USAID

FROM THE AMERICAN PEOPLE

POLICY BRIEF

IMPACTS OF USAID-SUPPORTED AGRICULTURAL PROGRAMS

METHODOLOGICAL LESSONS LEARNED

James F. Oehmke, Charles B. Moss, John Hoddinott, Melinda Smale, Athur Mabiso, and Neha Kumar

INTRODUCTION

In 2010, the Agency commissioned studies assessing the impact of the following Initiative to End Hunger in Africa and Global Food Security Response projects.

- East Africa Regional (intraregional maize trade facilitation activities): Regional Agricultural Trade Expansions Support (RATES) and Competitiveness and Trade Expansion (COMPETE) programs

- 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)

- Kenya (horticulture production and value chain development activities): Kenya Horticulture Development Program (KHDP)

- Kenya (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)

The studies were conducted by the Regional Strategic Analysis and Knowledge Support System and the Tegemeo Institute for Agricultural Policy and Development (East Africa study); the International Food Policy Research Institute (Ethiopia study); the Monitoring, Evaluation, and Technical Support Services Unit of the University of Cape Coast, Ghana (Ghana studies); the Tegemeo Institute (Kenya studies); and the

National University of Rwanda (Rwanda study). The projects were selected on the basis of mission interest, likely data availability, and indications of project success or lessons learned. There was a deliberate effort to examine likely success stories that might be scaled up under Feed the Future (FTF) and to examine key steps in the causal pathways from project activity to poverty reduction.

The objectives of the impact studies were

- to quantify the effect of USAID-supported projects on smallholder income and poverty status or child nutritional status;

- to provide empirical validation or falsification of the causal pathways from intervention to poverty reduction, by which the projects operate; and

- to learn lessons about what has made the projects most successful in augmenting smallholder income, particularly with respect to new activities to be funded under FTF.

The impact studies (listed in the bibliography) used quasi-experimental modeling methods with difference-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 livelihood descriptions (mango). Statistical specifics depend on the data used in the individual studies. Studies also used complementary methods to address questions posed by USAID missions related to their FTF programming needs.

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, which fed into the individual project designs. This brief focuses on the specific practical lessons learned in the selected projects' impact assessments. Two companion briefs describe the cost-effectiveness of these USAID projects and programmatic lessons learned.

LESSONS LEARNED AND RECOMMENDATIONS

It is possible to quantify income, poverty, and nutritional status. The detailed quantitative studies (for the East Africa region, Ethiopia, Kenya, and Rwanda) provided empirical evidence on at least one of the following indicators: income, poverty, and nutritional status. This evidence was drawn wholly (in the cases of East Africa, Ethiopia, and Kenya) or in part (in Rwanda) from primary data, which were collected in a manner designed to provide information relevant to USAID decisionmaking.

Recommendation: USAID should support the empirical measurement of impact indicators.

It is possible to quantify changes in income and poverty that can be attributed to USAID projects. The quantitative studies relied on quasi-experimental methods and difference-in-differences approaches in order to attribute changes in impact indicators to USAID-supported projects. The increase in a particular impact indicator—for example, measured household income—for beneficiaries (from the beginning of the project to its completion) was compared with the increase in that same indicator for nonbeneficiaries in a comparison group. Using multiple techniques to determine the robustness of the attribution, analysis of data from the Ethiopia, Ghana, Kenya, and Rwanda studies showed reasonably robust statistical results that contained useful information. The Kenya and Ethiopia studies accounted for initial differences between treatment and comparison groups; the Rwanda study used nonparametric statistics to corroborate results from parametric estimation; the Kenya study used nonparametric methods in the postestimation comparison of income distributions; the Ethiopia and Rwanda studies used qualitative livelihoods data (collected independently from the income data) that corroborate the quantitative findings from the income data; and the Ethiopia study used self-reported data (for example, on food gaps, meaning the number of months with insufficient food) and anthropometric data (for example, height, weight, and age) that will be used to triangulate findings.

Recommendations: Empirical measurement of impact indicators should be of sufficient quality to support statistical attribution of indicator changes to USAID programs. Empirical quantification of impact indicators should be supplemented with additional quantitative or qualitative information sufficient to validate 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 achieving the first Millennium Development Goal and the higher-level goals of FTF are income, poverty, and child underweight. Primary data collection at the household level is the most accurate way to measure these; it is also the most direct way to measure income and poverty. Although the assessments measured changes in income and poverty status attributable to USAID-supported projects at the household level, for analysis and reporting purposes the changes were aggregated to obtain the change in mean household income or in poverty rates. These aggregated changes were presented in terms of pre-specified subsamples (for example, districts in Rwanda) or through econometric analysis that quantified mean project effects (for example, in the Kenya and Rwanda projects as well as in the food gap project in Ethiopia).

Recommendation: Empirical quantification of impact indicators should rely on household data.

It is critically important to have good baseline data. The comparison of impact indicators before and after a project relies on baseline data. The Ghana 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: for example, in Ethiopia, data collection began in the first year of the project and still provided useful results. A review of related literature showed that slightly less than one half of the evaluation studies reviewed did not have any baseline data,

and the quality of the baseline data in the rest of the studies varied. None of the studies had baseline data on nutritional measures. Related to this is the need for baselines with knowledgeably defined beneficiary groups and expertly designed statistical sampling frames.

Recommendation: All projects should construct or cause to be constructed a baseline dataset.

It is critical to have an appropriate comparison group. The use of a comparison group helps to control for external factors, such as drought, that affect the entire population's income growth, poverty reduction, or hunger reduction. The East Africa, Ethiopia, Kenya, and Rwanda studies all rely on a comparison group against which the changes in the treatment group can be measured. The difference-in-differences 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, although the evaluation was designed to include treatment and comparison groups, evaluators felt that additional information was gained by also comparing the treatment group with smallholders in a nationally representative sample.

Recommendation: All projects should specify or cause to be specified treatment and comparison groups for evaluation and measurement purposes.

Longitudinal (panel) datasets 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 allows for statistical control of household idiosyncrasies and improves understanding of the household's emergence from poverty. Confidentiality is a concern—especially if using secondary data such as a national survey—but it can be solved. Longitudinal data allow for disaggregation to address specific interests and sometimes allow for combination with other datasets. Large

sample sizes are appropriate, especially when it is desirable to have results disaggregated by project activity or subactivity, beneficiary type (for example, poor versus nonpoor), gender, or agroclimatic conditions. In Kenya a sample size of around 600 households for all three activities led to small cell counts when data were disaggregated by activity, group (that is treatment versus nontreatment), and poverty status. In Ethiopia a large sample size was required to understand the relationship between the intensity of project activity and the level of impact.

Recommendations: USAID baseline surveys should be designed for follow-up surveys that generate a longitudinal household dataset. Sample sizes should be calculated to provide a full range of analytical opportunities for impact indicators and causal pathways (if possible). 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 end-of-project datasets. The National Institute of Statistics of Rwanda (NISR) conducted household surveys including income 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 were not yet available. In order to complete the Rwanda impact assessment study in a timely fashion, primary data were collected for 2010, facilitated by NISR's provision of information on sampling frames and survey instruments. The availability of prior data enabled this particular analysis, and more generally would lower the cost of any impact assessment. 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, which would still provide unique project information but at a lower cost.

Recommendation: Impact assessments should 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 to quantify the effect of programs on child nutritional status. While conducting the East Africa regional study of the impacts of trade flows, the Regional Strategic Analysis and Knowledge Support System (ReSAKSS) tried to access the Kenya DHS data on child underweight in order to quantify impact on child nutritional status. The working hypothesis was that smallholders living in food-deficit regions and in proximity to a trade route would realize improved child nutritional status as trade flows of staple foods, such as maize, 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 (which is mildly offset to preserve confidentiality) that likely would have served to delineate treatment and comparison households for a difference-in-differences analysis of maize trade flows and child nutritional status. In March 2011, ReSAKSS began a process for access to disaggregated DHS data including formal requests but has not yet been able to access the disaggregated data. It is unclear whether access will ultimately be granted.

Recommendation: To ensure that DHS data are not only useful but actually used (if applicable) to assess the impact of USAID agricultural projects on child nutritional status, the Bureau for Africa and the Bureau for Food Security should request a review of access procedures for DHS data.

Project activity reporting and accounting must be compatible with impact assessment needs. The TIPCEE project in Ghana had multiple activities taking place simultaneously, which made it difficult to determine how much money was spent on pineapple or mango activities versus how much money was spent on other activities. For example, within each project, the amount of time a scientist spent on any particular crop was difficult to quantify, as were

operational expenses such as vehicles or offices. In the halted assessment of PROFIT, it was difficult to determine which districts PROFIT worked in during any given year and subsequently which districts might or might not have benefitted from project activities. In Kenya each commodity value chain had its own project, which made it relatively easy to determine costs of dairy, for example. Issues remain in understanding which Kenyan project beneficiaries received which services and adopted which innovations, especially in the horticulture and maize activities. Therefore it is difficult to understand which innovations generated the largest impact and why. For example, some projects provide a field demonstration to many farmers but work in a hands-on fashion with a smaller number of farmers. Some smallholders participate in the program some years but not in others. In Ethiopia, the level of impact depended nonlinearly 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 disaggregation of benefits and costs may be desirable.

Recommendations: Projects should 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. Implementing partners might apply the suggestions on time, effort, and indirect cost reporting found in the Office of Management and Budget circulars A-21 and A-87, even if not contractually required to do so.

The selection of method is less important than the conscientious application of that method. There is currently debate about the use of randomized clinical trial (RCT) methods versus quasi-experimental models (QEM) to conduct assessments, and it is possible to generate lengthy arguments on the advantages of either method over the other. For practical reasons, the QEM methods were selected for the impact studies. RCTs cost more money, require clinical-type control over subject behaviors and actions, take longer than a year to complete (typically), do not account

for uncontrolled or unforeseen external factors that may influence results (typically), and face difficulties generating out-of-clinic results (possibly including those that would inform scaling-up of programs under FTF). The standard of solid and credible but not necessarily incontrovertible attribution of impacts to projects also influenced selection of QEMs over RCTs. Further, use of participatory methods may preclude use of RCTs, as seen in Ethiopia where the local community selected the project beneficiaries. Similar statements can be made about specific statistical and econometric tools: the keys are the appropriateness of the method selected to the project evaluation and the quality of the application.

Recommendations: USAID impact assessments should 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. USAID should 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-differences comparison between treatment and comparison groups. In 2000, Rwanda was very poor: even doubling the average smallholder's income left many short of the $1.25 standard poverty line. However, based on the rural poverty line (which is accepted by the United Nations for Millenium Development Goal 1 (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, which means that relatively small increases in income could help a household emerge from poverty. Regression estimates of poverty declines showed large differences in household income but only a modest difference—three percentage points—in the poverty rate between treatment and comparison

groups. This is because even the modest income increases in the comparison group 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 the exchange rates used), and at $2.00, the treatment group decrease in poverty was eight percentage points larger than the decrease in the comparison group poverty rate.

Recommendation: The assessment of headcount poverty reduction attributable to a USAID-supported project should 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 collection 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 is therefore considered more reliable. It was applied in the Ethiopia and Rwanda data collections. The Rwanda dataset showed an abnormally low 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 expenditure data available for triangulation. Consequently the quality of the income variable in each of the datasets may be very good but is currently unknown.

Recommendation: Further investigation should be conducted into the type and quality of information contained in different approaches to smallholder household income measurement.

The frequency of assessing impact depends on the projects being evaluated. The Ethiopia PSNP analysis (including primary data collection) is conducted every second year, the

Kenya USAID impact indicator dataset is updated every second year, and the Rwanda assessment is 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—baseline, midterm, and end of project—for the final evaluation. 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 household income and poverty trajectories. This allows for more robust estimation of the difference-in-differences estimator of project impact. Additional data, such as postproject data (three to five years later), could further improve estimation quality and comprehension of causal pathways.

Recommendations: Baseline, midterm, and end-of-project impact data should be gathered. 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 three to 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 are (1) specialists are often more knowledgeable about impact assessment than project implementers, (2) project managers do not have to manage impact assessments in addition to programmatic activities, and (3) third-party assessment is generally viewed as more objective. A disadvantage of the third-party approach is that organizations that specialize in impact assessment are often less knowledgeable than project implementers

about agricultural and value-chain specifics. Operational detail is informative programmatically within the impact assessment, and it also informs the way the impact assessment is designed and conducted. Having third parties work closely with programming partners to understand operational detail can be very important.

Recommendations: Impact assessments should generally be conducted by experienced third parties; if possible, these third parties could be local parties with knowledge of local agriculture and agricultural value chains.

Spillovers and indirect effects may be very important. The agricultural economic literature suggests that 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 nonfarm employment may have indirect effects 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, and its findings are consistent with an indirect income effect comparable in magnitude to the direct effect.

Recommendation: Future impact assessments should 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 conducted shortly after the main harvest. Crop-cut data are generally regarded as more reliable than farmer recall. However, crop cuts usually occur during harvest; socioeconomic surveys are generally postharvest surveys because farmers are too busy during harvest to spend time being interviewed or filling out surveys. 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 the smallholder is willing to sell, the number of buyers in the village or market, prior relationships the smallholder might have with buyers, cooperative memberships, smallholder bargaining acumen, and other factors. Smallholders may also sell products in sets 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 capable of delineating the effect of the project on prices.

Recommendation: In the context of impact assessment, further investigation into the most appropriate methods for quantifying smallholder yields and prices received should be conducted.

BIBLIOGRAPHY OF IMPACT ASSESSMENT STUDIES

East Africa

Guthiga, Paul, Stephen Wambugu, Maurice Ogada, Stella Massawe, Joseph Karugia, Hikuepi Katjiuongua, and James F. Oehmke. *An Assessment of the Impact of Cross-Border Trade in Maize on Welfare in the East African Region.* Washington, DC: USAID, 2011.

Ethiopia

Berhane, Guush, John Hoddinott, Neha Kumar, and Alemayehu Seyoum Tafesse. *The Impact of Ethiopia's Productive Safety Nets and Household Asset Building Programme: 2006–2010.* Washington, DC: International Food Policy Research Institute, 2011.

Ghana

Mensah, Albert O. *Mango Farmers Benefit from the Trade and Investment Program for a Competitive Export Economy (TIPCEE) Project.* Cape Coast, Ghana: University of Cape Coast, 2011.

Kenya

Oehmke, James F., Thomas S. Jayne, Sarma B. Aralas, and Mary K. Mathenge. *Impacts of USAID/Kenya Supported Agricultural Productivity Interventions on Household Income and Poverty Reduction.* Working Paper 38. Nairobi, Kenya: Tegemeo Institute of Agricultural Policy and Development, 2010.

Smale, Melinda, Mary K. Mathenge, Thomas S. Jayne, Eduardo Magalhaes, John Olwande, Lilian Kirimi, Mercy Kamau, and James Githuku. *Income and Poverty Impacts of USAID-Funded Programs to Promote Maize, Horticulture and Dairy Enterprises in Kenya, 2004–2010.* Nairobi, Kenya: Tegemeo Institute of Agricultural Policy and Development and Michigan State University, forthcoming.

Rwanda

Bihogo, Etienne, Dave D. Weatherspoon, and James F. Oehmke. *The Impact of PEARL and SPREAD Projects on the Development of the Rwandan Coffee Sector.* Kigali, Rwanda: USAID, 2011.

Moss, Charles B., James F. Oehmke, and Alexandre Lyambabaje. *An Economic Evaluation of SPREAD Impact on Rwanda's Rural Population.* Kigali, Rwanda: USAID, forthcoming.

Oehmke, James F., Alexandre Lyambabaje, Etienne Bihogo, Charles B. Moss, Jean Claude Kayisinga, and Dave D. Weatherspoon. *The Impact of USAID Investment on Sustainable Poverty Reduction among Rwandan Smallholder Coffee Producers: A Synthesis of Findings.* Kigali, Rwanda: USAID, 2011.

Literature

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

James F. Oehmke is an economic consultant based in East Lansing, Michigan, United States. **Charles B. Moss** is a professor in the Department of Food and Resource Economics at the University of Florida, Gainesville, United States. **John Hoddinott** is deputy director of the Poverty, Health, and Nutrition Division at the International Food Policy Research Institute (IFPRI), Washington, DC. **Melinda Smale** is a professor at Michigan State University, East Lansing. **Athur Mabiso** is a postdoctoral fellow in IFPRI's Development Strategy and Governance Division. **Neha Kumar** is a postdoctoral fellow in IFPRI's Poverty, Health, and Nutrition Division.

U.S. Agency for International Development

1300 Pennsylvania Avenue, NW
Washington, DC 20523
Tel: (202) 712-0000
Fax: (202) 216-3524
www.usaid.gov