------- DRAFT NOT FOR QUOTATION --------
12 November 2009

Evaluating Poverty Impacts of Bottom-of-the-Pyramid Irrigation Technology Supply: IDE's Rolling Baseline Approach to Household Income Impact Assessment[1]

Authors:
Giel Ton[2] - LEI Wageningen UR,
Stuart Taylor[3] – International Development Enterprises

In collaboration with:
Joost Vlaming – Envista Consulting
Sabine Hiller - LEI Wageningen UR

Corresponding author:
Giel Ton
LEI Wageningen UR
giel.ton@wur.nl
P.O. Box 35, 6700 AA Wageningen
the Netherlands
Phone: + 31 (0) 317 484420 / +31 (0) 70 3358201
Fax: + 31 (0) 317 484490

[1] This paper results from the Rural Prosperity Initiative (2007-2010), a program of International Development Enterprises (IDE) in Zambia, Ethiopia, Nepal, Myanmar and Vietnam, financed by the Bill and Melinda Gates Foundation and the Dutch Directorate of International Cooperation (DGIS). The Agricultural Economics Research Institute (LEI Wageningen UR) has been subcontracted by IDE to provide guidance on monitoring and evaluation, and to co-design the related household income impact assessment methodology. The authors want to thank Elizabeth Weight, Bob Nanes, Deepali Nath, Fritz Kramer, Jim Taylor, Loretta Payne and Michael Roberts working for IDE, in co-developing the monitoring and evaluation framework and related instrument, and, Daniel Kanyembe – IDE Zambia, Amitendra Chaudhary – IDE Nepal, and Alemayehu Eshete – IDE Ethiopia for testing and coordinating the implementation of the rolling baseline methodology in the field. The authors benefitted from critical comments on the initial design from Roy Steiner and Melissa Ho, of the Bill and Melinda Gates Foundation, and Lawrence Haddad of the Institute of Development Studies. However, all errors related to the design of the instrument and the data analysis in this paper are under exclusive responsibility of the authors.

[2] Giel Ton works as an agricultural economist in the Agricultural Economics Research Institute (LEI Wageningen UR), specialized in smallholder value chain development. The research for this paper has been made possible through an additional grant from the DGIS-WUR partnership project 'Improving livelihoods and resource management in the Central Rift Valley of Ethiopia' and the WUR-LNV knowledge base project 'Viable Value Chain Configurations'.
[3] Stuart Taylor works as executive director for IDE in Canada. Epidemiologist by training, he supervises the monitoring of household income impacts for IDE in the Rural Prosperity Initiative.

Abstract

Poverty reduction is the prime goal of many donor funded agricultural development projects. This creates an urgent need for instruments that can give credible figures on changes in the household incomes. However, the ways to measure these changes and translate these in attributable impacts is not straightforward and needs careful design. These items become part of a wider set of factors that create the poverty impact. Therefore, to credibly assess impact, a control for these exogenous and confounding factors is needed. Many traditional methods for impact evaluation are not suited for monitoring impacts in self-selecting client groups. It is difficult, if not impossible, to find proper 'treatment' and 'non-treatment' groups and derive the 'treatment effect' as a proxy for attributable household income impact. In self-selected client groups confounding variables make income comparisons between cohorts prone to biases. We present a novel methodology for assessing changes in household income attributable to technology adoption by smallholder farmers in Nepal. The rolling baseline survey methodology was applied by IDE, a non-profit organization that develops and promotes market-based supply chains of low-cost micro-irrigation equipment for households living at the bottom of the pyramid. Household income is calculated yearly by estimating the gross margins of farm and off-farm activities, before and after technology adoption. Pre-adoption household incomes of successive cohorts are used to construct a proxy control for exogenous factors such as price fluctuations and weather conditions. This paper describes the application of the rolling baseline method in Nepal and tests assumptions underlying the methodology related to inter-cohort variation and recall bias. The methodology was applied by IDE, a non-profit organization that develops and promotes market-based supply chains of low-cost micro-irrigation equipment for households living at the bottom of the pyramid. Household income is calculated yearly by estimating the gross margins of farm and off-farm activities, before and after technology adoption. Pre-adoption household incomes of successive cohorts are used to construct a proxy control for exogenous factors such as price fluctuations and weather conditions. This paper describes the application of the rolling baseline method in Nepal and tests assumptions underlying the methodology related to inter-cohort variation and recall bias. Test results indicated the need to adapt the formula used to calculate income impact with data of the counterfactual, to limit vulnerability of inter-cohort variation. Further research will have to be done on the recall bias, as test results point to differences in reported income sources as a result of longer or shorter recall periods. The adapted rolling baseline methodology is suited to evaluate income changes attributable to technological innovations with impact over short periods of time.

Key words:

Impact assessment – Rural development - Farm economics – Survey methodology – Experimental methods - Nepal

Introduction

There is growing recognition of the need for better impact assessment in value chain development support interventions. The Donor Committee on Enterprise Development (DCED, 2008) developed minimum standards for quantifying achievements in which monitoring income changes and calculating attribution to program interventions is a required practice. This information is necessary for private sector development programs to demonstrate their achievements and provide program management with regular monitoring information (DCED, 2008). First generation Bottom-of-the-Pyramid (BoP) ventures are typically strategies of companies to target poor consumers both to generate a profitable venture and to enhance the poor households' wellbeing or economic capacities (Simanis, Hart et al. 2004; London, Anupindi et al. 2009). The second generation of BoP strategies include interventions that target the poor with technologies and services to increase their productive capacities (Simanis and Hart 2008). While the first generation BoP has multinational companies as prime actor, the second generation calls for the complementary cooperation of a number of parties, like local NGOs, and local small and medium enterprises (Kandachar and Halme 2008). Increasingly, the poverty alleviation component attracts donor support to eliminate the initial start-up costs. And, like all donor-dependant public-private endeavours, these BoP support programmes will need to meet these minimum standards of DCED. The BoP Protocol (Simanis and Hart 2008) emphasizes the need to track the "triple bottom line" impacts associated with BoP enterprises. Landrum (2007) points to the lack of empirical evidence of the impact of BoP ventures on poverty. This lack of evidence does not only reflect a low priority on measuring impacts, but also the lack of appropriate, lean and credible instruments to do so. In this paper we focus on impact assessment of a value chain support strategy that emphasize co-creation and socially embedded business partnerships between enterprises, local entrepreneurs and poor people (London and Hart 2004; Danse, Vellema et al. 2005). IDE works in cooperation with local partners, private businesses and development organisations to set up supply chains for micro-irrigation technology and offers business support services to enhance horticulture production and marketing. The design of the micro-irrigation equipment introduced in the market is targeted to comply the requirements of 'the other 90%' of poor households (Polak 2008), and the distribution of technology is market-based through a supply chain of private companies assembling and selling the irrigation devices (Heierli and Polak 2000). The marketed technology is used to generate income from enhanced horticultural production. This irrigation-induced agricultural intensification process is meant to be self-sustaining and self-enhancing after a short period of external support. The start-up costs, the design process of micro-irrigation devices, the organisation of the supply chain and the training of farmers in getting the best out of the technology, are assumed by public and private grants. This expected outcome is that smallholders, clients of the irrigation supply chain, increase their incomes to a level that they can escape their poverty (Polak 2008).

Its strong emphasis on market-based technology sales to the poorest sections of the rural population gives it features that characterize many BoP-ventures, especially the need to assess impact in clients, not average populations. To present convincing evidence on this impact, IDE applied an impact assessment instrument that has the potential to combine operational efficiency with credibility in impact calculation. We describe this methodology and test some of the critical assumptions, using data from

Nepal. Finally, we indicate conditions under which the methodology will be most useful for measuring the poverty impacts of technology sales in client groups.

Rolling Baseline Survey methodology

Design

To assess the impact of technological innovation, the production and income levels of adopters have to be compared. However, along with a new technology, a range of other factors may have differed that influence these outcomes. Different (quasi)-experimental designs have been developed to control for this exogenous influences and assess how the treated person would have developed without the treatment - the so-called ‘counterfactual’. These quasi-experiments must be designed to rule out the most obvious threats to validity (Shadish, Cook et al., 2002). These threats include measurement error and biases in the data collection and analysis. We had three main reasons to develop the rolling baseline survey methodology as a preferable alternative to other experimental and quasi-experimental designs, such as randomized control trials (Duflo, Glennerster et al. 2006) and propensity score matching (Dehejia and Wahba 2002)

- First, a market-based approach to the distribution of its products and services implies that the “treatment” group are customers, that is self-selecting. Clients decide to use (or not) certain technology or services. It is therefore difficult to predict who will actually opt in. Products and services are in many cases delivered through third-party service providers and co-facilitated by a number of different local development organisations. Given this self-selection, randomized control trial methods are not appropriate.
- Second, the complex and intertwined livelihood strategies of poor households, their cropping patterns, the availability of water for irrigation of the fields, market access for horticultural products, and the institutional environment are all highly geographically specific and constrain the use of matching methods like Propensity Score Matching[4]. Matching models require information on a very wide range of background characteristics from a very large group of non-adopting households. Even if the resources to do this were available, the risk of missing an important latent unobserved external factor would be high (Heckman, 2005).
- Third, organizations or companies that are active in poor areas often operate on limited budgets in remote areas with poor communication and transportation infrastructure. They require a survey that is lean, affordable and feasible but with sufficient power to detect changes in annual income related to the adoption of specific agricultural technologies and/or services by their clients.

We designed the rolling baseline survey approach to assess income changes in customer households during the lifetime of a project. The survey provides annual data on the pre- and post-adoption income in a sample of client households.

[4] In PSM, by means of propensity scores, matches are made between households on similar pre-adoption characteristics. An important assumption for this matching process is the Conditional Independence Assumption (CIA): if the observables variables are controlled for, the difference in outcome is due to the treatment.

Following the vocabulary of Shadish, Cook and Campbell (2002), the rolling baseline survey uses a "non-random treated cohort design with a retrospective pretest and a single posttest." In OXO-notation it is written as:

	Year 0		*Year 1*		*Year 2*		*Year 3*
NR	$O_{1,r}$	*X*	O_2				
NR			$O_{3,r}$	*X*	O_4		
NR					$O_{5,r}$	*X*	O_6

In this notation, *O* indicates the observed outcome (household income), *X* the intervention (technology adoption), *NR* indicates that they are nonequivalent groups due to non-random assignment, and the dashed line separates the cohorts (adopters in successive years). The subscript *r* indicates that the outcomes are measured based on recall.[5]

The income changes calculated from the survey data reflect changes in the agricultural systems and livelihood strategies of the households. The impact estimate corrects for exogenous factors like output prices, currency rates and weather conditions. However, to attribute income changes to the interventions, we need to go one step further and explain how these changes in land use are produced, and what the role of the intervening agent is in relation to other stakeholders active in the intervention area. To evaluate this causal connection between the interventions and the resulting changes in farming practices and household income, we need causal process observations to complement the data-set observations from the survey. The causal mechanisms underlying this correlation must be checked by triangulating information from a wider body of evidence, combining the analysis of the quantitative data from the surveys with the qualitative information derived from sub-sector studies, focus group discussions and livelihood impact case studies that explore impacts of the intervention strategies (technologies) on livelihood strategies[6]. Statistically significant differences between baseline and attributable post-intervention income, supported by qualitative evidence for a causal pathway between the intervention and the observed income impact (the 'program mechanisms'), together, will support the evaluative conclusions about the income impact attributable to the intervention.

Income impact calculation

The survey reconstructs household income based on farm and non-farm activities. Farm income consists of crop production and animal husbandry. We estimated crop income based on self-reported household production and output prices (by crop) and on input costs (general), and further differentiated target crops (irrigated and promoted crops – mainly horticultural) and non-target crops (traditional crops – mainly field crops). The one-hour questionnaire captures the current year and the year prior to participation in the intervention (before adopting the technology). With irrigated horticulture, we expect that the 'before' and 'after' cropping systems are

[5] Essentially, all survey data are based on respondent recall as even the "current" year data are based on surveys conducted after the end of the second cropping cycle of the agricultural year, including harvesting and marketing. The "retrospective" pretest measurement (r) refers to a longer recall period - between one and two years.

[6] Focus group discussions and livelihood impact case studies are an integral part of IDE's monitoring and evaluation framework and collected by their field offices.

sufficiently different that respondents will recall the important details with reasonable precision.

	Year 0		*Year 1*		*Year 2*		*Year 3*
Cohort 1	$O_{1,\,r}$	X	O_2				
Cohort 2			$O_{3,\,r}$	X	O_4		
Cohort 3					$O_{5,\,r}$	X	O_6

Observed changes in income cannot be directly attributed to IDE-promoted technology adoption. Other exogenous variables - including weather, prices, inflation and other economic circumstances - can influence household income. The rolling baseline methodology intends to (partially) control for these exogenous influences by applying an index that incorporates the autonomous income change that households would have exhibited without IDE's intervention(s). When the two cohort match on their background characteristics, like field size, household, size and their main economic activities, the recalled pretest income of the second cohort could be used as the counterfactual. Next to income change (O_2-$O_{1,r}$), net income impact could be calculated with O_2-$O_{3,r}$.

Control on critical assumptions

The rolling baseline approach is based on the assumption that farmers can recall production data for the year preceding their adoption of the technology. We assume that differences in assessments of field size, yields, home consumption and sales prices will level-out when computing the averages. To test this assumption, we used data from a follow-up survey, which monitors income changes among a subsample of households in the three years following adoption of the technology or service. In Year 1 after the intervention (X), farmers reported their most recent production data (O_{2a}) and their previous, baseline production data ($O_{1,\,r}$). In Year 2, farmers reported their most recent production data (O_3) and recalled their previous (Year 1) production data ($O_{2b,\,r}$). This allowed us to compare Year 1 production data as reported after less than one year (O_{2a}) with Year 1 production as reported one to two years later ($O_{2b,\,r}$), and test for a significant difference and bias as a result of a longer recall period. The design of this recall bias test on the follow-up survey data can be summarized as:

	Year 0		*Year 1*	*Year 2*	*Year 3*
NR	$O_{1,\,r}$	X	O_{2a}	O_7	O_8
NR			$O_{2b,\,r}$		

Further, the rolling baseline methodology assumes that successive customer cohorts do not differ significantly in their fundamental characteristics. In other words, year 1 adopters are similar to year 2 and to year 3 adopters in terms of livelihood strategies, cropping patterns, marketing behaviour and pre-adoption poverty status. Thus, although outcomes are measured in different cohorts, they are assumed to be comparable. This assumption is tested by comparing and analyzing the characteristics of the customer cohorts with an independent sample T-test.

Application

Background

In the past, IDE has assessed household income impact of micro-irrigation technology sales through surveys asking respondents to self-assess the impact on their family income. The information proved useful for communication purposes but had limited validity as it was related to self assessment of total household income and attribution was assumed to be straightforward, without referring to a counterfactual. In developing the monitoring and evaluation framework for its Rural Prosperity Initiative (RPI), funded by the Bill and Melinda Gates Foundation (BMGF) and the Dutch Directorate for International Cooperation (DGIS), IDE emphasized the need for a more robust approach to estimate the average change in household income attributable to its interventions. The main objective of the RPI project is to raise the annual income of 40,000 low-income households by 200-250 US$ (IDE, 2006). Based on past experience, IDE expected much of this impact to be generated in the first year after adoption of technologies or services.

Data Collection

IDE customer households were registered as follows:

1. When a person purchases an IDE-promoted technology, the vendor records the customer name and address. IDE staff collect vendor records regularly and enter the household information into the registration form.
2. When a person purchases or receives an IDE-promoted service (training, agronomic support, market support, etc), IDE field staff record the household name and address on the registration form.
3. When a household joins an IDE intervention by adopting a recommended practice in relation to agricultural inputs, production, or marketing, IDE field staff record household information on the registration form.

Household information from the registration forms was entered into a database. As it was possible that the same farm household could be listed multiple times, IDE field staff reviewed the database of registered households to eliminate duplicate entries. To further control for duplication of customers registered, IDE assigned each household a unique household identification code. Survey samples for the annual rolling baseline survey were selected randomly from the list of newly registered customers in each year using two-stage cluster sampling. In the first stage, villages or catchment areas were selected randomly from the list of villages/ catchment areas in the IDE RPI project. In the second stage, respondents in these villages were randomly selected using weighted probabilities based on the ratio of customers in that cluster to the entire customer population. IDE used a yearly survey sample size of 200 households.

The survey questionnaire used a core set of questions to calculate household income based on production and prices of all crops and animals. IDE country programs add additional questions to the household-level survey based on specific program or project information needs. Additional questions include technology service adoption and use (e.g. use of manure, membership of marketing cooperatives, credit services); market outlets (buyers, location, knowledge); and/or social provisioning (expenditures on health care, education, etc). Each survey collects background information on age, social background, education, organization and development expectations. Data were entered, stored and processed in a customized MS Access-based application (www.mongi.org) that generated household income estimates based on an algorithm for processing the production and price data from the questionnaires. The data-set was exported to SPSS for reporting and statistical analysis.

Though the RBS was designed to be conducted by a team of contracted enumerators assisted by IDE, in most countries the first survey was implemented entirely by IDE field staff in order to acquaint them with the survey methodology. The 200 interviews took approximately 50 person-days per year (two weeks for five enumerators completing four surveys a day), with one-hour interviews of each respondent. We used Arshram (2009) to check for the minimal required sample size that detects income differences between (sub)groups. Based on the null hypothesis of no change and on the alternative hypothesis of a income change of 50 US$/household/year, well below the target impact (IDE, 2006), the minimum subgroup size resulted in 23. The sample size of 200 proved to be sufficient for IDE to disaggregate into subgroups, like districts and sex, and for detecting smaller than expected income impacts. Data entry and initial data clearance took an estimated of six weeks (8 questionnaires per day). Data cleaning on outliers was an iterative process that took more time than expected. Based on the year 1 experiences, a standard data cleaning procedure was implemented. Outliers were checked for consistency by IDE field staff and corrected, or removed before the final statistical analysis and reporting.

Household income impact

We tested the household income impact assessment methodology using the year 1 and year 2 survey data for Nepal. Nepal showed a high increase in average household income of 23.741 NPR (+51%). Currency fluctuations influenced the conversion to US$-figures[7]. Recalculated in US$-terms the yearly household incomes rose by 413 US$, representing an increase of 64% from the average 2006 starting point. The increase in PPP-terms was 871 PPP (+ 44%)[8]. More detailed analysis of changes in the income components showed that most of the increase was a result of price and yield increases. This underlined the need to control for the exogenous influence of price levels and climatic conditions. Applying the formula for calculating household income impact, this resulted in an average increase of 18,851 NPR (+40%), 302 US$ (+47%) or 726 PPP (+36%).

---

[7] The average US$ rate for each calendar year has been applied: 2006 = 72.05; 2007 = 66.38; 2008 = 69.66. Available on-line on http://www.oanda.com/convert/fxaverage.

[8] The Purchasing Power Parity rates are subject to periodic changes. PPP conversion allows national currencies to be compared on the basis of their purchasing powers free from differences in price levels across countries. We used the rates released in April 2009, http://www.imf.org/external/pubs/ft/weo/2009/01/weodata/index.aspx: 2006 = 23.48; 2007 = 24.63; 2008 = 25.98.

TABLE 1 AROUND HERE

The area-specific disaggregation of the income figures indicated important differences between the districts in average income impact. Applying the formula on the four districts where data from two surveys were available resulted in an average impact of 53% and 40% in the hilly regions of Kaski and Palpa respectively and attributable income increases of 27% and 66% in the Terai lowland areas of Kapilbastu and Rupandehi respectively.

Test on critical assumptions

Critical assumption 1: Recall gives reliable information

To test for an eventual recall bias we analyzed the correlation between the two measurements (expected to be highly significant as the data relates to the same reality) and tested for a structural difference between the measurements (searching for a recall bias to be controlled for). As expected, most variables showed a significant correlation between both measurements of farm income. However, the livestock income figures proved uncorrelated, indicating a serious problem of recall accuracy. Especially, the recall data on home consumption of livestock products and the accounting for animal losses seem to become unreliable. More important, however, are the large difference in the mean effects. There is an almost 50% overestimate in the longer recall observation compared with short recall observations of household production and income data. The major divergence resulted from differences in the reported income from target crops and off-farm income. We expect that a major part of this bias reflects changes in the questionnaire and interview process[9], and that the recall bias resulting from recall is smaller. However, we can deduce from our results that the methodology will have to continue monitoring for recall bias and incorporate a correction of bias due to long-term recall of production and price data.

TABEL 2 AROUND HERE

Critical assumption 2: Successive cohorts are comparable

To test the critical assumption that both cohorts had more or less the same characteristics, we explored the differences in background data between them with an independent-sample T-test (Table 2). The aggregate country data showed significant differences in several background variables The first cohort comprised more male household heads, with higher education level, in more remote areas.

[9] The data cleaning on the follow-up survey sample is still under way at the moment of writing this paper. The comparison is based on the rough data trimmed on outliers with 47 valid observations of a total of 56.

TABLE 3 AROUND HERE

While these differences in qualitative background characteristics were not problematic, the differences in income composition i(Table 4) indicated that the new cohort was more dedicated to irrigated horticulture than the first cohort, where traditional agriculture played a bigger role in their livelihood strategies. A disaggregation of the data to the four different districts involved in the project, showed that this change in cohort characteristics was especially manifest in the Kapilbastu district. In that district the income levels of cohort 2 could not be used to calculate impact, The three other districts showed difference in income composition between the cohorts that could be expected, with a higher income from horticultural crops for the cohort that applied the micro-irrigation equipment compared with the second cohort before adoption.

TABLE 4 AROUND HERE

Notwithstanding, the overall positive outcomes of the test on inter-cohort differences, it is clear that, when applying the RBS, a selection bias will have to be checked for.[10] Therefore, to limit the vulnerability of the RBS to these cohort differences, we modified the calculation formula for income impact. The income changes in first cohort is multiplied by index for exogenous influences $O_{1,r}/O_{3,r}$. The index is based on the district averages and, as a fixed ratio, is applied to correct individual responses for the influence of exogenous factors like weather and prices. The advantage of the index-based calculation is the fact that it enables disaggregation and subgroup comparisons. The coefficient is calculated as the average pre-adoption income in year A divided by the average pre-adoption income in year B. The formula proposed to calculate the average annual household income impact in year 1 is the following:

$$\Delta Y_{year1,impact} = (O_2 - O_{1,r}) \times (O_{1,r} / O_{3,r})$$

	Year 0		*Year 1*		*Year 2*		*Year 3*
Cohort 1	$O_{1,r}$	X	O_2				
Cohort 2			$O_{3,r}$	X	O_4		
Cohort 3					$O_{5,r}$	X	O_6

Discussion

The Nepal experience indicates the logistic feasibility of the rolling baseline survey instrument as an income measurement tool. The tool generates annual dat during the

[10] The same methodology and tests applied in Zambia detected important differences between cohorts, which motivated the modification of the calculation formula also for Nepal.

project period that can be used to adjust intervention strategies. The sample size of 200, used to trace income changes in clients, allowed subgroup disaggregation and is manageable and cost-effective. The retrospective pretest-posttest design is easy to grasp for non-experts (in contrast to most matching and regression designs). The availability of data for two years on each individual household make data control procedures more rigorous, as 'before' and 'after' situations of each respondent will have to show logic, and abnormalities and outliers can be detected and checked accordingly; something very difficult to do when only comparing test averages of different groups of respondents in different time-place contexts. However, the use of the resulting data for deriving attributable impact is less straightforward than assumed in the desig. In Nepal, the test on the recall bias pointed to an apparent overestimation of recalled household income. The large differences in reported income between the two surveys can be partly attributed to the changes in the survey questionnaire design and interview process, which was made more user-friendly after the first survey experiences[11]. A (forthcoming) second recall test may prove more reliable in detecting and calculating the magnitude of recall bias. If the recall bias proves manifest in the next tests, the formula needs to be amended to correct for this over- or sub-estimation of income as a result of longer recall periods. In that case, we propose to correct the income impact calculation formula with a with a recall coefficient ($O_{2a}/O_{2b,r}$), reflecting the size of the bias.

A more comprehensive and synthetic quasi-experimental design, including the recall control element can be written as follows:

		Year 0		*Year 1*		*Year 2*		*Year 3*
RBS	*NR*	$O_{1,r}$	*X*	O_2				
	NR			$O_{3,r}$	*X*	O_4		
	NR					$O_{5,r}$	*X*	O_6
Recall control	*NR*	$O_{1a,r}$	*X*	O_{2a}		O_{8a}		O_{9a}
				$O_{2b,r}$		$O_{8b,r}$		$O_{9b,r}$

The Nepal experience indicates also that the way to translate the registered income changes into income impact, correcting for exogenous factors, proved to be problematic when important differences in start-off conditions exist between the cohorts, as happened in Kapilbastu. However, considering the overall results of the test on inter-cohort differences, we still think that successive client cohort data is useful to construct a credible counterfactual. The indirect use of the inter-cohort income data through the index, instead of using direct subtraction, makes the method more robust to cope with inter-cohort differences.

The rolling baseline methodology is applicable in other development settings as a credible assessment tool, when several conditions are met. First, the tool assesses 'first-year-after-adoption'-impact. This assumes that impact of technology manifests itself in the first year after adoption. Increasing the time interval for measuring two- or three-year impact introduces even more serious recall issues. The solution of surveying the respondents in two distinct opportunities, once directly after the purchase and later a year after having used it, seems a good option to prevent the recall bias, but presents a logistical challenge and needs sufficient budget. In the near future, cell phone interviews may be a cheap option to increase the efficiency of repeated data collection in one-and-the-same respondent.

[11] It took between 2 and 3 hours to register the production data on all crops in the extremely diversified farms in Nepal. This affected the accuracy of data-collection. In the second round of surveys the detail on inputs used was streamlines, lumping then together instead of detailing them per crop.

Second, the interventions being tested must have a plausible causal pathway to support the attribution of impact (Hailey and Sorgenfrei, 2004). Many bottom-of-the-pyramid ventures sell technologies that are necessary but not sufficient to generate changes in the agricultural system and related income streams. In the case of irrigation technologies, the causal connection seems quite obvious. In agricultural input supplies, like pest control, quality seed provisioning or new fertilizer formulas, the constellation of factors that will produce increasing crop incomes is more complex and intertwined. Attribution of poverty reducing effects will depend much more on information collected on indirect indicators, like sales or use of the technologies, and on qualitative ‘causal process observations’ (Brady, Collier et al., 2006) of the resulting change processes in household livelihood strategies; survey-based direct income measurements alone are useful but insufficient for causal inferences of poverty impact.

Conclusions

The promise of poverty alleviation as an important additional impact of Bottom-of-the-Pyramid strategies, coupled with an increase in private and public development support to co-finance of establishing of BoP-ventures, increases the demand for impact assessments. The Donor Committee on Enterprise Development proposes minimum standards for such reporting. This creates an urgent need for instruments that can give credible figures on changes in the customers poverty levels. However, the ways to measure these changes and translate these in attributable impacts is not straightforward and needs careful design. Impact attribution is especially difficult in BoP ventures that target poor households with products that are not directly consumed but incorporated in strategies of the poor as producers (London, Anupindi et al. 2009). These items become part of a wider set of factors that generate the poverty impact. Many traditional methods for impact evaluation are not suited for monitoring impacts in self-selecting client groups. The rolling baseline survey methodology is an approach designed to add a much-needed tool to the quasi-experimental tool box. With the careful control for recall bias and inter-cohort differences, the tool is useful to measure outcomes for impact evaluation. The sampling from a companies’ client registration system, the comparison of the household income before and after adoption, and the application on an index to reflect autonomous income change, together, generate poverty impact figures of BoP technology sales with increased validity and credibility.

References

Arshram, H. 2009. Statistical Thinking for Managerial Decisions: Sample Size Determination. http://tinyurl.com/nq5kgr [30 July, 2009]

Bamberger, M., J. Rugh, J., Mabry, L. 2006. RealWorld evaluation: working under budget, time, data, and political constraints, Sage Publications Inc.

Brady, H. E., Collier, D., Seawright, J. 2006. "Toward a Pluralistic Vision of Methodology." Political Analysis **14**(3): 353-368.

Chelimsky, E., Shadish, W.R. 1997. Evaluation for the 21st century: A handbook, Sage.

DCED. 2008. Quantifying Achievements in Private Sector Development: Control Points and Compliance Criteria, Donor Committee for Enterprise Development.

Danse, M., S. Vellema, Plant, A. 2005. "Small-scale Farmer Access to International Agri-food Chains: A BOP-Based Reflection on the Need for Socially Embedded Innovation in the Coffee and Flower Sector." Greener Management International **51**: 39.

Dehejia, R. H., Wahba, S. 2002. "Propensity score-matching methods for nonexperimental causal studies." Review of Economics and Statistics **84**(1): 151-161.

Duflo, E., Glennerster, R., Kremer, M., Center, L., Floor, T. 2006. "Using Ramdomization in Development Economics Research: A Tool Kit." NBER Working Paper.

Hailey, J., M. Sorgenfrei , M. 2004. Measuring Success: issues in performance measurement. Oxford, INTRAC.

Heckman, J. J. 2005. "The scientific model of causality." Sociological Methodology **35**(1): 1-98.

Heierli, U., Polak, P. 2000. "The Market Creation Approach to Development." Swiss Agency for Development and Cooperation (SDC). http://www.intercooperation.ch/sed/product/heierli/main.html. [30 August 2009]

IDE. 2006. A Path Out of Poverty: Connecting Dollar-a-Day Farmers to Afforable Small Plot Irrigation and Markets. Denver, International Development Enterprises.

Kandachar, P., Halme, M. 2008. Sustainability Challenges and Solutions at the Base of the Pyramid, Greenleaf Publishing.

Landrum, N. E. 2007. "Advancing the "Base of the Pyramid" Debate." Strategic Management Review **1**(1).

London, T., R. Anupindi, Sheth, S. 2009. "Creating mutual value: Lessons learned from ventures serving base of the pyramid producers." Journal of Business Research. In press [Online: 19 July]

London, T., Hart, S. 2004. "Reinventing strategies for emerging markets: beyond the transnational model." Journal of International Business Studies **1**(2).

Polak, P. 2008. Out of poverty: what works when traditional approaches fail, Berrett-Koehler.

Ravallion, M. 2009. "Should the Randomistas Rule?" The Economists' Voice **6**(2): 6.

Shadish, W. R., Cook, T.D., Campbell, D.T. 2002. Experimental and Quasi-Experimental Designs for Generalized Causal Inference, Houghton Mifflin Co. Boston, MA.

Simanis, E., S. Hart, S, Enk, G., Duke, D., Gordon, M., Lippert, A. 2004. Strategic Initiatives at the Base of the Pyramid: A Protocol for Mutual Value Creation. Rancine, Base of the Pyramid Protocol™ Workshop Group,October 19-22.

Simanis, E.,Hart, S. 2008. The Base of the Pyramid Protocol: Toward Next Generation BoP Strategy, second edition, Center for Sustainable Global Enterprise, Cornell University.

Tables

Table 1: First year household income impact in IDE-Nepal RPI-areas

RPI-area	$O_{1,r}$ NPR (2006)	O_2 NPR (2007)	$O_{3,r}$ NPR (2007)	income change 2006-2007	%	income impact 2007	%
NEPAL	**46,688**	**70,429**	**58,800**	**23,741**	**51%**	**18,851**	**40%**
Kaski	65,560	99,973	74,478	34,413	52%	30,292	46%
Palpa	66,312	79,108	45,526	12,796	19%	18,638	28%
Kapilbastu	20,705	35,867	61,168	15,162	73%	5,132	25%
Rupandehi	48,637	80,288	54,134	31,651	65%	28,437	58%

Table 2 Control on recall bias comparing short and long recall observations on the same household characteristics (2007)

		Mean	N	Std. Error Mean	Significance of correlation	Significance of difference (2-tailed)
Pair 1	survey 1 Age household head	48,9	47	2,14	0.000	.016
	Survey 2 Age household head	44,2	47	1,66		
Pair 2	survey 1 Total area (m2)	7823	47	755	0.000	.675
	survey 2 Total area (m2)	8161	47	982		
Pair 4	survey 1 Household income (NPR)	$51,572	47	$8,539	0.000	.008
	survey 2 Household income (NPR)	$76,284	47	$9,800		
Pair 5	survey 1 Crop income (NPR)	$32,799	47	$5,972	0.004	.472
	survey 2 Crop income (NPR)	$36,959	47	$4,251		
Pair 6	survey 1 Target-crop income (NPR)	$4,908	47	$1,511	0.037	.000
	survey 2 Target-crop income (NPR)	$19,186	47	$2,531		
Pair 7	survey 1 Livestock income (NPR)	$-5,098	47	$4,478	0.327	.153
	survey 2 Livestock income (NPR)	$3,584	47	$4,668		
Pair 8	survey 1 Off-farm income (NPR)	$23,872	47	$4,935	0.014	.098
	survey 2 Off-farm income (NPR)	$35,739	47	$7,065		

Table 3: Significant differences in background characteristics of cohort 1 compared with cohort 2 (2007)

	Period of analysis (2007)	N	Mean	Std. Error Mean	Significance of difference (2-tailed)
Gender of household head (male=1; female=2)	cohort 1	178	1,06	,018	0.000
	cohort 2	193	1,20	,029	
Age of household head [yrs]	cohort 1	178	46,8	1,02	0.548
	cohort 2	193	47,7	1,06	
Education level household head [years]	cohort 1	178	3,09	,276	0.000
	cohort 2	193	1,78	,174	
Distance to main market [km]	cohort 1	178	31,73	8,25	0.001
	cohort 2	193	4,26	,65	
Total area (m2)	cohort 1	178	7239	388	0.808
	cohort 2	193	7111	355	
Total number household members	cohort 1	178	8,53	,30	0.771
	cohort 2	193	8,40	,30	
No. of children per woman in reproductive age	cohort 1	170	1,96	,11	0.321
	cohort 2	193	1,82	,10	

Table 4: District disaggregated analysis of income characteristics of cohort 1 (posttest) compared with cohort 2 (pretest) in 2007

RPI Intervention District		Period of analysis (2007)	N	Mean	Std. Error mean	Significance of difference (2-tailed)
Kaski	Off Farm Income total (NPR)(2007)	cohort 1	28	56,821	12,846	.490
		cohort 2	28	43,616	13,993	
	Livestock income (NPR)(2007)	cohort 1	28	8,586	5,383	.612
		cohort 2	28	12,665	5,912	
	Non-target crop income(NPR)(2007)	cohort 1	28	25,053	5,643	.133
		cohort 2	28	15,399	2,800	
	Target crop income (NPR)(2007)	cohort 1	28	9,510	2,039	.006
		cohort 2	28	2,796	1,091	
	Household income ($)(2007)	cohort 1	28	1,506	245	.270
		cohort 2	28	1,122	242	
Palpa	Off Farm Income total (NPR)(2007)	cohort 1	34	29,244	6,938	.135
		cohort 2	31	16,471	4,778	
	Livestock income (NPR)(2007)	cohort 1	34	-6,430	6,289	.141
		cohort 2	31	4,263	3,394	
	Non-target crop income(NPR)(2007)	cohort 1	34	46,458	22,412	.281
		cohort 2	31	20,726	2,798	
	Target crop income (NPR)(2007)	cohort 1	34	9,836	1,607	.027
		cohort 2	31	4,064	2,008	
	Household income ($)(2007)	cohort 1	34	1,192	330	.149
		cohort 2	31	686	95	
Kapilbastu	Off Farm Income total (NPR)(2007)	cohort 1	51	26,773	12,101	.997
		cohort 2	49	26,823	5,322	
	Livestock income (NPR)(2007)	cohort 1	51	-6,263	2,237	.746
		cohort 2	49	-4,766	4,017	
	Non-target crop income(NPR)(2007)	cohort 1	51	18,251	5,753	.765
		cohort 2	49	20,861	6,545	
	Target crop income (NPR)(2007)	cohort 1	51	-2,895	538	.000
		cohort 2	49	11,216	1,693	
	Household income ($)(2007)	cohort 1	51	540	202	.272
		cohort 2	49	816	143	
Rupandehi	Off Farm Income total (2007)	cohort 1	65	23,010	4,046	.706
		cohort 2	85	21,047	3,324	
	Livestock income (NPR)(2007)	cohort 1	65	-3,147	2,253	.690
		cohort 2	85	-1,787	2,421	
	Non-target crop income(NPR)(2007)	cohort 1	65	48,586	6,375	.005
		cohort 2	85	28,362	3,073	
	Target crop income (NPR)(2007)	cohort 1	65	1,1839	1,720	.518
		cohort 2	85	1,3546	1,891	
	Household income ($)(2007)	cohort 1	65	1210	117	.052
		cohort 2	85	921	89	