, consulting documents from the World Bank and the International Monetary Fund at the country level may be warranted (Rompaey, Metreau and Kouame, 2021).

Spillovers: Intervention may generate positive or negative externalities, which may have implications for the costs (or effect sizes) estimated by the analyses. If the intervention generates spillovers, the externality should be documented and implications for cost estimates should be documented and accounted for.

4.2.3 Cost sharing

Interventions are situated within an organization. An organization will likely implement interventions simultaneously, presenting the opportunity for resources to be shared. An organization may also view integrated intervention implementation as an efficient way to implement interventions. For example, labor can be shared for interventions where labor time is devoted across interventions (Lopetegui *et al.*, 2014); equipment can be shared in terms of time; and a classroom can serve as a learning environment or meeting space. Overhead costs such as buildings, human resources, and equipment should be divided by the proportion of the organization's budget that the intervention under evaluation utilizes. If the total costs of these items and the share of employed units is known, apportioning by proportional use is recommended.

Cost sharing is one of the more difficult aspects of costing, and past guidelines have suggested simple methods of dividing costs across interventions. Annexure 2 explores several relatively more advanced approaches.

4.3 Uncertainty and sensitivity

How should we manage uncertainty in cost data? Should we consider alternative assumptions for adjustment parameters? At the time of design, project planners should anticipate uncertainties during project implementation, for instance, that certain portions of project sites will incur different costs, and that costs may change during implementation.

Often, the total costs in a costing analysis will contain information that is based on assumptions, for which alternative assumptions are feasible or data are obtained from samples. The cost analyst should report the base parameters for assumptions, including point estimates for costs obtained using base assumptions and mean values.

Cost analysts should also conduct sensitivity analyses around discount rates and exchange rates (McEwan, 2011). Occasionally, depreciation values as well as the longevity of capital goods should be varied. If researchers are concerned that parameters are likely to be wrong, then sensitivity analysis should be conducted. Usually only two or three parameters will be adjusted. If there are values that were obtained from a sample, then confidence intervals should be noted. Uncertainty and sensitivity analyses provide ranges for the true costs of a project.

5. Reporting costs and impact

5.1 Cost methods

As with any impact study, cost studies should contain a methodology section that documents cost perspective, data sources, adjustments, and assumptions. Although many of the parameters used in costing, such as inflation rates, interest rates, exchange rates, and longevity of capital goods, may follow internationally recognized guidelines (such as the World Bank or the International Monetary Fund), the presentation of assumptions clarifies how the results were derived. Methods should be clearly specified so that the results can be replicated and generalizable to other contexts.

5.2 Cost summaries

Before reporting costs with impact, cost summaries should be made. In Table 3, cost summaries are presented to demonstrate how cost data can be summarized. Cost summaries for the STARS Case Study are presented in Box 7.

Table 3: Types of costs summaries

Total cost	Total costs are the value of all costs at the relevant unit. Costs can be reported for the entire intervention, or for some small units of the intervention that are representative of the larger implementation (see Section 1.2.1, 'Unit of coverage') or at a regional level. The sampling strategy and sample size should be noted. An example would be the total costs to the state for a vaccine program implemented at the district level.
Average cost	Costs may vary, so it can be informative to report the average costs per unit; for example, the average cost of a vaccine program per district, or the average cost per 1,000 children vaccinated.
Costs by cost center	The activity-based costing approach emphasizes the importance of organizing costs by cost center. Total costs for the major components of the intervention should be presented and can be a critical planning resource for future programming, for example, the costs per vaccine camp set-up.
Costs by resource usage	For program planning purposes, it may be helpful to report costs by the type of cost to understand the cost drivers, for example, reporting total and average labor costs, capital costs, variable costs, etc.

Box 7: Presenting cost summaries for the STARS Cost Analysis

The STARS program was successfully implemented, and the evaluation team has finalized end line data collection. The team is ready to compile and analyze cost and effectiveness data to answer the research questions developed in Box 2.

To answer our first research question, we report our cost estimates.

RQ1: What is the total cost of the STARS program?

	D0	D1	D2
Total cost	$600,000.00	$700,000.00	$1,000,000.00
Number of participants	10,000	8,000	6,000
Average cost per participant	$60	$87.50	$166.66

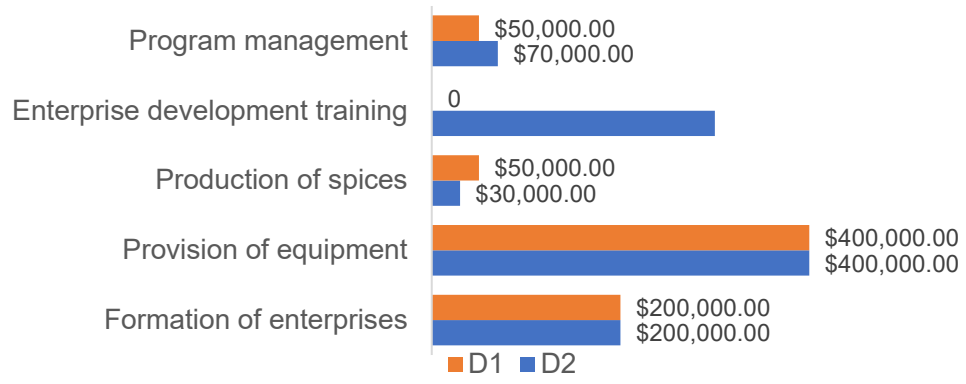
STARS D2 (forming enterprises and conducting EDL training) is much more expensive than both the comparator ('light touch' training) and D1 (forming enterprises) due to the EDL training component of the program.

Our Ministry partners also expressed interest in disaggregating STARS costs by region. Because the costs we captured were disaggregated by region in our CCT (Box 4), we are able to report total costs and average costs per participant by treatment arm and region. We show an illustrative example of regional disaggregation for D1 in the table below.

	Region 1	Region 2	Total
Cost (D1)	$300,0000	$400,000	$700,000.00
Number of participants	5,000	3,000	8,000
Average cost per participant	$60	$133.33	$87.50

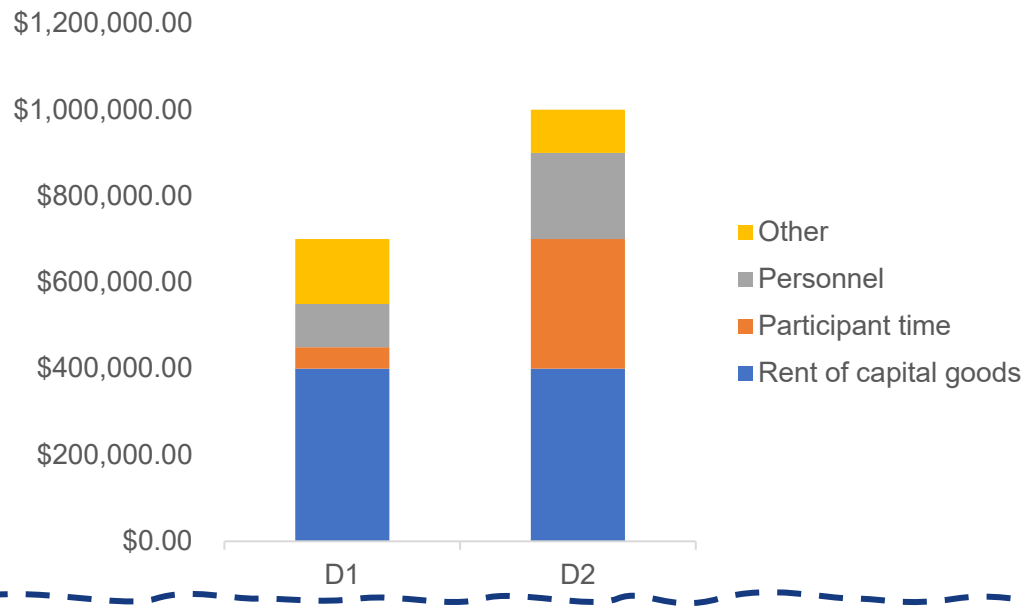
Costs can also be presented by cost center. For example, we observe that STARS D1 and D2 made identical investments in the 'formation of enterprises' and 'provision of spice processing equipment.' Based on differences in the 'production of spices' costs, conducting the EDL training (STARS D2) may have contributed to cost savings.

Figure 5: STARS Cost by Cost Center



Another suggestion for presenting costs is by resource type. For example, one of the cost-drivers of the STARS D2 program, compared to D1, is the cost of participant time. D2 was time-intensive for participants and involved traveling upwards of 30 kilometers to training in centralized locations. We observe that much of the additional expense of D2 is driven by participant time.

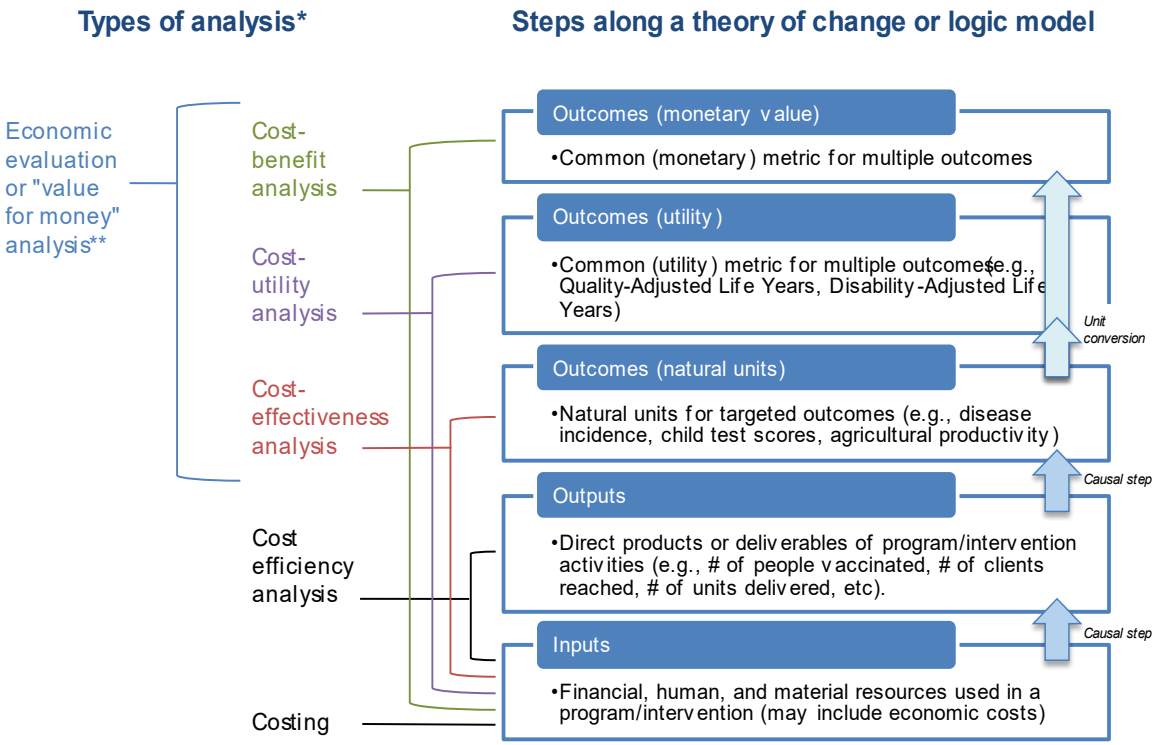
Figure 6: STARS Costs by Resource Type



5.3 Understanding cost in relation to impact

Costs and impact can be compared through several methodologies – they imply different interpretations of how impacts stand in relation to costs and answer slightly different research questions (Figure 7). For example, while cost efficiency analysis estimates the ratio of project costs to outputs, cost-effectiveness compares the cost per unit of change in an outcome of interest (Drummond *et al.*, 2015). Although these methods propose various strategies for systematically assessing impact, a commonality among them is that they all rely on efficient and accurate collection of cost data. Point estimates are the starting point of cost-effectiveness analyses and, similar to outcome measures, should be reported with confidence intervals.

Figure 7: Key analytical components of analyses using cost evidence along the theory of change



*all listed analyses may either beex ante (modeling/predicting a future intervention) or ex post (characterizing a completed intervention)
**outcomes for these analyses should be causally attributed to the intervention based on construct of a valid counterfactual

Source: Glandon et al., 2023

5.4 Bringing cost and effectiveness together

Cost and effectiveness measures must be brought together using a consistent unit of measurement, such as cost per district per impact, or cost per person per effect. Effect sizes in many impact analyses are derived from regressions, and should be interpreted to fit the cost analysis. For some situations, multiple measures may be appropriate. For instance, an agricultural intervention affects individual farmers, but may also have village level impact.

This handbook recommends that researchers agree on units of measurements that are consistent across both the cost and impact studies. Cost and impact measures should come from the same sites and the same time period. However, there may also be a need for stratified sampling for both cost and effectiveness. It is also possible that cost data is more homogenous, thus requiring fewer data collection sites, or that purposive sampling is required to ensure heterogeneity in costs captured. Ideally, these considerations should be accounted for during the 'Define the Project' phase of the cost analysis (Section 1.2).

5.4.1 The comparator

As discussed in Sections 1.2 and 3.1, costs and effects must be reflected as relative to other activities, for example, the differences in outcomes observed in treatment and control groups. In cost-effectiveness, costs should be presented relative to a comparator. The comparator is usually the status quo or an alternative intervention or project. Cost

data for the comparator should be collected or estimated, but in some cases, the intervention does not replace a similar project, and it is possible that there are no documented costs for the comparator. For example, if the project is targeting a new or historically under-researched intervention, cost data on comparators may be difficult to find, especially in low- and middle-income country contexts (Tan-Torres Edejer *et al.*, 2003). As a result, many CEA studies do not provide appropriate comparators (Elliott *et al.*, 2014).

If it is feasible to assume that the 'status quo' generates zero costs, or if nothing is known about the cost of the comparator, the team will have to consider how to estimate the change in impact. Ideally, baseline outcome measures can be used to estimate effect sizes, and the CE ratio can be calculated by dividing the change in an outcome by the average cost of the intervention. If there is no baseline outcome metric, the team will need to determine whether a CE ratio (the total cost divided by the effect size) is useful to the implementing partner, perhaps by using a threshold (see Section 4.4.3).

5.4.2 The incremental cost-effectiveness ratio

How should one understand the cost-effectiveness ratio, the main result of the cost analysis, after costs and impacts are estimated and adjusted? ICER is stated as the following:

$$ICER_{New} = \frac{Cost_{New} - Cost_{Old}}{Effect_{New} - Effect_{Old}}$$

The decision process can be understood better if one views this through the cost-effectiveness plane (Drummond *et al.*, 2015), as shown below (Figure 7). To interpret the quadrants (hereafter, Q), assume that an organization spends $100,000 on a project to affect 1,000 people, and that the impact in terms of well-being is well known. A new project is proposed that will cost $95,000 and affect 1,100 people. The incremental cost-effectiveness ratio is the following:

$$\frac{\$95000 - \$100000}{1100 - 1000} = \frac{-5000}{100} = -\$50$$

An ICER of -$50 indicates that the new program is cheaper than the older program. Thus, replacement is warranted – in this case the new program helps more people at a lower cost. That is, for $5,000 less in costs, it was observed that an additional 100 people were being helped. This is the Q3 where there are positive gains for lower costs.

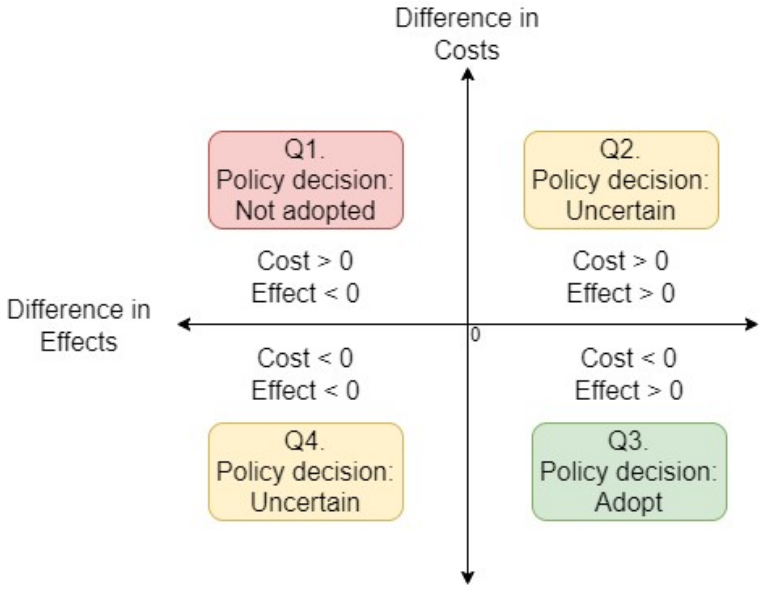
Supposing the new costs of the new project were $115,000. In this case, the ICER is $150, that is, for an additional $150 per person, the new project benefits more people. Another way to interpret the result is that for an additional $15,000 spent, 100 more people gain benefits. This ICER is placed in Q2, and decision-makers must decide as to whether this project is worth the effort.

If costs are higher with smaller benefits, the ICER is placed in Q1, and the decision-maker will dismiss the new intervention from continuing or being replicated. Let us suppose that for costs of $95,000 only 950 people were affected, and the ICER is placed in Q4. Then the program saves $5,000, but fails to affect 50 people than it did before. The decision-maker may consider whether these savings that require a reduction in welfare is acceptable.

Some projects are considered ‘add-on,’ or new programs that incur additional costs and are expected to yield positive effects on well-being. Suppose an intervention costs $100,000 and affects 500 children to gain one more year of school, that is, a year of school is gained by spending an additional $200 per child. The policymakers must ask whether such an intervention is worth funding. There is an ICER here, interpreted as the difference from 0 prior effects and costs. This is a cost-effectiveness ratio which has the same interpretation as average costs.

Impact evaluation studies often conduct three-arm studies, with one of the arms being the status quo and the other two affecting the same type of welfare. The program with the higher costs must also produce higher benefits to be considered as an alternative. One assesses the two alternatives through the method shown in Q2.

Figure 8: The Cost-effectiveness plane



Source: Drummond et al., 2015

It is possible for interventions to have produced statistically significant improvements and still be seen as inefficient. For example, modest gains that have been seen from the “Millennium Villages” interventions, implemented in some parts of rural sub-Sahara Africa, have yet to be seen to have warranted the costs of the interventions (Masset, Hombrados and Acharya, 2020). One suggestion for decision-making on a project whose ICER falls within Q2 and Q4 is to ask whether society would accept a particular amount of costs (or savings) in gaining (or losing) a unit of outcome. Box 8 provides an example of ICER from the STARS Case Studies.[10]

[10] Calculations of ICER are easier to interpret if there are current interventions in place and the comparator is well-defined. However, in some cases, even if there are current interventions in place, it may be hard to discern what the present (comparator) costs are. Suppose a new program is introduced to strengthen the resolve of the female students to finish high school. There may already be some effort towards that end. However, in such a situation, the costs of the present effort would most likely not be known. Thus, assessing the CEA of the new project would yield a cost-effectiveness ratio that reports the ratio of total costs of the program over the total increased number of female youth matriculating from high school. ICER is the relevant measure for at least one of the projects, if two interventions are being compared with the status quo.

Box 8: Calculating the ICER, STARS Case Study

The evaluation team worked with the ministry to develop a tool to estimate women's economic empowerment in the treatment and control groups using a point score test. Women who score above a certain threshold level of the scale are, for the purpose of this example, considered 'empowered'. The impact evaluation reports the change in the proportion of women in our sample who are 'empowered'. This change can be causally attributed to the STARS program.

- RQ 1: What is the total cost of the STARS program?

To answer the first research question, cost estimates for D0 (the 'light touch' training), D1 (forming enterprises), and D2 (forming enterprises and conducting EDL training) are reported. As discussed in Box 7, we find that D2 is much more expensive than D1.

Program	D0	D1	D2
Cost	$600,000.00	$700,000.00	$1,000,000.00
Number of participants	10,000	8,000	6,000
Average cost per participant	$60	$87.50	$166.66

- RQ 2: What is the effect of the STARS program on women's economic empowerment?

The impact evaluation collected data on women's economic empowerment (WEE), reported as the number of women considered 'empowered,' measured at both baseline and endline, for each intervention group (D1, D2) and the comparator (D0). The primary outcome of interest was the change in WEE that was attributable to each intervention (D1 and D2), relative to the comparator. For instructional purposes, we present the attributable change as the difference in the number of women 'empowered' at endline relative to baseline in each of the intervention arms (D1, D2) relative to the comparison arm (D0). The evaluation reports that:

Program	D0	D1	D2
Proportion of women empowered at baseline [95% CI]	40% [38.8%, 41.9%] (4,000/10,000)	40% [37.5%, 42.5%] (3,200/8,000)	40% [36.9%, 43.1%] (2,400/6,000)
Proportion of women empowered at endline [95% CI]	50% [49.1%, 50.9%] (5,000/10,000)	60% [58.2%, 61.6%] (4,800/8,000)	70% [68.8%, 71.9] (4,200/6000)
Change in proportion (and number) of women empowered (Endline – Baseline) [95% CI]	10% (1,000) [9%,11%]	20% (1,600) [17.2%, 21.2%]	30% (1,800) [28.2%, 31.5%]
Change in WEE attributable to STARS (D1, D2) relative to D0* [95% CI]	--	600 [536, 698]	800 [692, 890]

- RQ 3: Which STARS treatment arm is relatively more cost effective?

To compare cost-effectiveness, the relative cost effectiveness of each treatment arm must be calculated. The following formulas are used to calculate ICERs for D1 and D2:

ICER D1*: Comparing CEA of D1 and D0

$$\frac{Cost\ of\ D1\ \ (\$700{,}000) - Cost\ of\ D0\ (\$600{,}000)}{Effect\ of\ D1\ (1{,}600)\ \ - Effect\ of\ D0\ (1{,}000)}$$

= Investing an additional $100,000 on D1 is expected to empower 600 additional participants

= D1 costs $167 per additional woman empowered.

ICER D2*: Comparing CEA of D2 and D0

$$\frac{Cost\ of\ D2\ (\$1{,}000{,}000) - Cost\ of\ D0\ (\$600{,}000)}{Effect\ of\ D2\ (1{,}800) - Effect\ of\ D0\ (1{,}000)}$$

= Investing an additional $400,000 on D2 is expected to empower 800 additional participants

= D2 costs $500 per additional woman empowered.

Relative to D0, both D1 and D2 are more expensive and more effective (Costs > 0; Effects > 0) and both ICERs are mapped to Q2 of the ICER plane (Figure 7). However, when compared to D0, we observe that D1 is more cost effective than D2.

Because STARS is both more effective and more expensive than 'business as usual,' additional analyses may help policymakers to determine whether the initiative is worthwhile. Box 9 provides a few examples of possible interpretations of the ICER. Though beyond the scope of this guide, it may be useful to consult resources on conducting willingness-to-pay analysis (Gabor and Granger, 1979), budget impact analysis (Sullivan *et al.*, 2014) and others.

Below, we summarize findings from the cost-effectiveness analysis.

	D0	D1	D2
Total cost	$600,000.00	$700,000.00	$ 1,000,000.00
Number of participants	10,000	8,000	6,000
Change in proportion (and number) of women empowered (Endline - Baseline) [95% CI]	10% (1,000) [9%,11%]	20% (1,600) [17.2%, 21.2%]	30% (1,800) [28.2%, 31.5%]
Marginal cost per woman empowered, relative to D0	--	$167	$500

*For instructional purposes we report nominal difference in observed change between treatment arms and control groups. In practice the attributable effect will be calculated based on the econometric specifications.

Box 9: Additional interpretation of findings from the STARS Cost Analysis

What should we recommend to our government partners? In this box, a few possibilities are presented for applying findings from the STARS CEA results. Note that this analysis has been simplified for learning purposes and is not exhaustive. Interpretation will vary by stakeholder interests and scope of cost study defined in the 'Define the Intervention' phase.

1. Heterogeneity in costs

As discussed in Section 4.2 (Box 7), it is observed that participant time and personnel costs were cost drivers for D2. D2 was much more expensive than D1 because it required hundreds of hours of respondent participation, resulting in lost wages. D2 also cost more per participant because fewer participants were mobilized.

It is also observed that there is heterogeneity in costs by region. Region 2 was harder to reach than Region 1, and the implementation team incurred higher costs and reached fewer people. These differences should be accounted for in subsequent analyses, including scale-up modeling.

2. Comparing costs and effects

Both iterations of the STARS intervention significantly increased the number of participants empowered relative to the D0 'light touch' intervention. After dividing the total cost of each intervention by the number of participants empowered, the average cost per empowered participant is greater for D0 than D1 or D2.

When presenting this information, we may encourage the ministry to consider replacing the status-up (D0) program with one of the STARS interventions, but caveats should also be presented. For example, D2 reached fewer participants than D1 or D0, and additional analyses will be required to determine how the effectiveness of the program will change if the program is scaled. Average costs should be presented using sensitivity analyses (see Section 3.3).

3. Interpreting the ICER: D1 vs D2

D1 is more cost-effective than D2 as the intervention costs less per additional woman empowered, but caveats should again be presented. Both STARS interventions are more effective and more expensive than the status quo, and per the ICER plane (Figure 7), the decision to adopt is 'uncertain' (see Section 4.3.2). The cost team may need to work with the decision-maker to determine if this additional investment is valued by the ministry stakeholders, possibly by conducting additional scale-up or sensitivity analyses to explore how estimated costs and effects will change over time.

For example:

- When defining the scope of our cost analysis, the ministry stated their commitment to increasing women's economic empowerment such that an additional 40% of women are considered 'empowered' by 2025 (Box 1).
- In this scenario, the STARS D2 alternative may be the stronger choice because in two years, it is expected to empower 30% of participants, while D1 only reaches 20%.
- This target could be used to help the ministry decision-makers determine whether they value the gain in women's economic empowerment more than the additional investment ($500 per woman empowered).

5.4.3 Thresholds

Realistically, many projects do not automatically warrant an 'adopt' decision as suggested by Q3 of the ICER plane (Figure 7). It is also possible that there is no 'status-quo' intervention or policy in place to serve as a comparator to the intervention under evaluation. In these cases, the research team may consider alternative approaches to support the interpretation of CEA, such as by using cost-effectiveness thresholds.

Any decision regarding whether to adopt or reject an intervention that is both more effective and expensive involves an implicit consideration of a threshold (Drummond *et al.*, 2015). Cost-effectiveness thresholds are commonly used in health economics to compare an ICER (the output of the CEA) to a normative CE threshold to identify whether the new intervention is a good value-for-money (Thokala *et al.*, 2018). However, these threshold values (which generally assign value to quality of life) are highly contested (that is, reasonable and informed people can disagree as to what the values should be), and various values have been proposed in different contexts (Hirth *et al.*, 2000; Hyewon and Levine, 2012). Practically, it may be challenging to expect policymakers to identify a threshold to justify adaptation of a policy. If used, the approach to determining an appropriate threshold should be evidence-informed (Culyer *et al.*, 2007). The resources cited in this section should be consulted for more information on using thresholds in CEA.

A conceptual explanation of thresholds

A cost-effectiveness threshold is the maximum cost a society will be willing to pay for acquiring an additional amount of an outcome. It can also be conceptualized as the marginal benefit that society gains when producing an outcome.

It is possible that marginal benefits decrease as more of an outcome is produced, just as one sees in a demand curve. More of an outcome should be produced so long as the extra cost of producing this outcome is lower than this threshold. If possible, the amount produced should be at a level where the extra cost of production at that level equals the threshold amount.

5.5 Scaling-up

Cost-effectiveness analyses are often initially conducted at a smaller scale to evaluate the impact of interventions. During this phase, it is crucial to perform a thorough process evaluation to determine the replicability of the program. If the cost-effectiveness ratio proves favorable, decision-makers might consider expanding the program, replicating it in new locations, or maintaining the current level of intervention.

Within a sufficient timeframe, economists usually assume that any production process can be repeated or that learning-by-doing reduces costs. However, there may be some constraints, such as hard-to-reach participants, skilled labor or environmental and resource constraints that prevent economics of scale. CEA can help decision-makers anticipate these constraints through accurate and disaggregated costing that can be used in modeling or forecasting (these methods are not discussed in this handbook).

If considering scale costs in cost analysis, one of the assumptions for the validity of conducting CEA is that the intervention does not affect the prices of intervention inputs. Yet, most interventions are part of a budget. As an intervention is scaled, it may affect interventions that are more efficient than the intervention being evaluated or require additional specialized skills. In both cases, labor costs may rise as scaling up takes place and will need to be adjusted.

5.6 Multiple outcomes

In recent years development strategies have increasingly emphasized multi-sectoral collaboration (Glandon *et al.*, 2019). It has been recognized that interventions implemented simultaneously can induce a 'take-off' or 'big push' for change in well-being. On smaller scales, the literature suggests that integrating multi-sectoral interventions regionally can create synergy and efficiencies that reduce costs and improve outcomes (Banerjee *et al.*, 2015; Gelman *et al.*, 2022). It is also possible that an intervention yields multiple outcomes. For instance, education interventions contribute to changes in outcomes related to student health, income, and civic participation. Integrated or complex interventions involving multiple sectors should yield cost savings or produce multiple outputs.

6. Conclusion

In a world of resource scarcity, comparing the impacts of policies or projects without considering costs "is like one hand clapping" (Gaarder and Linn, 2023). Despite its potential for informing policy and program design, cost evidence is often not incorporated into impact evaluations (Brown & Tanner 2019).

This handbook focuses on CEA conducted as **ex-post analysis using prospective cost capture**, in which cost data is collected as the project progresses. This approach is essential as actual costs can vary significantly from budgeted, expected or secondary cost data sources.

The CEA framework developed in this handbook comprises four empirical tasks: 1) defining the intervention and scope of study; 2) designing a cost collection template to capture costs incurred throughout the intervention lifecycle, 3) collecting and adjusting cost data as the intervention is implemented; and 4) reporting costs and impact. The key steps of each task are summarized in Annex 1.

The hope is that these resources will contribute to more transparent cost reporting, facilitate better resource allocation, and strengthen the credibility of policy decisions. By standardizing cost analysis methods, policymakers can more reliably compare the cost-effectiveness of different interventions, leading to more informed decisions that optimize impact.

Appendix A: Costing Checklist

Key Tasks of CEA in IE		Project Lifecycle			
		D	I	R	
1 Define the project and study scope					**Project resources needed:**
	Develop research questions	X			• Impact evaluation design • Pre-analysis plan • Theory of Change
	Identify cost metrics to be generated by the analysis	X			
	Identify evaluation design				
	Develop project Theory of Change	X			
	Define comparator	X			
	Define perspective	X			
	Define unit of coverage	X			
	Define time horizon				
	Define scope of costs to be included in the analysis	X			
2 Develop cost collection template					**Project resources needed:**
	List cost centers	X			• Implementation partner to co-develop cost centers and ingredients
	List cost ingredients	X			
	Specify cost data sources, data collection timelines	X			
	Identify cost adjustments	X			
3 Collect and adjust costs					**Project resources needed:**
	Incorporate cost collection in IE data collection	X			• CEA Expert to support cost adjustments
	Conduct time adjustments		X		
	Conduct non-time adjustments		X		
	Allocate shared costs		X		
	Account for uncertainty		X		
	Conduct sensitivity analyses			X	
4 Report costs and impact					**Project resources needed:**
	Report cost summaries			X	• Effect sizes estimated by IE
	Incorporate impact; calculate CER or ICER			X	
	Scale-up			X	
	Multiple outcomes			X	

D - Design; I - Implementation; R - Reporting

Appendix B: Cost Sharing

For many projects some inputs will be shared with other projects. It is also possible that some outputs that would have been produced separately would be produced jointly. Production processes can impose externalities, both negative and positive. This section does not examine externalities, only uses of joint inputs are examined. Although any sharing of inputs can be thought of as a joint production, this will most likely involve substantial usage of shared inputs.

Inputs when not completely used up in a production process can be thought of as a club good or a public good – the good can be used for some other purpose without being non-rivalrous. An input can be used for some other purposes when there is excess capacity. Dividing the cost of jointly used inputs can be a complicated topic (Acharya *et al*. 2022). A brief introduction is made here.

Dividing by use intensity: One can measure how much of an input is used toward different purposes. Some analysts (Drummond *et al.*, 2015) consider as a natural example of cost sharing overhead costs in a hospital that can be divided across the many health interventions that are undertaken within the hospital. For example, if the total housekeeping cost in a hospital is known, it can be divided by the floor space that each cluster in the hospital occupies.

Some have suggested taking the total cost of input and apportionment of the costs equally to each usage that it has (Shepherd, Zeng, and Nguen 2015).

For many sharing of labor inputs time and motion studies can be useful. Studies take place to enumerate detailed data on the duration and movement required to accomplish a specific task. The portions used are assigned to a project or activity.

One should note that when input or a production process is used for multiple outputs there is joint production that must induce cost savings in comparison to production processes where outputs are stand-alone products.

References

Banerjee, A. *et al.* (2015) ‘A multifaceted program causes lasting progress for the very poor: Evidence from six countries’, *Science*, 348(6236), pp. 1260799–1260799. Available at: https://doi.org/10.1126/science.1260799.

Brown, E.D. and Tanner, J.C. (2019) *Integrating Value for Money and Impact Evaluations: Issues, Institutions and Opportunities*. World Bank Working Paper, 9041. Available at: https://documents1.worldbank.org/curated/en/862091571145787913/pdf/Integrating-Value-for-Money-and-Impact-Evaluations-Issues-Institutions-and-Opportunities.pdf.

Brown, L. (2022) *Mercury Project First Virtual Convening Recording*. Available at: https://drive.google.com/drive/folders/1jJGQlBrbWJVfDKeDEL0ey6m46L_4q1V- (Accessed: 25 July 2023).

Byford, S. and Raftery, J. (1998) ‘Perspectives in economic evaluation’, *BMJ : British Medical Journal*, 316(7143), pp. 1529–1530.

Cost-Benefit Discounting (2014). Available at: https://www.youtube.com/watch?v=Mol1yT7tczY (Accessed: 25 July 2023).

Culyer, A. *et al.* (2007) ‘Searching for a threshold, not setting one: the role of the National Institute for Health and Clinical Excellence’, *Journal of Health Services Research & Policy*, 12(1), pp. 56–58. Available at: https://doi.org/10.1258/135581907779497567.

Dhaliwal, I. *et al.* (2012) ‘Comparative Cost-Effectiveness Analysis to Inform Policy in Developing Countries: A General Framework with Applications for Education’. MIT, J-PAL. Available at: https://www.povertyactionlab.org/sites/default/files/research-paper/CEA%20in%20Education%202013.01.29.pdf.

Drummond, M. *et al.* (2015) *Methods for the economic evaluation of health care programmes*. 4th Edition. Oxford, United Kingdom ; New York, NY, USA: Oxford University Press (Oxford medical publications).

Elliott, R.A. *et al.* (2014) ‘A Review of the Methodological Challenges in Assessing the Cost Effectiveness of Pharmacist Interventions’, *PharmacoEconomics*, 32(12), pp. 1185–1199. Available at: https://doi.org/10.1007/s40273-014-0197-z.

Gaarder, M. and Linn, J.F. (2023) ‘Upping the game: adding costs to impacts an introduction to the special issue on costing in the field of development effectiveness’, *Journal of Development Effectiveness*, 15(1), pp. 1–4. Available at: https://doi.org/10.1080/19439342.2023.2171062.

Gabor, A. and Granger, C.W.J. (1979) ‘The Attitude of the Consumer to Prices’, *Management Decision*, 17(8), pp. 619–634. Available at: https://doi.org/10.1108/eb001215.

Gelman, A. *et al.* (2022) ‘Reconciling Evaluations of the Millennium Villages Project’, *Statistics and Public Policy*, 9(1), pp. 1–7. Available at: https://doi.org/10.1080/2330443X.2021.2019152.

Glandon, D. *et al.* (2019) ‘Methodological gaps and opportunities for studying multisectoral collaboration for health in low- and middle-income countries’, *Health Policy and Planning*, 34(Supple-2), pp. ii7–ii17. Available at: https://doi.org/10.1093/heapol/czz116.

Glandon, D. *et al.* (2023) ‘The State of Cost-Effectiveness Guidance: Ten Best Resources for CEA in Impact Evaluations’, *Journal of Development Effectiveness*, 15(1), pp. 5–16. Available at: https://doi.org/10.1080/19439342.2022.2034916.

Hirth, R.A. *et al.* (2000) ‘Willingness to Pay for a Quality-adjusted Life Year: In Search of a Standard’, *Medical Decision Making*, 20(3), pp. 332–342. Available at: https://doi.org/10.1177/0272989X0002000310.

Hyewon, H.L. and Levine, M. (2012) ‘Determining the threshold for acceptability of an ICER when natural health units are used’, *Journal of Population Therapeutics and Clinical Pharmacology*, 19(2). Available at: http://www.jptcp.com/index.php/jptcp/article/view/429 (Accessed: 7 February 2024).

IRC (2014) *Protection Cost Effectiveness Brief - Girls Empowered by Microenterprise*. Available at: https://www.rescue.org/sites/default/files/2022-10/Livelihoods%20Cost-Effectiveness%20Brief-%20GEM_0.pdf (Accessed: 12 February 2024).

Levin, H.M. and McEwan, P.J. (2001) *Cost-effectiveness analysis: methods and applications*. 2nd ed. Thousand Oaks, Calif: Sage Publications.

Levin, H.M. and McEwan, P.J. (2002) *Cost-effectiveness and educational policy*. Eye on Education Larchmont, NJ. Available at: https://www.academia.edu/download/52568316/Cost-Effectiveness_and_Educational_Polic20170410-28103-1qzve3g.pdf (Accessed: 30 January 2024).

Lopetegui, M. *et al.* (2014) ‘Time motion studies in healthcare: What are we talking about?’, *Journal of Biomedical Informatics*, 49, pp. 292–299. Available at: https://doi.org/10.1016/j.jbi.2014.02.017.

McEwan, P. (2011) ‘Cost-Effectiveness Analysis of Education and Health Interventions in Developing Countries’. Inter-America Development Bank, SPD Working Papers 1102. Available at: https://ideas.repec.org/p/idb/spdwps/1102.html.

Muralidharan, K. and Sundararaman, V. (2011) ‘Teacher Performance Pay: Experimental Evidence from India’, *Journal of Political Economy*, 119(1), pp. 39–77. Available at: https://doi.org/10.1086/659655.

O’Mahony, J.F., Newall, A.T. and van Rosmalen, J. (2015) ‘Dealing with Time in Health Economic Evaluation: Methodological Issues and Recommendations for Practice’, *PharmacoEconomics*, 33(12), pp. 1255–1268. Available at: https://doi.org/10.1007/s40273-015-0309-4.

Rompaey, C.V., Metreau, E. and Kouame, C. (2021) ‘Accounting for parallel exchange rates in World Bank classifications of countries by income level’, *The World Bank*, 17 August. Available at: https://datatopics.worldbank.org/world-development-indicators/stories/accounting-for-parallel-exchange-rates-in-world-bank-classifications-of-countries-by-income-level.html (Accessed: 31 January 2024).

Sullivan, S.D. *et al.* (2014) ‘Budget impact analysis—principles of good practice: report of the ISPOR 2012 Budget Impact Analysis Good Practice II Task Force’, *Value in health*, 17(1), pp. 5–14.

Tan-Torres Edejer, T.T.-T. *et al.* (eds) (2003) *Making choices in health: WHO guide to cost-effectiveness analysis*. Geneva: World Health Organization.

Thokala, P. *et al.* (2018) ‘Cost-Effectiveness Thresholds: the Past, the Present and the Future’, *PharmacoEconomics*, 36(5), pp. 509–522. Available at: https://doi.org/10.1007/s40273-017-0606-1.

USAID Global Health Supply Chain Program (2021) *Activity-based Costing Handbook*. Available at: https://www.ghsupplychain.org/activity-based-costing-handbook (Accessed: 23 April 2024).

USAID, O. of E. (2018) ‘Cost Reporting Guidance for USAIDFunded Education Activities’. United States Agency of International Development. Available at: https://www.edu-links.org/sites/default/files/media/file/USAID%20Cost%20Reporting%20Guidance%202018%20FINAL%20%281%29.pdf.

White, H. (2013) ‘Using Causal Change to Make Sense of Numbers’. 3iE. Available at: https://www.3ieimpact.org/blogs/using-causal-chain-make-sense-numbers.

Working Paper Series

International Initiative for Impact Evaluation
Unit no. 306, 3rd Floor, Rectangle-1
D-4, Saket District Centre
New Delhi – 110017
India

3ie@3ieimpact.org
Tel: +91 11 4989 4444

www.3ieimpact.org