

The Impacts of USAID-Supported Agricultural Programs: Methodological Lessons Learned

James F Oehmke
Charles B Moss
John Hoddinott
Melinda Smale
Athur Mabiso
Neha Kumar¹

Prepared for USAID/AFR/SD and USAID/BFS

Sep. 29, 2011
Revised Oct. 11, 2011

¹ Consultant, University of Florida, International Food Policy Research Institute (IFPRI), Michigan State University, IFPRI, and IFPRI, respectively. For questions and comments please contact Prof. James F Oehmke, 909 Harrington Lane, East Lansing, MI 48823 USA; oehmkej@gmail.com. The authors would like to thank the large number of individuals who helped conduct the individual impact studies, and Juan Belt, Ron Greenberg, Jeff Hill, Susan Thompson and Jerome Wolgin for useful comments and discussion. The views presented in this paper represent the views of the authors and do not necessarily represent the views of USAID or the United States Government or of the authors' home institutions. The responsibility for all errors remains with the authors.

Introduction

In FY 2010 the United States Agency for International Development (USAID), Bureau for Africa (AFR), Office of Sustainable Development (SD) commissioned a set of studies assessing the impact of USAID African agricultural projects. The initial plan was to carry out impact assessments of the following activities/programs:

- East Africa Regional: intra-regional maize trade facilitation activities, Regional Agricultural Trade Expansions Support (RATES) program and Competitiveness and Trade Expansion (COMPETE) program
- Ethiopia: food and income support activities, Productive Safety Nets Programme (PSNP) and Household Asset Building Program (HABP)
- Ghana: pineapple and mango value chain enhancement activities, Trade and Investment Program for a Competitive Export Economy (TIPCEE)
- Kenya: dairy production and value chain development activities, Kenya Dairy Development Program (KDDP) and Kenya Dairy Sector Competitiveness Program (KDSCP); horticulture production and value chain development activities, Kenya Horticulture Development Program (KHDP); and maize production and value chain development activities, Kenya Maize Development Program (KMDP).
- Rwanda: coffee value chain development activities, Partnership for Enhancing Agriculture in Rwanda through Linkages (PEARL I & II) and Sustainable Partnership to Enhance Rural Enterprise and Agribusiness (SPREAD)
- Zambia: beef cattle value chain development activities, Production, Finance, and Improved Technology (PROFIT).

The studies were selected on the basis of mission interest, likely data availability, and indications of project success or of lessons learned—there was a deliberate effort to examine likely success stories that might be scalable under Feed the Future (FtF), and to examine key steps in the causal pathways from project activity to poverty reduction. The exception to the likely project success was Ghanaian pineapple, selected because the outgrower pineapple scheme provided a contrast to the Ghanaian mango producer-group structure and therefore potential lessons learned even though the pineapple activity had not achieved all its intended targets. The Zambia assessment was stymied by lack of project information and a change in mission interest.

The objectives of the impact studies were:

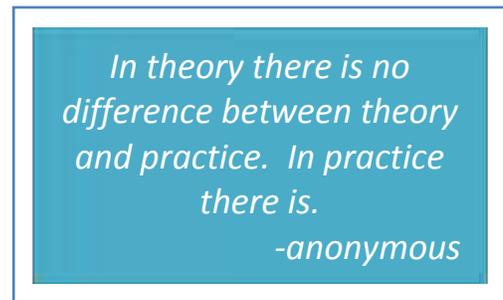
- 1) To quantify the effect of value chain improvements on smallholder income and poverty status or child nutritional status;
- 2) To provide empirical validation/falsification of the causal pathways from intervention to poverty reduction, by which the program expects to operate; and
- 3) To learn lessons about what has made the value chain improvements most successful in augmenting smallholder income, particularly with respect to new programs to be funded under FTF.

The objective of this policy brief is to summarize the methodological lessons learned from these impact studies for practical quantification and attribution of impact to USAID projects in sub-Saharan Africa. There are large and growing literatures on the meaning and quantification of poverty, and on project evaluation, and these literatures fed into the individual project designs. However, this brief focuses on

practical lessons learned in the impact assessments of the selected projects (the assessment studies are listed in the bibliography).

The impact studies used quasi-experimental modeling methods with differences-in-differences based attribution of impact to USAID supported projects. The Ghana studies were exceptions due to insufficient baseline data; they relied on changes in gross margins over time (pineapple) and livelihoods descriptions (mango). Statistical specifics depend on the specific data used in the individual studies. Studies also employed complementary methods to address questions posed by USAID missions related to their FtF programming needs. A companion brief summarizes the quantitative results of the studies.

The remainder of this brief draws out the specific lessons learned.



Lessons Learned

It is possible to quantify income, poverty and nutritional status

Each of the detailed quantitative studies was able to provide empirical evidence on at least one of these indicators (East Africa regional, Ethiopia, Kenya, Rwanda). This evidence was drawn wholly (East Africa, Ethiopia, Kenya) or in part (Rwanda) from primary data collection designed to provide information relevant to USAID decision making.

It is recommended that USAID support the empirical measurement of impact indicators.

It is possible credibly to quantify changes in income and poverty attributable to USAID projects

The quantitative studies relied on quasi-experimental methods and differences-in-difference approaches to attribute change in impact indicators to USAID supported projects. The fundamental comparison is the increase in the impact indicator (e.g. measured household income) among beneficiaries from the start to the end of the project compared to the increase in the impact indicator among a non-beneficiary comparison group. The Ethiopia, Ghana, Kenya and Rwanda studies analyzed the data using multiple techniques to determine the robustness of the attribution, and all found that statistical results were reasonably robust and contained useful information. The Kenya and Ethiopia studies accounted for initial differences between treatment and comparison groups; the Rwanda study used non-parametric statistics to corroborate results from parametric estimation; the Kenya study used non-parametric methods in the post-estimation comparison of income distributions; the Ethiopia and Rwanda studies have qualitative livelihoods data (collected independently from the income data) that corroborate the quantitative findings from the income data; the Ethiopia study has both self-reported (food gap) and anthropometric (height, weight and age) data that will be used to triangulate findings.

It is recommended that empirical measurement of impact indicators be of sufficient quality to support statistical attribution of changes in impact indicators to USAID programs.

It is recommended that empirical quantification of impact indicators be supplemented with additional quantitative and/or qualitative information sufficient to corroborate (or invalidate) the quantified changes in impact indicators attributed to USAID projects.

It is critically important to have household data on impact indicators

The three impact indicators relevant to MDG1 targets and FtF higher-level goals are income, poverty, and child underweight. Each of these is most accurately measured through primary data collection at the household level. Changes in income and poverty are directly measurable at the household level. Although the assessments measured changes in income and poverty status attributable to USAID supported projects at the household level, for analysis and reporting these changes were aggregated through analysis of changes either in mean household income or in poverty rates. These changes were presented in terms of pre-specified subsamples (e.g. districts, as in Rwanda) and/or through econometric analysis that quantified mean project effects (Kenya, Rwanda; Ethiopia for the food gap).

It is recommended that the empirical quantification of impact indicators rely on household data.

It is critically important to have good baseline data

The comparison of impact indicators before and after the project relies on baseline data. The Ghana pineapple analysis is the weakest of the analyses in part because the most relevant baseline data available were gross and net margins; baseline data on income or nutritional status were not available. It is likely not necessary to have baseline data collected prior to the onset of the USAID project: in Ethiopia data collection began in the first year of the project and yields useful results. Literature review showed that slightly less than one-half of the evaluation studies reviewed did not have baseline data; of those with baseline data the quality varied, and none had nutritional measures. Related to this is the need for baselines with knowledgeably defined beneficiary groups and expertly designed statistical sampling frames.

It is recommended that all projects designed to generate high-level impact (some may be designed to improve the enabling environment in support of other activities, e.g.) should construct or cause to be constructed a baseline data set.

It is critically important to have an appropriate comparison group

The use of a comparison group helps to control for external factors that affect the entire population in terms of income growth, poverty reduction or hunger reduction, such as drought. The East Africa, Ethiopia, Kenya and Rwanda studies all rely on a comparison group against which the progress of the treatment group can be measured. The difference-in-difference approach requires construction of an appropriate comparison group that, at the baseline, is similar to the treatment group. Very few of the other project evaluations in the literature review had a comparison group. In Kenya, despite the evaluation designed specifically to include treatment and comparison groups, the evaluators felt that additional information was gained by comparing the treatment group to smallholders in a nationally representative sample.

It is recommended that all projects designed to generate high-level impact (some may be designed to improve the ambient environment in support of other activities, e.g.) should specify or cause to be specified treatment and comparison groups for evaluation and measurement purposes.

Longitudinal (panel) data sets are very useful for impact assessment

Ultimately the quantification of impact is designed to measure changes in household status attributable to the USAID supported projects, and there are no better data for this purpose than longitudinal household data. The ability to identify households in some way allows for statistical control of household idiosyncrasies and improves understanding of the household's emergence from poverty. It

allows for specific disaggregation to address specific interests, and sometimes allows combination with other data sets. Confidentiality issues are a concern, especially if using secondary data such as a national survey. Large sample sizes are appropriate especially when it is desirable to have results disaggregated by project activity or sub-activity, beneficiary type (e.g. poor v. non-poor), gender, agro-climatic conditions, etc. In Kenya a sample size of around 600 households for three USAID-supported activities led to small cell counts when data were disaggregated by activity, treatment v comparison group and poverty status. In Ethiopia a large sample size was required to understand the relationship between the intensity of the project activity and the level of impact.

It is recommended that USAID baseline surveys be designed for follow-on surveys that generate a longitudinal household data set. Sample sizes should be calculated to provide a full range of analytical opportunities for both impact indicators and if possible causal pathways to the impact. To improve the credibility of the data, response rates should be reported according to American Association of Public Opinion Research guidelines, and sample validity and reliability statistics should be calculated.

National agricultural surveys may be able to provide useful baseline and/or end of project datasets

The National Institute of Statistics Rwanda (NISR) conducted household surveys including income (expenditure) information in 2000, 2005 and 2010, which coincidentally corresponded to the starting points of PEARL and SPREAD and the ending point of SPREAD, respectively. Upon request NISR provided 2000 and 2005 data; 2010 data are not yet available. The Rwanda study collected primary data for 2010 in order to complete the assessment in a timely fashion; the NSIR provided information on sampling frames and survey instruments that facilitated the primary data collection. The availability of prior data enabled this particular analysis, and more generally would lower the cost of impact assessments. In contrast, the Ethiopian and Kenyan analyses relied on primary data collections. Project-specific primary data collection is an effective but at times costly approach. Kenya is considering nesting the treatment group for USAID programs within the overall sampling design for the nationally representative survey, providing unique project information but at a lower cost.

It is recommended that impact assessments explore the possibility of using national household survey data if available and suitable.

The Demographic and Health Surveys (DHS) data are not easily accessible at the level of disaggregation necessary for quantifying the effect of programs on child nutritional status.

ReSAKKS, conducting the East Africa regional study of the impacts of trade flows, tried to access the Kenya DHS data in order to quantify impact on child nutritional status as quantified by DHS data on child underweight. The working hypothesis was that smallholders living in food deficit regions and in propinquity to a trade route would realize improved child nutrition status as staple food (maize) trade flows increased. A treatment-comparison contrast was defined by geographic distance to a trade route. The Kenya DHS data contain global positioning information for each household—mildly offset to preserve confidentiality—that likely would have served to delineate treatment and comparison households for a difference-in-difference analysis of maize trade flows and child nutritional status. In March, 2011 ReSAKKS began a process for access to disaggregated DHS data including formal requests, but to date has been unable to access the disaggregated data and it is unclear if access will ultimately be granted.

It is recommended that AFR and BFS request a review of DHS data access procedures to insure that the data are used and useful in assessing the impact of USAID agricultural projects on child nutritional status while preserving the confidential nature of the data.

Project activity reporting and accounting needs to be compatible with impact assessment needs

The TIPCEE project in Ghana had multiple activities ongoing, making it difficult to determine how much was spent on pineapple or mango v. other activities. E.g., it was difficult to know how much of a project scientist's time was spent on any particular crop, and similarly for operational expenses such as project vehicles or offices. In the halted assessment of PROFIT, it was difficult to determine even which districts PROFIT worked in during any given year. Thus it became difficult to determine which districts might have benefitted from project activities, and which might not. In Kenya each commodity value chain had its own project, which made it relatively easy to determine costs spent on dairy, e.g. However, there remain issues especially in the horticulture and maize activities in understanding which project beneficiaries received which services and adopted which innovations, and therefore it is difficult to understand which innovations generated the largest impact and why. E.g., some projects provide a field demonstration to many farmers, but work in a hands-on fashion with a smaller number. Some smallholders participate in the program in some years but not others. In Ethiopia the level of impact depended non-linearly on the intensity of contact, so that treating all beneficiaries similarly would have been inaccurate. It is generally expected that the different levels of engagement will generate different impacts at different costs, so that disaggregation of benefits and costs may be desirable.

It is recommended that projects provide reasonable estimates of costs and beneficiaries disaggregated by the type and level of project activity as specified in a causal pathway or logical framework. OMB circulars A-21 and A-87 offer suggestions on time and effort and indirect cost reporting that the implementing partner might apply even if not contractually required to do so.

The selection of method is less important than conscientious application of the method selected

There is currently debate about the use of Randomized Clinical Trial (RCT) methods v. Quasi-Experimental Model (QEM) methods. It is possible to generate pages on the advantages of either method over the other. The QEM methods were selected for the impact studies for practical reasons: RCTs are expensive, require clinical-type control over subject behaviors and actions, in development application are rarely completed in a year, typically do not account for uncontrolled or unforeseen external factors that may influence results, and face difficulties generating out-of-clinic results possibly including those that would inform scale-up of programs under FtF; the standard of solid and credible but not necessarily bullet-proof attribution of impacts to projects also influenced selection of QEMs v. RCTs. Further, use of participatory methods may preclude use of RCTs, as in Ethiopia where the local community selected the project beneficiaries.

Similar statements can be made about specific statistical and econometric tools: the key is in the appropriateness of the method selected to the evaluation at hand and the quality of the application.

It is recommended that USAID impact assessments select the best method(s) for the questions at hand, including those generated by missions and project managers. It is possible and even likely that multiple methods will be applied simultaneously within a single general framework.

It is recommended that USAID monitor the quality of the impact assessments conducted.

The definition of poverty can have an important influence on results

Project impact as quantified by poverty reduction may be influenced by the choice of poverty line, even in a difference-in-difference comparison between treatment and comparison groups. In 2000 Rwanda was very poor: even doubling the average smallholders' income left many short of the \$1.25 standard.

However, based on the rural poverty line (which is accepted by the United Nations (UN) for MDG1 measurement purposes in Rwanda) the income increase attributed to SPREAD is sufficient to generate a large and statistically significant impact on poverty—a 14.3 percentage point decline in the poverty rate. The converse is true in Kenya. The poverty gap in Kenya at \$1.25 is relatively small, so that relatively small increases in income could help a household emerge from poverty. Regression estimates of poverty declines showed large differences in household income changes but only a modest difference in the change in poverty rate between treatment and comparison groups—three percentage points—because even the modest income increases in comparison group income were sufficient to pull many of those households above the \$1.25 poverty line. However, the rural poverty line in Kenya is probably closer to \$2.00 (depending on exchange rates used etc.), and at \$2.00 the treatment group decrease in poverty is eight percentage points larger than the decrease in the comparison group poverty rate.

It is recommended that assessment of headcount poverty reduction attributable to a USAID-supported project calculate the poverty reduction at the UN-specified MDG1 poverty line (usually \$1.25) and at one or two other poverty lines relevant to the country and project circumstances.

The method for quantifying income is important.

The most commonly used primary-data method for quantifying income is to ask respondents about expenditures and infer income. This approach is viewed as less invasive than asking directly about income and therefore more reliable. It was applied in the Ethiopia and Rwanda data collections. The Rwanda data set showed an abnormally low reported figure for home consumption; the Ethiopia figure has not yet been analyzed. In Kenya, smallholders were asked directly to quantify income from various sources, which provides information not just on income but on the role of farm income and income diversification in households emerging from poverty. However, the sample has no reported tests of validity or reliability, nor are there expenditure data available for triangulation. Consequently the quality of the income variable in each of the datasets may be very good, but is unknown.

It is recommended that further investigation be conducted into the type and quality of information contained in different approaches to smallholder household income measurement.

The frequency of impact assessment depends on the projects evaluated

The Ethiopia PSNP analysis is conducted every second year including primary data collection, the Kenya USAID impact indicator data set is updated every second year, and the Rwanda assessment was based on data collected at five year intervals. In each case the data were sufficient for impact assessment, and the timing was appropriate for project evaluation and management needs.

Methodologically there is an advantage to collecting at least three rounds of longitudinal data for the final evaluation, e.g. baseline, mid-term and end-of-project. The first two rounds provide not just a baseline ‘snapshot’ of where households are at the beginning of the project, but together the baseline and midterm data provide an initial picture of the household income and poverty trajectories. This allows for more robust estimation of the differences-in-differences estimator of project impact. Additional data could further improve estimation quality and comprehension of causal pathways.

It is recommended that impact data be gathered at the project baseline, mid-term, and end-of project. A preliminary impact assessment can be carried out at project midterm if desired, in part to quantify impact but also to show movement along the causal pathway from project activity to impact. A full impact assessment should be conducted upon project completion. An additional assessment could be

conducted to address longer term impact and sustainability issues 3-5 five years after project completion.

There are advantages to third party impact assessments

All of the impact assessments were conducted by third parties, usually local partners. It takes skill, time and effort to implement a high-quality impact assessment. Advantages of third-party assessment teams were that specialists are often more knowledgeable about impact assessment than were the project implementers, that project management did not have to manage impact assessments in addition to programmatic activities, and that third party assessment was generally viewed as more objective. A disadvantage of the third-party approach was that organizations that specialize in impact assessment are often less knowledgeable about the agricultural and value chain specifics than are the project implementers. Operational detail is both programmatically informative within the impact assessment, and informs the way in which the impact assessment is designed and conducted. Having third parties work closely with programming partners to understand operational detail can be very important.

It is recommended that impact assessments generally be conducted by third parties with experience in impact assessment; if possible, these third parties could be local parties with knowledge of local agricultural value chains.

Spillovers and indirect effects may be very important

The agricultural economic literature suggests that the indirect effects of smallholder development through spillover of agricultural techniques to other smallholders, increased demand for labor throughout the commodity value chain, and increased local consumption leading to rural non-farm employment may be as large as or larger than the direct effects of agricultural technology on smallholder incomes. The Kenya study is the only impact assessment to address this; the findings are consistent with an indirect income effect comparable in magnitude to the direct effect.

It is recommended that future impact assessments consider and if appropriate quantify the indirect effects of smallholder agricultural growth on income increases and poverty reduction.

The method for quantifying yields and prices is important

Yields were not an impact indicator in any study, but were an intermediate indicator in the Ethiopia, Ghana, Kenya and Rwanda analyses. The Kenya study found problems with crop yields measured by farmer recall, even though the data collection was usually right after the main harvest. Crop cut data is generally regarded as more reliable than farmer recall. However, crop cuts usually occur during harvest; socio-economic surveys are generally post-harvest surveys because farmers are too busy to spend time being interviewed or filling out surveys during harvest. Due to the large number of external influences on yields and the ability to control many external influences in field trials, it may be useful to quantify project effects on crop yields in a randomized trial design. On a broader scale, dairy yield data may need to be collected from daily farmer records; staple crop yields will be measured once or twice a year at harvest.

Similarly, it is difficult to quantify prices received and the influence of projects on prices received because of the nature of pricing mechanisms. Prices received are influenced by the timing of the sale, the quality of the output, the amount that the smallholder is willing to sell, the number of buyers in the village or market, prior relationships the smallholder might have with buyers, cooperative membership, smallholder bargaining acumen, etc. Smallholders may also sell product in a set of small amounts when they need cash, receiving a different price each time. Particularly for value chain projects seeking to

improve smallholder value, it is important to have price data that are capable of delineating the effect of the project on prices.

It is recommended that further investigation be conducted into the most appropriate methods for quantifying smallholder yields and prices received in the context of impact assessment.

Bibliography of SD Impact Assessment Studies

East Africa

Karugia, J., et al., *An assessment of the impact of cross-border trade in maize on welfare in the East Africa region*. 2011: Regional Strategic Analysis and Knowledge Support System, Eastern and Central Africa: Nairobi, Kenya.

Ethiopia

Berhane, G., et al., *The impact of Ethiopia's Productive Safety Nets and Household Asset Building Programme: 2006-2010*. 2011, International Food Policy Research Institute.

Ghana

Mensah, A.O., *Mango farmer benefit from the Trade and Investment Program for a Competitive Export Economy (TIPCEE) project*. 2011, University of Cape Coast: Cape Coast, Ghana.

Oehmke, J.F., *TIPCEE: Impact overview and pineapple assessment*. 2011, USAID: Accra, Ghana.

Kenya

Oehmke, J.F., et al., *Impacts of USAID/Kenya supported Agricultural Productivity Interventions on Household Income and Poverty Reduction*, in *WPS 38/2010*. 2010, Tegemeo Institute of Agricultural Policy and Development: Nairobi, Kenya.

Smale, M., M. Mathenge, T.S. Jayne, E. Magalhaes, J. Olwande, L. Kiriimi, M. Kamau, and J. Githuku, *Income and poverty impacts of USAID-funded programs to promote maize, horticulture and dairy enterprises in Kenya, 2004-2010*. 2011: Working Paper (in review) Tegemeo Institute of Agricultural Policy and Development, and Michigan State University: Nairobi, Kenya and East Lansing, MI.

Rwanda

Oehmke, J.F., et al., *The impact of USAID investment on sustainable poverty reduction among Rwandan smallholder coffee producers: A synthesis of findings*. 2011: Kigali, Rwanda.

Bihogo, E., D.D. Weatherspoon, and J.F. Oehmke, *The impact of PEARL and SPREAD projects on the development of the Rwandan coffee sector*. 2011, USAID: Kigali.

Moss, C.B., J.F. Oehmke, and A. Lyambabaje, *An Economic Evaluation of SPREAD Impact on Rwanda's Rural Population*. 2011 (forthcoming): Kigali.

Literature

Katjioungua, H., *A Review of Evaluated Projects*. 2011: Washington, DC.