How Do We Gather Baseline Data for Impact Evaluation

IFAD is placing increasing emphasis on impact evaluation of the projects that it supports. Good-quality impact evaluation not only provides a justification for the investment in the project (assuming that the evaluation shows a worthwhile positive impact), but we can also generate some useful lessons on what development initiatives work and what do not work.

The problem is how do we do this? What is needed is evidence that goes beyond a story or an anecdote, which can convince readers of the document that the project has generated certain results. This evidence is usually generated by a some sort of sample survey. Other means of gathering evidence can also be used, such as participatory rural appraisal, focus group discussions and case studies, but these lack the rigor of a sample survey in that it cannot be proven (by statistical analysis) that the information gathered is really representative of the entire group under investigation. However, these methods are very useful in understanding processes and why change happens (or does not happen).

A formal survey uses a questionnaire to gather data from a randomly selected sample of project participants with a sufficient number being interviewed to produce results that are representative of all participants. To measure change, some sort of comparison is needed. Two types of comparison are possible. First, the "before and after": compares current information on the project sample with information from before project interventions took place. In this way, the change that took place can be measured. Second, the "with and without" compares information on the project sample with that on a sample of households that did not take part in project interventions (a "control group").

The drawback of relying only on changes measured "before and after" is that changes may well have taken place without the interventions of the project. A simple "with and without" comparison does not provide information on change. We need to estimate the "change without the project", known as the "counterfactual." This can be done be comparing "before and after" changes for a project sample with changes for a control sample. This is known as the "difference in difference" (DID) approach.

To measure these changes, information on the pre-intervention (baseline) situation is needed from both the project and control samples. For ongoing projects, this can be difficult. Baseline information may not have been collected or may not cover the indicators needed to measure project impact. Although projects often collect baseline data on project participants, they may not have it for a control group.

One way of obtaining baseline information is to ask survey respondents to recall the pre-intervention situation—in other words, to make their own estimate. This "recall" method is, of course, less accurate than having real data collected at the time (when it was "fresh"). However, it can be surprisingly precise. Last year, an IFAD project in Bangladesh, the Micro-Finance for Marginal and Small Farmers Project (MFMSFP) came to an end. An impact study was commissioned, carried out by the Nielsen Company Bangladesh Ltd. This used the recall measure to collect data on the pre-project situation. However, the project also had pre-project information on some indicators as households profiles were completed at the time people joined the project groups (mostly in 2006 and 2007), and an analysis of a sample of 600 of these profile forms had been made as a baseline study.

The following table compares recall baseline data from the impact study with actual baseline data from the household profiles. This shows that, for most of these indicators, indicator data at entry into the project (i.e., at baseline) using the actual baseline profiles is within 10% of that collected in the impact survey using recall estimates. The difference is significantly greater for the number of poultry, use of latrines, and total household income. These are indicators that would seem to be more difficult to recall with any degree of accuracy, especially income.

Table 1. Comparison of recall baseline data with actual baseline data.

			Before	After
Average land owned	Acres	Baseline	1.39	
		Impact	1.32	1.61
Paddy production	Acres	Baseline	2.71	
		Impact	2.50	2.89
	Total tons	Baseline	5.33	
		Impact	4.90	5.68
Number of animals per household	Cattle	Baseline	1.59	
		Impact	1.70	2.00
	Sheep/goat	Baseline	1.37	
		Impact	1.30	1.40
	Poultry	Baseline	14.72	
		Impact	13.20	19.20
Household income	Taka/year	Baseline	77,271.00	
		Impact	92,934.00	108,781.00
Open pit or no latrine	% of households	Baseline	24.00	
		Impact	27.00	5.80

This analysis suggests that recall data can be used to reconstruct a pre-intervention baseline, providing indicators that respondents can reasonably remember.

Recall also has the advantage that the same households are providing both baseline and impact-level data. The data in the table show that changes between the "before" and "after" situations can be quite small. Collecting baseline and impact data from separate samples means that different households may be interviewed in baseline and impact surveys. Inevitably, there will be underlying differences in these samples (i.e., not caused by project interventions).

For example, the "before" average number of sheep and goats per household in the baseline profiles was 1.37, compared with 1.30 recalled in the impact survey. This difference is only 5.4%, but the "after" project

average number only increased to 1.40. In this case, recall gives a much bigger increase of 0.10 (7.7%) than if the increase was calculated using the baseline profiles (0.03 = 2.2%). Ideally, real baseline data are collected from a sample of households at the start of the project, with these same households being re-interviewed for mid-term and/or completion surveys. This is known as a "panel sample", but it adds another level of complication (finding the same households for each round of survey) to the data collection process.

The conclusion is that, even if baseline data were not collected at the start of the project, an impact evaluation can still be carried out using a reconstructed baseline based on recall. But this evaluation should be limited to indicators that respondents can recall with reasonable accuracy. Secondary data can also be useful to confirm if the changes apparent in the control group reflect the changes in the larger population.

Reference

IFAD Asia Blog

Bio-sketch and contact details

Mr. Edward Mallorie is an economist with extensive experience of project design and implementation, monitoring and evaluation, and policy formulation for rural development and poverty reduction. His experience includes micro-credit, agriculture, market access, water resources, small farmer development, agroindustries, marketing, forestry, livestock, fisheries, food aid and environmental protection. He can be reached via email at EMALLORIE@aol.com.