Guidelines on Estimating the Costs and Cost-Effectiveness of Women’s Groups in International Development

DECEMBER 2019

Garima Siwach, PhD | Thomas de Hoop, PhD | Giulia Ferrari, PhD | Yulia Belyakova

Source: Gates Archive / Jiro Ose
Acknowledgements

We would like to thank co-members of the Evidence Consortium on Women’s Groups, specifically Leigh Anderson, Gary Darmstadt, and Sapna Desai, for their support throughout this project. We would especially like to thank Helen Connolly at the American Institutes for Research for her valuable comments and feedback on previous versions of these guidelines and tools. Our thanks go to the Gender Equality team at the Bill & Melinda Gates Foundation for their financial and technical support, and specifically to Shubha Jayaram, Katherine Hay, Sybil Chidiac, and Yamini Atmavilas for their continuous support and insights during the writing of these guidelines. Finally, our acknowledgements would be incomplete without mentioning our team of very able editors of the American Institutes for Research, including Amy deBoinville, Emma Ruckley, Sharon Wallace, and Jane Garwood.

The Evidence Consortium on Women’s Groups is funded by a grant from the Bill & Melinda Gates Foundation.
Background

There is a wide body of evidence examining the effectiveness of different types of women’s groups, particularly economic self-help groups (SHGs), Village Savings and Loans Associations (VSLAs), and Participatory Learning and Action (PLA) groups. Most commonly, research examines the impact of women’s groups on access to savings and credit; income, asset ownership, and household consumption; women’s economic empowerment and mobility; political empowerment; and health and well-being.

With the growing emphasis on implementing different types of women’s groups in South Asia and Africa, investors and donors are interested in understanding not just the overall funding requirements, but also the costs of replicating, scaling up, or sustaining a women’s group program. Information about costs and benefits can help investors and donors to compare cost-effectiveness across different types of women’s groups to guide resource allocation decisions. However, research on the costs and cost-effectiveness of women’s groups is scarce, possibly because of the lack of consistent cost data linked to different activities and outcomes.

These guidelines are intended to support researchers, stakeholders, and program implementers to collect and analyze data on program costs and estimate program cost-effectiveness, with the goal of supporting better practices in cost data collection and the use of these data to inform future investments.

Current Evidence

The most comprehensive evidence on the costs of women’s groups comes from SHGs in India. In 2007, for example, the Consultative Group to Assist the Poor (CGAP) analyzed the costs of nine SHGs in India to understand the determinants of financial sustainability. The study team collected cost data on SHGs by gathering information on different processes in service set-up and delivery, including the costs of launching, providing training, and monitoring SHGs for about 3 years. The study found wide variation in SHG costs across programs. SHG programs that were multipurpose and promoted empowerment of members had the highest costs ($443 per SHG over 3 years, or approximately $37 per person, assuming an average of 12 members per SHG). An earlier study by Tankha (2002) estimated SHG program costs of 10 leading nongovernmental organizations (NGOs) in India, which ranged from 4,500 INR to 25,000 INR per SHG ($90 to $500 in 2002 prices). Finally, a study by Harper (2002) found that the costs of developing an SHG from scratch, to the point where groups were formally linked with a bank account, ranged from 1,350 INR to 16,000 INR (approximately $27 to $320 in 2002 prices).

The evidence base suggests that both the benefits and the costs of SHGs increase when an increased number of activities are delivered. Brody et al. (2015) showed that SHGs’ impact on women’s empowerment increased significantly when they included training components—for example, training related to business skills, community development, or employment and leadership. Isern et al. (2007) found that the costs of SHGs that focused on empowerment, social change, and livelihood generation were much higher than the costs of SHGs that focused solely on lending. The lower cost SHGs also focused on literate, less-poor women who lived...
closer to bank branches, and their promotion costs were mainly limited to launching the SHG, plus initial orientation and training costs and overall program administration costs.\textsuperscript{14} It is important to exercise caution when interpreting the results of Isern et al. (2007)\textsuperscript{15} because the study intentionally focused on robustly implemented, more sustainable SHGs, which are not representative of all SHGs in India. However, Tankha (2002)\textsuperscript{16} found similar results when computing the costs of SHGs delivered by 10 leading NGOs in India. His findings showed that minimalist SHGs with a sole focus on lending and bank linkage had the lowest costs, and that the costs of SHGs that included women’s empowerment as a focus area were significantly higher.\textsuperscript{17}

Other factors that contribute to SHG costs include social cohesion, the type of promoter, and the SHG development stage. For example, Harper (2002)\textsuperscript{18} notes that the level of cohesion within a community is an important determinant of SHG costs. The functioning of group mechanisms is often tied to kinship ties, in the absence of which social mobilization costs are likely to be high.\textsuperscript{19} He also shows that SHG development accounts for the largest percentage of SHG costs, followed by support costs and then social mobilization costs.\textsuperscript{20}

A more recent study conducted by the Evidence Consortium on Women’s Groups (ECWG) suggests that SHG costs decrease significantly with scale. The ECWG analyzed expenditure statements and audit reports for the JEEVIKA program and the Jharkhand State Livelihood Promotion Society (JSLPS), which carried out state-level implementation of the National Rural Livelihoods Mission (NRLM) (a national poverty alleviation program that uses SHGs as the primary delivery vehicle) in Bihar and Jharkhand respectively.\textsuperscript{21,22} The analysis found that annual expenditure (in 2018 prices) on group formation and basic group activities in Bihar was approximately $60 per household covered when the project began in 2007. The per household cost had fallen to just over $13 by 2016, when the project reached scale. With an average group size of 11 women, these costs ranged from $143 to $660 per group in 2018 prices, or from $50 to $232 per group in 2002 prices\textsuperscript{1} (similar to the range of estimates generated by Tankha (2002)\textsuperscript{23}). These results are suggestive of large variation specifically due to scale.

In terms of the return on investment or cost-effectiveness of women’s groups, we found only three studies with evidence on cost-effectiveness for economic SHG programs: two with a focus on India,\textsuperscript{24,25} and one with a focus on Ethiopia.\textsuperscript{26} Deininger and Liu (2009) and Chandrashekhar et al. (2019) both showed that SHGs had a positive return on investment when benefits to members were calculated in terms of economic outcomes, such as increased income and consumption. Chandrashekhar et al.’s (2019)\textsuperscript{27} study also found that the Parivartan program in India resulted in the prevention of 23 neonatal deaths, at a cost of $3,825 per life-year saved.\textsuperscript{2} Unfortunately, the impact findings from Ethiopia reported by Venton et al. (2013)\textsuperscript{28} have a high risk of selection bias due to the absence of a strong experimental or quasi-experimental

\textsuperscript{1} Expenses per household covered in the JEEVIKA study (Siwach et al., 2019) were based on cumulative program outreach. Estimates shown here are based on annual program outreach to make the figures consistent with earlier literature, which estimates costs of group formation and basic activities. The authors assume that these costs will be divided among new members only.

\textsuperscript{2} The cost-effectiveness analysis focused on the adoption of maternal and newborn health behaviors, promoted by integrating health behavior change communication with SHGs.
evaluation design. This means that the reported benefits may have been overestimated (if women from higher income households were more likely to join SHGs) or underestimated (if women from lower income households were more likely to join SHGs).

A further limitation of the current literature is that all of these studies estimated costs from the funders’ perspective, measuring outcomes against total investments in the program. None of the studies accounted for the private expenditure that SHG members may have incurred to attend program activities, or the opportunity costs of their time. This could result in an underestimation of overall program costs from a societal perspective, which measures overall costs associated with a program irrespective of which party incurs those costs.

**Scope of These Guidelines**

Considering the scant evidence on the costs and cost-effectiveness of women’s groups, the ECWG—in collaboration with the Bill & Melinda Gates Foundation—prepared the following guidelines for collecting cost data on programs that focus on women’s groups. Much of these guidelines build upon two existing sources: the Abdul Latif Jameel Poverty Action Lab’s (J-PAL) cost-effectiveness guidelines, and the London School of Hygiene and Tropical Medicine’s (LSHTM) guidelines for conducting cost analyses of interventions to prevent violence against women.

Our guidelines offer key methods, suggestions, and tools to support cost data collection, including suggestions on how to use these data to conduct an economic evaluation that compares program costs against program benefits, and how these measures can be used to guide resource allocation decisions. Information on program effects or benefits can be drawn from existing impact evaluations. It is not possible to calculate a precise return on investment or cost-effectiveness in the absence of impact evidence, but a robust costing analysis can be conducted to identify the minimum program impact (in monetary terms) needed for a program to break even. This estimate can be compared with existing impact evaluations of similar programs in other contexts to give practitioners an approximate idea of the likelihood of program return on investment.

The remainder of the guidelines lays out the key steps in conducting a costing analysis. We also present two cost data collection tools—a basic tool and an advanced tool—and detail specific scenarios in which each tool may be appropriate for use. These tools were created by modifying the J-PAL and LSHTM tools for the specific purpose of collecting cost data on women’s groups, and are available for download at https://womensgroupevidence.org.
Key Steps

These guidelines address two key goals of collecting cost data:

- Determining how much a certain program costs to replicate in a different setting, or to scale up in the same setting.
- Understanding resource needs and associated costs which can inform program implementers about potential areas for cost savings, suggesting scope for building in cost-efficiencies.
- Estimating overall cost-effectiveness through the “effect” per dollar invested in the program.

Keeping these goals in mind, a comprehensive costing analysis should start with a clear description of the program or intervention to help understand the inputs required and the associated outputs. Laying out the program features, target population, and program and study goals before starting any analysis helps to map out the planned timing and format of data collection, as explained in Figure 1 and the sections that follow.

Figure 1: Key Steps in a Costing Exercise

1. Describe the purpose of the study
   a. What is the goal of the study?
   b. Given the study goals, what perspective (i.e., funders’ or societal) should be taken when measuring costs and outcomes?

2. Describe the intervention
   a. Identify key activities
   b. Identify key resources needed to implement each activity
   c. Identify the target population
   d. Identify the geography and duration of the project

3. Describe the analysis
   a. Which evaluation methods will be used?
   b. Which outcomes will be evaluated?
   c. When will the analysis be conducted?
   d. What are the key costs and indicators that need to be collected?

4. Data collection
   a. Modify the cost data collection template to match evaluation needs
   b. Build in data collection and quality checks

5. Data analysis
   a. Cost data analysis and presentation
   b. Impact analysis
   c. Combining cost and impact analyses to estimate cost-effectiveness

Describing the Purpose of the Study

The first step in conducting an economic evaluation is to clearly define the end goal of the exercise. This involves understanding the key stakeholders that are interested in the exercise and their objective(s). Are governments trying to determine optimal resource allocations across multiple interventions? Are they trying to seek private funding for these programs, and do they need to demonstrate return on investment? Are they interested in information on participants’ costs and time that has not yet been recorded? Answering these questions will ultimately help define the perspective (i.e., funders’ or societal) of the exercise.
Delivering a women’s group program involves several actors, including donors, the government, and the participants who will ultimately be affected by the program. A crucial question is whether the research should only include costs to program funders and providers, or should include costs to the participants as well. In its most basic form, the cost perspective can be divided into the provider perspective and the societal perspective. The provider perspective considers costs borne directly by the parties involved in service delivery (including donors, investors and implementers), while the societal perspective also includes costs borne by the participants or other members of society. The provider perspective does not consider costs (and benefits) to society at large and can therefore generate misleading indicators of social welfare.

Much of the current literature on the costs of women’s groups largely ignores costs to participants and only considers costs from the perspective of “superstructures” (like SHG promoters, including NGOs, governments, or banks). However, donors and investors—including governments and other stakeholders—are often interested in learning whether and to what extent their investments lead to benefits for program participants, in which case the benefits of the program extend beyond financial viability. These guidelines recommend adopting the societal perspective when estimating the cost-effectiveness of different types of women’s groups, based on the assumption that the primary goal of these groups is to achieve improvements in women’s empowerment and well-being.

Describing the Intervention

The next step is to clearly lay out the features of the intervention, including the implementation components, theory of change, specific activities, and resource needs for each activity, as well as the type, geography, and size of the population directly and indirectly targeted. Table 1 provides an example of key women’s group activities supported under the NRLM in India. As shown, the intervention includes a number of different components, beginning with group mobilization and incorporating different forms of financial inclusion and livelihoods interventions.

Table 1: Example of an Intervention Description

<table>
<thead>
<tr>
<th>Program Components</th>
<th>Activities</th>
<th>Resources</th>
<th>Planned Outreach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Program set-up</td>
<td>• Capacity building</td>
<td>• Staffing</td>
<td>Intermediary: 600 program staff members</td>
</tr>
<tr>
<td></td>
<td>• District-level management</td>
<td>• Capacity-building support</td>
<td>Final: 1 million women (total number of</td>
</tr>
<tr>
<td></td>
<td>• Institution building</td>
<td>• Travel support</td>
<td>women targeted)</td>
</tr>
<tr>
<td></td>
<td>• Staff/resource person training</td>
<td>• Building space</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Furniture</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Office supplies</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Computers</td>
<td></td>
</tr>
<tr>
<td>Social mobilization</td>
<td>• Community resource person rounds</td>
<td>• Salaried staff</td>
<td>Intermediary: 400 community</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Community volunteers</td>
<td></td>
</tr>
<tr>
<td>Program Components</td>
<td>Activities</td>
<td>Resources</td>
<td>Planned Outreach</td>
</tr>
<tr>
<td>--------------------</td>
<td>------------</td>
<td>-----------</td>
<td>------------------</td>
</tr>
</tbody>
</table>
| and formation of SHGs | • Community resource person development  
• Facilitation and capacity building of SHGs/village organizations/cluster-level federations | • Capacity-building support  
• Travel support  
• Membership fee from women  
• Group meetings (time and out-of-pocket costs for women) | resource persons  
Final: 1 million women |
| Financial inclusion initiatives | • Initiation of savings and internal lending  
• Exposure visit of bank officials from resource blocks  
• Bookkeeping and opening of bank accounts  
• Financial literacy and credit counseling | • Staffing  
• Computers/laptops  
• Travel support  
• Institution-building support  
• Office supplies | 1 million women |
| Livelihood initiatives | • System for Crop Intensification (SCI)  
• Introduction of new financial products  
• Set-up of community health centers  
• Construction of toilets | • Staffing  
• Capital to set up physical infrastructure for shops/health centers  
• Capacity-building support  
• Seeds and manure to support SCI trainings  
• Labor for toilet construction  
• Construction equipment and resources | 200,000 women |

Source: Hypothetical example based on the Ministry of Rural Development’s (2015) framework for implementation for the NRLM.

Table 1 provides examples of different activities conducted at the group level under the NRLM, along with the resource needs for each of these activities. It is important to estimate the number of participants directly targeted by each component of a program. Initial program set-up activities determine a program’s minimum basic set-up costs and should be apportioned to
(1) the number of trainers, facilitators, community mobilisers, and loan officers who need to be trained to reach the number of women across all program components (in this case, 1 million); and (2) the total number of participants targeted by the program. This helps program staff understand the costs of training the cadre of people who will be needed to deliver or support a program when scaling up, and to estimate the costs per beneficiary. For example, if an SHG program envisions providing 1 million women with some basic services, program set-up costs will be divided across this entire population when estimating the per-beneficiary costs. If the goal is to form groups and provide bank linkage to all members but to only introduce livelihoods activities to 20% of total participants, set-up costs are needed for each program activity to be able to accurately apportion costs to group members participating in each activity. The layout in Table 1 helps to clearly estimate the cost per woman through each specific activity included in the intervention.

Describing the Analysis

The next step is to define the type of analysis that needs to be conducted to achieve the stated evaluation goals. All economic evaluations include the measurement of costs and outcomes. To understand how these costs and outcomes should be estimated and analyzed, the following questions need to be answered:

- What is the evaluation intended to achieve?
- How are program “costs” defined?
- Which costs need to be included to achieve the end goal of the evaluation?

We discuss two broad types of economic analyses of women’s groups, which offer different ways of combining benefits with cost information:

- A cost-effectiveness analysis, which generates a cost-effectiveness ratio
- A cost-benefit analysis, which generates a return on investment

Which evaluation should be conducted? Table 2 provides a summary of these two kinds of economic evaluations where costs and effects are compared. A measure of cost-effectiveness estimates the cost required to achieve one unit of a given outcome, enabling an assessment of the relative cost-effectiveness of different programs that seek to affect the same outcome. A cost-benefit analysis estimates return on investment by looking at multiple outcomes simultaneously, monetizing those outcomes, and generating a common financial unit of “benefit.”

Cost-benefit analyses of social interventions use two key valuation techniques: the revealed preference and the stated preference. Revealed preference techniques represent a market-based approach, in which monetary values are assigned to an outcome based on its market value or price. This approach is straightforward when assessing economic indicators such as increases in asset ownership or consumption, but it cannot be applied directly to outcomes that are not traded on a market. For example, women’s participation in decision making is often used as an indicator for measuring empowerment. However, monetizing the value of an increase in
women's participation in household decisions is not straightforward because that participation does not have a market price. Instead, the stated preference approach may be used, where stakeholders (donors or beneficiaries, depending on the perspective) respond to contingent valuation surveys. These surveys ask respondents to report their willingness to pay for a given increase in an outcome—in this case, women’s participation in household decisions—to measure its monetary value. However, contingent valuation surveys suffer from several biases because respondents do not have an incentive to reveal their true willingness to pay.\textsuperscript{34}

A cost-effectiveness analysis, on the other hand, estimates the costs of achieving a given impact for a single outcome. While this avoids the need to monetize outcomes, it can be difficult to estimate these costs when multiple program outcomes hold value, especially if a societal perspective is used. Combining the monetary value of all outcomes and estimating a program’s return on investment might be appealing in these circumstances, but this approach still has several problems. First, when making resource allocation decisions by comparing return on investment across programs, it can be difficult to find impact evaluations that estimate impacts for the same list of outcomes. Second, since the ROI methodology combines multiple outcomes, this method may lead to a lack of transparency in identifying outcomes for which the program delivers particularly well, which is especially problematic if stakeholders assign different weights to different outcomes. For example, consider an SHG program which includes financial inclusion and livelihoods training for women, and which was shown to generate – 1) a positive income for women who were not earning prior to the program; and 2) an increase in savings due to lower costs of borrowing. When estimating the ROI of the program, the two benefits are summed up and divided by the total cost. However, given that women who had no earnings are now trained and brought into income generating activities, this benefit is likely to lead to greater economic and social empowerment as vocational and life skills are accumulated over the life cycle. Social investors may therefore place a higher value on women’s earnings than on lower costs of borrowing, which cannot be conveyed by the point ROI estimate. Third, as mentioned above, monetizing the value of outcomes related to women’s group interventions comes with significant uncertainty since empowerment outcomes do not have an associated market price.

We recommend conducting a cost-effectiveness analysis when evaluating women’s groups because it allows stakeholders to assign their own values to different outcomes. When a program impact is evaluated on more than one outcome, constructing an impact inventory can be an effective and transparent way of communicating these multiple impacts.\textsuperscript{35} The inventory should include an exhaustive list of all direct and spillover outcomes, from a societal perspective. For example, in an impact evaluation of JEEVIKA—the SHG program in Bihar, India—Hoffman et al. (2018)\textsuperscript{36} found that while the program did not show evidence of a direct impact on participants, it did show significant effects on outstanding debt by lowering interest rates for informal borrowing. The impact inventory for JEEVIKA should therefore include lower costs of borrowing as a program outcome for all beneficiaries (beyond the direct target population of the program). The main purpose of the impact inventory is to ensure that all outcomes are consistently considered when estimating a cost-effectiveness ratio. To support this, the ECWG is developing guidelines for consistent outcome measurement in impact evaluations of women’s groups.\textsuperscript{37}
Table 2: Economic Evaluations Using Cost Data

<table>
<thead>
<tr>
<th></th>
<th>Cost-Effectiveness Analysis</th>
<th>Cost-Benefit Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost measurement</td>
<td>US dollars in constant year prices</td>
<td>US dollars in constant year prices</td>
</tr>
<tr>
<td>Outcome measurement</td>
<td>Single “natural” unit outcome measure</td>
<td>Multiple outcomes monetized and summed up (US dollars)</td>
</tr>
<tr>
<td>Comparison across programs</td>
<td>Single outcome so comparison is feasible</td>
<td>Comparison is feasible if every program has evidence on the same set of outcomes, which is more difficult</td>
</tr>
</tbody>
</table>

Source: Babigumira (n.d.)

How are program “costs” defined? Program costs generally include both economic costs and accounting costs. Accounting costs (or expenditures) measure the financial units spent on physical or human resources. For example, salaries for staff (staffing resources in Table 1) are accounting costs because they are realized directly in monetary units. Economic costs, on the other hand, also include opportunity costs, such as time spent on program activities or the time of unpaid volunteers and other donated resources. Many existing SHGs may use resources that were donated by a third party, which means that the program did not directly spend money on these resources. For example, a donated vehicle may have been used for travel support, as shown in Table 1. This resource may not appear in accounting books, but its opportunity cost can be determined by calculating its replacement cost. Similarly, the time of unpaid volunteers such as community mobilizers can be monetized by estimating the prevailing wages for paid community mobilizers. To estimate the costs of replicating a program, the costs of every resource that was used needs to be estimated, irrespective of the source of the expenditure. This means that all donated resources should be included in cost calculations. Similarly, if beneficiaries substitute their labor hours with time spent in meetings, any foregone wages should be included as an additional cost to the program (under the societal perspective on program costs). This is explained in more detail below.

Which costs should be included given the end goal of the evaluation?

- **Costs to beneficiaries**: Women’s groups often involve women’s time, as well as out-of-pocket expenses for group-related activities. We strongly recommend including these costs in total program costs when they are a program requirement. Karduck and Seibel’s (2004) study of 78 SHGs in Karnataka, India, sought to understand the transaction costs borne by SHG members (with a focus on SHGs that were linked to banks or cooperative societies). The study provided important insights—specifically, that regular meetings were an important factor of group dynamics, with 55% of groups meeting weekly and 31% meeting monthly. The study found that transaction costs, including both real expenditure (like transportation costs) as well as the costs of time spent in meetings, calculated at the local wage labor rate, amounted to $27 per group. 51% of the $27 were out-of-pocket expenditures and 49% were opportunity costs.
• **Costs of resources with overlapping use:** In many scenarios, fixed assets like buildings and computers may be used for more than one program at the same time. Following Dhaliwal, Duflo, Glennerster, and Tulloch (2013), we recommend that if these resources are necessary for program implementation, the associated costs should be included when estimating the program cost-effectiveness. However, this specific cost ingredient should be clearly labeled to enable the end user of the analysis to easily estimate program costs after excluding these costs. This is useful when applying the results to resource allocation decisions in scenarios where these costs are already incurred, for example, when implementers are interested in scaling up a program to a small group of beneficiaries and do not require to set up new infrastructure.

• **High-level management costs:** A special category of resources with overlapping uses includes high-level management and administrative costs. For example, every state in India is responsible for implementing their own model of the NRLM program through the State Rural Livelihoods Mission (SRLM). Although most of the direct costs related to program implementation are borne by the SRLMs, the Executive Committee of the NRLM includes several members of the Ministry of Rural Development under the Government of India. However, these individuals were not recruited specifically for the NRLM and have several other responsibilities in terms of overseeing poverty alleviation programs and policies related to rural development. In general, such costs can be excluded from total program costs for two reasons – 1) These costs will already be included as the “counterfactual” costs against which the additional expenditure of a given program is measured; and 2) Most programs related to women’s groups are usually a part of a larger organization either under the state, or under private organizations. Therefore, such high-level administrative costs will almost always be inevitably incurred irrespective of the specific women’s group program outreach.

• **Fixed costs and variable costs:** Fixed costs usually include start-up costs, such as capacity building and adapting the intervention to the local context; as well as capital costs related to physical infrastructure such as buildings and offices, which cannot be changed in the short run. Variable costs, on the other hand, include day-to-day activities and increase with the level of output (e.g., for every additional training or for every additional beneficiary). Determining whether both types of costs should be included in the analysis depends on the goal of the analysis. If the goal is to understand whether a program will be cost-effective if replicated in another context or scaled up, fixed costs should be included. If the goal is to estimate the additional cost of reaching a small number of participants under the current set-up, such costs may be excluded. For example, start-up costs (which are usually included in fixed costs) may include expenses for basic infrastructure, such as a building and a small number of vehicles. Reaching new participants within the same cluster of villages may not require additional buildings or vehicles. However, such costs will likely be incurred when replicating the project in new districts, at which point the fixed costs should be revisited. In summary, the distinction between fixed and variable costs depends on the context and scale of a given project; some costs that may be fixed in the short term with a small target population might become variable over the longer term as a program scales up.
Table 3 summarizes costs or expenses that require additional consideration when deciding whether to include or exclude them from analyses.

### Table 3: Factors to Consider When Including or Excluding Costs

<table>
<thead>
<tr>
<th>Factors to Consider</th>
<th>When to Exclude</th>
<th>When to Include</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monitoring and evaluation (MEL) costs versus research costs</td>
<td>Which costs are necessary to implement the program?</td>
<td>Exclude research costs that are not related to implementation—for example, a process evaluation.</td>
</tr>
<tr>
<td>Goods provided for free</td>
<td>Were these goods used to meet resource needs during implementation?</td>
<td>Exclude goods or services that were provided for free if they were not directly used for service delivery and/or were not necessary for implementation.</td>
</tr>
<tr>
<td>Costs to beneficiaries</td>
<td>Did participants spend any money in the form of fees or other direct expenses, such as travel or opening bank accounts?</td>
<td>Exclude costs to participants if the analysis is taking the provider perspective.</td>
</tr>
<tr>
<td>Nonmonetary costs (or opportunity costs)</td>
<td>Did participants spend major amount of time on activities related to women’s groups?</td>
<td>Exclude nonmonetary costs if only taking accounting costs into consideration.</td>
</tr>
</tbody>
</table>
| Start-up or fixed costs versus variable costs            | Do stakeholders want to understand the total resource need to replicate this program?  
Do stakeholders want to understand the cost of delivering services to additional participants under the existing program? | Exclude start-up costs when estimating current resource needs under the current program. | Include start-up costs when estimating total resource needs.  
Include start-up costs when estimating the return on investment or cost-effectiveness of a program. |
Data Collection

Cost data collection often comes with a number of challenges, as expense data are usually incomplete or inaccurate and verifying the source of expenditure can be cumbersome. While it may be easiest to collect these data through program budgets, this approach may result in critical inaccuracies. For example, a given program’s budget may exclude resources with overlapping purposes that have already been budgeted for by another program.

A useful way of collecting cost data is to take a top-down or bottom-up ingredients approach. Under a top-down ingredients approach, cost ingredients are defined a priori for intervention activities described in Step 2 (see Figure 1) that are deemed necessary for program implementation or replication purposes. The total cost of a program is then broken down across these ingredients using predetermined allocation criteria. The allocation criteria are usually developed based on an understanding of project implementation and resource requirements. For example, a community resource professional (CRP) may have a standard cost per full-time equivalent (FTE). In a top-down ingredients approach, implementers usually include different buckets of project activities and then allocate CRP FTEs to each bucket based on their anticipated level of effort.

A more comprehensive way of collecting cost data follows the bottom-up ingredients approach, where each input in the production process is listed alongside its unit price, and the cost of each activity is derived by summing all inputs that make up the activity, multiplied by their unit prices. This generates the unit cost of each category, which can be broken down into its smallest components for adjustment in other settings or scenarios. The bottom-up ingredients approach is the most accurate way of collecting cost data and is more resource-intensive in terms of time and costs incurred on data collection.

Building on the J-PAL guidelines for conducting cost-effectiveness analyses, we suggest using the following 10 categories, at a minimum, to classify cost ingredients for women’s groups:

- Capital costs (including buildings, vehicles, etc.)
- Targeting costs
- Staffing costs (staff salaries, as well as volunteer staff with an equivalent wage)
- Travel costs (including fuel costs)
- Training costs for staff
- Training costs for beneficiaries
- Materials and resource costs (including office supplies and software)
- Monitoring costs (related to standard monitoring of program operations, such as routine tracking)
- Other direct costs (including maintenance, utilities, etc.)
- Beneficiaries’ costs
Data can then be collected across each input from activity records or expense reports. When considering the opportunity costs of beneficiaries’ time, time-use surveys can be used to collect information on participants’ time spent on group-related activities. These surveys should include questions such as: How far are group meetings from your home? When and how often do women meet? What would women be doing in the absence of these group meetings? When detailed time-use surveys are not feasible, such data can also be collected qualitatively through interviews with the implementation staff. This information can then be combined with local wage-rate data to estimate the opportunity cost of women’s time. It is important to clearly state the method chosen to estimate a foregone time value, its rationale, and the value ultimately chosen, even if the individual is not in paid employment at that time.41

To collect data on time-use and resource costs, we recommend using the protocols on direct observation or semi-structured interviews that are included in the guidelines developed by Ferrari et al. (2018).42 Most of the expenditure data needed to fill in the tools can be obtained from the NGO/company books, personnel reimbursement records, and petty cash records. The value of donated time (volunteers) or goods (e.g., vehicles, spaces) can be determined using market wages for individuals with similar qualifications or replacement costs for donated goods. Where possible, it is preferable to collect cost data prospectively, while the intervention is being delivered and developed.

Data collection can be prospective (occurring alongside implementation) or retrospective (occurring after program delivery). Prospective data collection is generally more desirable because it captures more accurate information; retrospective data collection is affected by recall bias and may be further limited by imperfect or incomplete financial records. In these guidelines, we present two tools that use the bottom-up ingredients approach to measure costs, tailored to suit different levels of data availability. Specifically, we offer one basic tool and one advanced tool, which were originally developed by J-PAL43 but have been modified to include components specific to women’s groups. In addition, we also provide a cost summary and analysis tool, which uses data on costs and impact to generate cost-effectiveness ratios. These tools are available for download at https://womensgroupevidence.org.

Table 4 summarizes factors to consider when deciding which of the two tools to use for an evaluation, based on the evaluation goal and data availability. The basic tool requires the user to identify the total costs associated with each ingredient. It then uses the ingredients to calculate the average cost per beneficiary in US dollars, based on the appropriate exchange rate. The advanced tool requires more detail, providing line items that are likely to be applicable in the program’s context, against which the user can fill in the associated cost. This “micro-costing” approach is preferable if calculating cost-efficiency is one of the goals of an economic evaluation. It can also help to identify ingredients with overlapping uses, enabling users of the analysis to determine which costs can be excluded when thinking of replicating a program.

We have adapted these tools to suit the context of women’s groups and have added data collection elements on different program components (informed by the more comprehensive tools developed by the LSHTM; see Ferrari et al., 2018).44 Determining the proportion of resources spent on different program components is key for women’s groups, which often start
as savings or credit groups but over time serve as vehicles for delivering multiple services, including health, nutrition, and farm and nonfarm-based livelihoods components.

**Table 4: Summary of Basic and Advanced Cost Data Collection Tools**

<table>
<thead>
<tr>
<th></th>
<th>Basic</th>
<th>Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retrospective data collection</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Prospective data collection</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>No access to detailed expense reports</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>No restriction on resources for data collection</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Interested in learning specific cost breakdown across categories</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Interested in learning the quantity of input across each category</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Evaluation needs to align costs to different program components</td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

Data Analysis

When combining costs with estimates of program impact, costs need to be accumulated from the start of the program (the base year) until the year when the impact analysis was conducted (the end year). Once data on cost ingredients have been collected, these costs need to be adjusted for exchange rates, inflation rates, and time value, to be expressed in a standardized currency and year.

**Adjusting for exchange rates:** To ensure that all costs are expressed in a common currency, the norm is to report costs in both local currency and US dollars (USD). There are two common mechanisms for currency conversion: purchasing power parity (PPP) and the market exchange rate. PPP adjusts for differences in purchasing power for a standard basket of goods across economies. However, the actual value or “worth” of these goods may not be directly comparable across countries. The market exchange rate is based on the demand and supply of international currencies on the market and is the preferred choice when financial flows are involved. For a costing analysis, where the goal is to determine the overall need for financial resources, we recommend using market exchange rates to convert all local prices to US dollars.

**Adjusting for inflation:** Since program costs are often realized over multiple years across the lifespan of an intervention, these costs need to be stated in base-year prices, after adjusting for inflation. For example, if an intervention lasted for 3 years from 2008 to 2010, the costs realized in 2009 and 2010 need to be converted into 2008 prices. This can be done by using the GDP deflator or the Consumer Price Index (CPI). We recommend using the CPI for two reasons. First, unlike the GDP deflator, the CPI captures price changes in imported goods in addition to goods produced domestically. Second, the CPI captures the change in prices for a fixed basket of goods, while the basket of goods varies under the GDP deflator. Therefore, the GDP deflator
tends to understate the decrease in consumer welfare that results from price increases of goods that are no longer included in the GDP deflator basket.

In this example, to express costs incurred in 2010 in 2008 prices, we use the following formula:

\[ \text{Cost}_{2008} = \text{Cost}_{2010} \times \frac{\text{CPI}_{2008}}{\text{CPI}_{2010}} \]

We recommend that for tradeable resources that are generally traded on the international market (e.g., laptops, vehicles), the conversion to US dollars should be made at the year of purchase, and the inflation adjustment should be done in US dollars. For non-tradeable goods/services (such as the wage of a local worker), inflation adjustments should be done in the local currency, and the exchange into US dollars should only take place in the reporting year, because the good is only traded on the local market.

**Discounting:** Project costs that are realized in different years need to be adjusted to base-year prices to reflect different time preferences using a discount rate—an indicator for the rate at which the future value of a resource is discounted. The discounting takes care of the fact that individuals value spending money in the present higher than in the future, implying that the future value of a commodity will be lower than the present value. There are multiple approaches to assigning parameters to this “discounting.” A “descriptive” approach is based on the opportunity costs of resources and actual behavior, while a “prescriptive” approach relies on ethical considerations of intergenerational equity. For example, assigning a positive rate of time preference may result in placing a higher value on current programs at the expense of investments that may make future generations better off.\(^{45}\) Cost-effectiveness analyses in international development generally use the social opportunity cost of capital to determine discount rates, due to a lack of empirical data on rates of time preference across countries. Using a standard discount rate based on opportunity costs of 10\(^{\%}\),\(^{46}\) the present discounted value (PDV) in 2008 of costs incurred in 2010 (3 years from the base year) is calculated as:

\[ PDV = \frac{\text{Future Costs}}{(1 + 0.10)^3} \]

Finally, before applying costs to the calculation of the cost-effectiveness ratio, they should be converted into end-year prices (which is when program impact is estimated). Figure 2 lays out a sequential summary of estimating program costs.
Figure 2: Computing Intervention Costs After Data Collection

**Non-tradeable resources**

- Apply local inflation rate to indicate annual costs in base-year prices in local currency
- Estimate PDV of future cost streams in base year in local currency
- Apply base-year USD inflation rate to estimate costs in base-year USD prices
- Apply USD inflation rate to indicate costs in end-year USD prices

**Tradeable resources**

- Apply exchange rates to indicate annual costs in USD
- Apply USD inflation rate to indicate annual costs in base-year USD prices
- Estimate PDV of future cost streams in base year in USD prices
- Apply USD inflation rate to indicate costs in end-year USD prices

**Calculating a program’s return on investment or cost-effectiveness ratio:** Once overall program costs have been estimated, they can be compared to estimates of program “effectiveness” from impact evaluations to generate a program’s return on investment or cost-effectiveness ratio. To calculate return on investment, all potential benefits must be converted into monetary units and summed up to generate a total program benefit in the same currency and year used for the costing analysis. The program’s return on investment (ROI) can be generated using the following formula:

\[
ROI = \frac{\text{Total program benefits}}{\text{Total program costs}}
\]

A cost-effectiveness ratio, on the other hand, denotes the cost-per-unit impact generated for a given outcome. For women’s groups, for example, the cost-effectiveness ratio (CER) can be generated using the following formula:

\[
CER = \frac{\text{Total program costs}}{\text{Number of additional women empowered because of the program}}
\]
When conducting a cost-effectiveness analysis, it is useful to estimate an incremental cost-effectiveness ratio by dividing the difference in cost between a program and its “comparator” by the difference in their outcomes. A comparator in this case is the outcome that would exist in the absence of the program. This can be crucial when estimating the cost-effectiveness of women’s groups that have multiple components. For example, when analyzing the cost-effectiveness of adding farm-based livelihoods activities to SHGs that originally focused only on savings, the basic costs of SHG mobilization will be incurred irrespective of the livelihoods component. As a result, the cost-effectiveness ratio will be estimated by dividing the cost of adding the livelihoods component by the additional outcome derived from this component compared to savings-based SHGs:

\[
\frac{(Costs \ of \ savings \ and \ livelihoods \ SHGs) - (Costs \ of \ savings \ SHGs)}{(Women \ empowered \ in \ savings \ and \ livelihoods \ SHGs) - (Women \ empowered \ in \ savings \ SHGs)}
\]

**Generalizability**

When comparing the cost-effectiveness of programs that are implemented in different settings, it is important to note the location-specific parameters that could influence program cost-effectiveness. Usually, impact estimates can be standardized to outcome levels in the absence of the intervention and expressed in a common unit (typically a standard deviation) to facilitate comparisons between programs. However, other factors could have a direct impact on resource needs and costs—for example, the skill or experience of local staff. For such comparisons, it is often useful to select the appropriate variables that could influence program costs and provide a range of cost estimates for different values these variables can be expected to take on. For example, the District Poverty Initiative Project (DPIP) in Telangana, India—which had been implementing an NRLM-like model for many years prior to the implementation of the NRLM in other states—may report that its staff receive CRP trainings twice a year. States that are newly implementing the NRLM, however, may need a higher number of trainings to achieve the same level of community mobilization skills. In this case, we can estimate how the program costs, and therefore the cost-effectiveness, of the Telangana program would change if trainings were increased to four times a year or six times a year.

**Conclusion**

In recent years, international organizations and donors like the Department for International Development (DFID), the United States Agency for International Development (USAID), the World Bank, and the Bill & Melinda Gates Foundation have been increasingly demanding that impact evaluations of development programs be accompanied by rigorous measures of cost-effectiveness or return on investment.\(^{47,48}\) There have also been an increasing number of guidelines highlighting possible ways of conducting such evaluations, two of which were developed by J-PAL and the LSHTM (around education and gender-based violence interventions, respectively). Our guidelines aim to build on this literature and specifically apply
these lessons to women’s groups, which are a common vehicle for delivering poverty alleviation interventions in Africa and South Asia.

Several challenges have precluded researchers from reporting information on costs or cost-effectiveness. First, cost data are hard to collect after program implementation. Second, even when cost data can be collected through expense sheets or audit reports, it is difficult to match these data with program activities if information about those activities has not been recorded during implementation. Third, there is limited knowledge about how to use cost data in a standardized way, which makes it challenging to compare costs across different settings. Finally, impacts are not consistently measured across programs, which makes it challenging to compare costs against impacts in order to make resource allocation decisions. While there is no single solution to these challenges, these guidelines offer possible ways to address them by laying out key steps—and associated assumptions—in conducting different types of cost-effectiveness analyses for different evaluation goals, providing two cost data collection tools (a basic tool and an advanced tool) and a cost analysis tool, and laying out specific scenarios in which each tool may be appropriate for use.
References


Guidelines on Estimating Costs and Cost-Effectiveness of Women’s Groups in International Development


