



USAID
FROM THE AMERICAN PEOPLE

ASSESSING THE IMPACT OF NEW GENERATION PRIVATE SECTOR DEVELOPMENT PROGRAMS

IMPACT ASSESSMENT PRIMER SERIES
PUBLICATION #1

PRIVATE SECTOR DEVELOPMENT IMPACT ASSESSMENT INITIATIVE

DECEMBER 2006

This publication was produced for review by the United States Agency for International Development. It was prepared by Don Snodgrass of Development Alternatives, Inc.

ASSESSING THE IMPACT OF NEW GENERATION PRIVATE SECTOR DEVELOPMENT PROGRAMS

IMPACT ASSESSMENT PRIMER SERIES
PUBLICATION #1

PRIVATE SECTOR DEVELOPMENT IMPACT ASSESSMENT INITIATIVE

DISCLAIMER

The author's views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

CONTENTS

- What Are New Generation Private Sector Development Programs? 1**
- What is Impact Assessment?3**
- How Does Impact Assessment Differ from Other Types of Program Evaluation?3**
- Why Are Impact Assessments of PSD Programs Needed?4**
- What Are the Steps in Implementing an Impact Assessment?6**
- What is a Credible Impact Assessment?9**
- What Are Methodological Challenges of Doing Impact Assessments? 10**
- When Should Impact Assessments Be Done? 11**
- What Does Impact Assessment Cost? 11**
- What Are the PSD Impact Assessment Initiative and the Impact Assessment Primer Series? 11**

WHAT ARE NEW GENERATION PRIVATE SECTOR DEVELOPMENT PROGRAMS?

The private sector is the engine that drives economic growth. It is a large and growing segment of nearly all economies as well as the main source of the innovation and dynamism on which economic growth depends. Economic growth in turn is essential for reducing the prevalence of poverty in developing and transition countries. The poor benefit from economic growth in four distinct ways:

1. **As farmers and entrepreneurs.** Economic growth raises end-market demand for the goods and services produced by micro and small enterprises (MSEs) owned and operated by the poor, increasing the income that they earn.
2. **As workers.** Economic growth raises the demand for labor, increasing wage payments received by those who have nothing to sell but their labor.
3. **As consumers.** Rising production from MSEs increases the supply of lower-priced goods and services consumed by the poor, raising their purchasing power.
4. **As potential recipients of tax-funded services or transfers.** Economic growth boosts government revenues, which can finance expansion of services like elementary education and basic health care if policy makers so decide.

Recognizing these facts, USAID and other development agencies have invested hundreds of millions of dollars in a variety of *new generation private sector development* (PSD) programs in dozens of developing and transition countries. These programs intervene at selected points within the constellation of forces that shape private sector development.¹

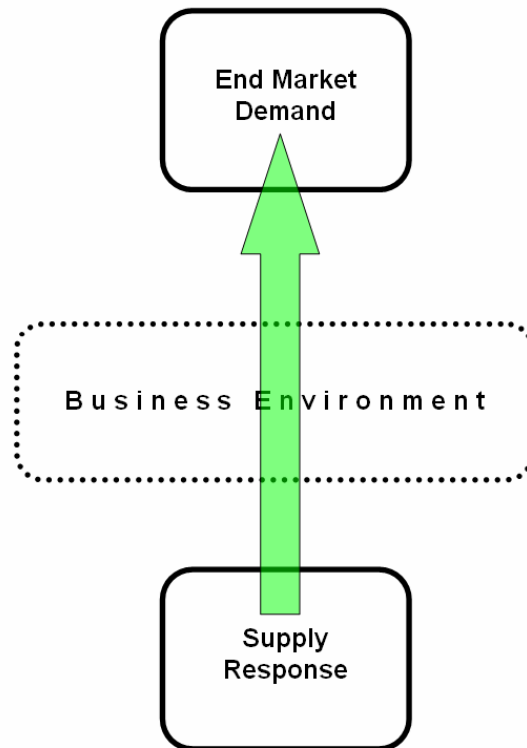
As Figure 1 indicates, the ultimate driver of private sector development is end-market demand (domestic or international), which creates incentives to increase production and upgrade enterprises, including MSEs. Since local markets in many developing countries are small and grow relatively slowly, overseas markets provide attractive alternatives in many cases. Demand incentives are filtered through the business environment, which has local, industrial, national, and global components and can significantly affect the risks and transaction costs involved in doing business. Industries and firms respond, with greater or lesser vigor, to the demand signals that reach them. Competitiveness is fundamental to the successful response of industries and/or business networks to market demand. It can be achieved by improving productivity, through product differentiation, or by emphasizing unique product characteristics (e.g., by shifting from commodity coffee to specialty coffee).²

Networks of firms are also important for achieving competitiveness. They can take the form of clusters of enterprises producing similar or complementary items or of value chains that move goods and services from their primary producers to end markets through a series of stages. They may also involve external suppliers, processing agents, marketing firms, think tanks, banks or other financial institutions, consulting firms, government entities, and many others. MSEs need linkages to such networks to achieve competitiveness, especially when striving to meet overseas demand.

¹ For a comprehensive review of recent PSD programs, see Donald Snodgrass, (2005), “Inventory and Analysis of Donor Sponsored MSE Development Programs,” microREPORT #15, Washington, DC: USAID.

² For an in-depth discussion of the forces shaping private sector development, see Jeanne Downing, Donald Snodgrass, Zan Northrip, and Gary Woller, (2006), “The New Generation of Private Sector Development Programming: The Emerging Path to Economic Growth with Poverty Reduction,” microREPORT #44, Washington, DC: USAID.

Figure 1. A Model of Private Sector Development



Supporting markets, including finance and non-sector specific business services and products, help industries and enterprises respond to changing patterns of end-market demand. Demand for the goods and services traded in supporting markets is derived from end-market product demand. In other words, firms buy those services and inputs that are expected to help them boost earnings in their product markets.

Within this constellation of forces, new generation PSD programs usually seek to improve aspects of the business environment and/or strengthen the supply response from industries and firms. Traditional PSD programs were limited to the supply side and usually worked with individual firms. Later, improving financial services was emphasized. Now more attention is being paid to inter-firm linkages, improving the business environment, and facilitating market solutions to constraints on expansion and upgrading, rather than direct provision of services. Many programs seek to develop specific value chains and promote MSE participation in them, while others feature a small number of specific interventions that deal with barriers hypothesized to be critical for PSD (e.g., improving small business access to finance or boosting general managerial capacity). Still other programs emphasize the strengthening of business member organizations. New generation PSD programs often combine multiple interventions aimed at promoting private sector development in different ways.

New generation PSD programs typically seek impact at national, regional, sector, enterprise, and/or household levels. Impacts sought by new generation PSD programs include accelerated economic growth; a better business climate; upgraded firm performance; strengthened competitiveness of selected value chains and clusters; stronger supporting markets for finance, inputs, and business services; employment creation; and a reduction in the incidence and severity of poverty. Unlike many past efforts, new generation PSD programs emphasize *sustainable* improvements by working through markets and private sector agents and by reducing an eventually eliminating subsidies provided to participating businesses.

Notwithstanding donors' extensive investment in new generation PSD programs, there is little credible evidence on the extent to which their goals are achieved. Do current donor interventions intended to improve the performance of the private sector actually work? Do they increase the demand for labor and raise wages? Do they lead to improved competitiveness and firm upgrading? Do they help raise the participation of MSEs in more promising value chains and bolster the rewards that they receive for their participation?

Impact assessment can provide donors with vital information on program effectiveness, a *sine qua non* for measuring cost-effectiveness. High-quality impact assessments of selected PSD programs would facilitate adjustments to increase the impact of on-going programs and improve future programming.

WHAT IS IMPACT ASSESSMENT?

Impact assessment is a form of analysis that can determine, within a reasonable margin of error, (1) the impacts of a PSD program, including both intended and unintended effects, (2) the magnitude of the impacts, and (3) the causal factors underlying the impacts. Beyond impact assessment, cost-benefit or cost-effectiveness analysis would be needed to determine whether the measured impact justified the cost of the intervention.

The central problem of impact assessment is establishing attribution. This requires the analyst to go beyond other forms of evaluation (such as performance monitoring, contractor evaluation, or USAID's performance measurement plan) to determine not only whether the desired outcomes occurred but also whether those outcomes occurred *because the program was implemented*. That in turn requires the definition of a *counterfactual*, or what would have happened if the program had not been implemented. The plausibility of an impact assessment is determined by the success achieved in establishing a counterfactual.³

HOW DOES IMPACT ASSESSMENT DIFFER FROM OTHER TYPES OF PROGRAM EVALUATION?

Any development program is based (explicitly or implicitly) on a causal model that shows how program activities can lead in turn to desired outputs, outcomes, and ultimately impacts (see Figure 2).⁴ Many program designs provide for monitoring of program activities and outputs to see whether activities took place on time and whether they resulted in the anticipated outputs and outcomes. Impact assessment differs from this kind of *performance monitoring* in two ways. First, it focuses on the program's higher-level objectives: the outcomes and impacts achieved at the regional, sub-

³ For a representative review of PSD impact assessments, see Lily Zandniapur, Jennefer Sebstad, and Donald Snodgrass, (2004), Review of Evaluations of Selected Enterprises Development Projects," microREPORT #3, Washington, DC: USAID.

⁴ The causal model in Figure 2 is taken from Don Snodgrass and Jennefer Sebstad, (2005), "Assessing the Impact of Kenya BDS and Horticulture Development Center Projects in the Tree Fruit Value Chain in Kenya: Baseline Research Report," microREPORT #33, Washington, DC: USAID.

sector, firm, and household levels (Columns 4 and 5 in Figure 2). Second, as already mentioned, it attempts to measure changes that occurred *because the program was implemented* and would not have happened otherwise.⁵

In contrast, performance monitoring is primarily concerned with program activities and outputs (Columns 2 and 3 in Figure 2). It thus involves the left side of the causal chain that leads from program activities to desired impacts, while impact assessment deals with the right side of the causal chain. Performance monitoring is an important management function that is best carried out by internal program staff with quick feedback to allow mid-course adjustments to be made in program operations. Impact assessment is best done by independent analysts, albeit in close consultation with program management, and has a longer turnaround time.

In addition to performance monitoring, programs also conduct mid-stream or ex-post evaluations. Contractor evaluation is a common example. Like performance monitoring, contractor evaluation focuses on the left side of the causal chain, usually limiting itself to asking whether the contractor carried out its assigned functions punctually and competently. While it is important to ask such questions, good contractor performance does not guarantee—for example, if the program design is faulty or the resources devoted to the program are inadequate.

Tracking indicator performance, as in the performance measurement plan (PMP), does not ensure program impact. It may be, for example, that defined targets at the sub-sector and firm levels are achieved after the program's inception, yet they would have been met anyway because their movement is explained primarily by factors other than the program. By the same token, failure to meet indicator targets may also be explained by external factors; performance could have been even worse in the program's absence.

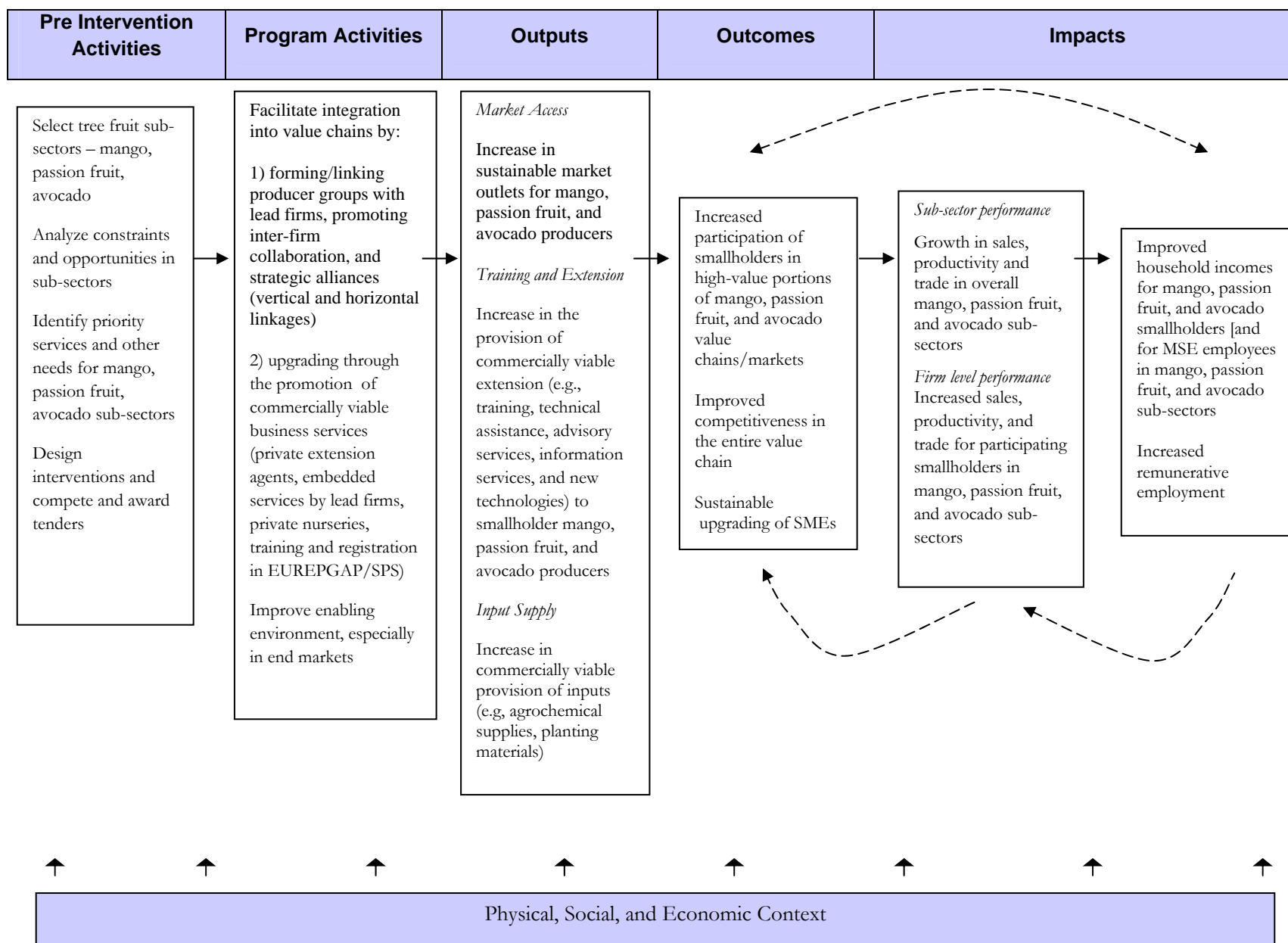
WHY ARE IMPACT ASSESSMENTS OF PSD PROGRAMS NEEDED?

There is a growing need to measure the contributions that new generation PSD programs make to private sector development and economic growth with poverty reduction. USAID, for example, has declared its desire to invigorate the “culture of evaluation” within the Agency and become more of a “learning organization.” Director of U.S. Foreign Assistance and USAID Administrator Randall Tobias recently referred to “new responsibilities to focus on performance results [and] accountability,” indicating a renewed commitment at USAID to more cost-effective programming. Pressure for USAID to show that positive results are being achieved also comes from Congress and the Government Accountability Office (GAO).

Other donors are likewise feeling pressure to demonstrate the impacts of their PSD programs, if only to justify additional commitments of funds to their assistance programs. Impact assessment is already widely applied in other areas of development—for instance, in programming for health, education, and social welfare—but its application to PSD programs is in its infancy.

⁵ For more on causal models, see the forthcoming Impact Assessment Primer Series article by Gary Woller and Jeanne Downing, “Developing and Using Causal Models in Conducting Impact Assessments of Private Sector Development Programs.”

Figure 2. Sample Causal Model (for Kenya BDS and Fintrac Programs)



Implementation of a set of high quality impact assessments of new generation PSD programs would offer a new body of credible knowledge about what works and what does not work, bringing multiple benefits not only to donors but also to policy makers and practitioner organizations:

- It would help donors and policy makers design better PSD programs and achieve more impact for given resources.
- It would help donors and policy makers modify or redesign ongoing programs to produce better results.
- It would help donors and policy makers allocate funds among competing development programs more effectively.
- It would give practitioner organizations, such as NGOs, a better way to demonstrate results than anecdotal examples.

Doing impact assessments of all PSD programs would be unreasonably costly. To ensure cost-effectiveness, impact assessments should be targeted strategically at a limited number of large, multimillion dollar programs and/or innovative programs with a large potential for generalized learning that can be used to improve future program designs.

WHAT ARE THE STEPS IN IMPLEMENTING AN IMPACT ASSESSMENT?

Conducting a good impact assessment of a PSD program involves the following steps:

1. Selecting the program(s) to be assessed
2. Conducting an evaluability assessment
3. Preparing a research plan
4. Contracting for and staff the impact assessment
5. Carrying out a baseline assessment and analyze its results
6. Carrying out follow-up assessment and analyze its results⁶
7. Reporting on and disseminating assessment findings

These steps are outlined in this section and discussed at greater length in later papers in the Impact Assessment Primer Series (see below).

Selecting the Program(s) to be assessed. Impact assessments are carried out because someone needs to know what results particular programs or intervention approaches are achieving. This demand for information is most likely to arise from the higher levels of donor organizations (those who plan aid programs and allocate funds), but it can also come from Mission Directors, program officers, and implementers, such as NGOs. Specific reasons for doing impact assessments of new generation PSD programs have already been mentioned. As noted earlier, large programs and those that take innovative and promising approaches are strong candidates for assessment. Another important

⁶ This paper emphasizes “quasi-experimental” impact assessments, in which a control group is selected so as to be comparable to a previously defined group of program participants. As discussed below, an alternative methodology is the “experimental” approach, in which *both* the participant group and the control group are picked at random from a qualified population. If the experimental design is used, a baseline survey is not required because the evaluation rests on outcome differences between the two groups at a point of time that allows for the program being assessed to achieve significant impact.

criterion is a cooperative attitude among program leadership and the USAID Mission. Finally, the availability of finance to do the assessment is obviously critical.

Conducting an evaluability assessment. This is an initial assessment of whether an impact assessment should be conducted on the program and, if so, what is the most appropriate methodology to do so.⁷ An important part of the evaluability assessment involves sitting down with program staff to work through a causal model for all the program activities to be covered in the impact assessment. This means determining what exactly the program is doing or will do, over what time period, and with what expected outputs, outcomes, and impacts. If this discussion indicates that the program is unrealistic in its postulated relationships between activities and impacts, or if the time period is unsuitable (a follow-up assessment can generally not be conducted in less than two years after the program commences), then the impact assessment is not worthwhile.

The evaluability assessment should also consider the purposes that would be served by an impact analysis, the audience for the research, its cost-effectiveness, its potential credibility, and the best timing.⁸

Preparing a research plan. The research plan should include the causal model of the impact assessment and a practical plan for carrying out the study. The causal model is used to generate a set of hypotheses about outcomes and impacts that will be tested in the study. Typically, impacts of several different types will be anticipated at three levels:

1. In the value chains and markets involved, including product markets and sometimes also supporting markets for inputs, business services, and/or finance
2. Among participating MSEs
3. In the households associated with participating MSEs

Once testable hypotheses have been identified, the next step is to define measurable indicators that can be used to determine whether impact has been achieved. After that, sources of information for measuring the indicators must be identified. In the quasi-experimental approach (see definitions below), a longitudinal survey serves as an important source of information for determining whether there is impact at the MSE and household levels. This involves selecting a sample of program participants (explicitly defined in a manner consistent with the program's structure and approach) and matching it with a sample of non-participants who are as similar as possible to the program participants in all relevant characteristics (the control group). This must be done carefully to minimize the effect of selection bias—the tendency for people who would have done better anyway to become program participants—which leads to overstatement of the program's impact. In a quasi-experimental impact assessment, the two groups of survey respondents form a panel that will be interviewed at least twice, with a minimum interval of two years between survey rounds. To allow for attrition in the sample between rounds, over-sampling is required in the baseline round. In an experimental assessment, the two groups are selected at random and then interviewed just once at the conclusion of the study.

The survey is the quantitative part of the impact assessment. It can be combined with qualitative research to get a richer view of impact at the MSE and household levels, as well as to obtain some idea of what the program's impact is at the value chain and market levels. At these higher levels, finding a satisfactory control group is likely to be difficult if not impossible, so impact cannot be proven so definitely as at the MSE and household levels. The qualitative research consists of a program of structured interviews, focus group discussions, and other qualitative methods with

⁷ For more on evaluability assessments, see the forthcoming Impact Assessment Primer Series article by Elizabeth Dunn, "Evaluability Assessment: The First Step in Assessing Impacts."

⁸ For examples of evaluability assessments, see the PSD Impact Assessment Initiative website at www.microlinks.org/psdimpact.

persons who participate in various ways in the relevant value chains and markets. Their views and insights are then triangulated in an attempt to get a coherent picture of the structure of the markets concerned and changes over time that may be attributable to program activities.

The research plan should also include detailed specifications for the questions to be asked on the survey questionnaire and guidelines for the interviews and focus group discussions.⁹

Contracting for and staffing the impact assessment. Once a draft research plan is drawn up, the next step in a quasi-experimental impact assessment is to make arrangements to carry out the baseline research.¹⁰ This will involve some combination of external consultants (senior and/or junior) and local research partners (firms and/or individual consultants). Typically, a local research firm is hired to carry out the quantitative and qualitative parts of the baseline study under the guidance of the assessment's sponsors and designers. In this case, a competitive bidding process is desirable. Potential partners submit a bid based on a scope of work (SOW) that clearly defines the responsibilities, time schedule, and budget for the baseline study.¹¹ Selection of the local contractor will consider several factors, including past experience, technical expertise, the quality of the proposal, recommendations, timing, and cost.¹²

Carrying out a baseline assessment and analyzing its results. With the guidance and participation of the consultants, the local research partners carry out the baseline research, including the survey, interviews, focus group discussions, and other qualitative methods. The survey data are then entered, cleaned, and tabulated, while careful records of the interviews and focus group discussions are prepared. The consultants next analyze the results of the baseline research. The analysis at this stage is largely descriptive; the aim is to create an accurate picture of the program's setting and participants near the start of program implementation that will serve as a standard against which changes can be measured after the follow-up round of research is completed.

Carrying out follow-up assessment and analyzing its results. After an interval long enough to create a reasonable expectation that program impacts might have become measurable (at least two years), follow-up research is carried out. This closely follows the pattern of the baseline round and, to the extent possible, involves the same respondents. Program impact is assessed using the “difference in difference” method—that is, changes for the participant group are compared to changes for the control group. Impact is inferred if the changes for the participant group are significantly more favorable than those for the control group. The analysis must also take account of “mediating variables” that might affect this comparison—for example, differences in wealth, age, gender, or educational attainment between the two samples.

Reporting on and disseminating assessment findings. Since the impact assessment is likely to generate information that has value beyond the particular PSD intervention analyzed, it is vital that the lessons learned through the study be transmitted effectively to all those who are in a position to use them. Possible means of dissemination include web postings, seminar or conference presentations, workshops, and published papers.

⁹ For examples of research plans, see the PSD Impact Assessment Initiative website.

¹⁰ The research plan will continue to undergo refinement as the details of the impact assessment are worked out by the assessment designers and the local research partner.

¹¹ For examples of SOWs for local research partners, see the PSD Impact Assessment Initiative website.

¹² For more on selecting local research partners, see the forthcoming Impact Assessment Primer Series article by Lucy Creevey, “Methodological Issues in Conducting Impact Assessments of Private Sector Development Programs.”

WHAT IS A CREDIBLE IMPACT ASSESSMENT?

A credible impact assessment will include a longitudinal survey that satisfies the following minimum acceptable methodological standards:

1. It will include observations on a group of participants (treatment group) and a matched group of non-participants (control group).
2. It will assess the status of both treatment and control group members at a time after impacts can reasonably be expected to have occurred (follow-up).
3. It will be based on a causal (logical) model in which clearly stated hypotheses link program activities to expected impacts.
4. It will be rigorous, in that all methodologies used are well documented and their weaknesses identified.
5. It will use data collection methods that follow accepted good practice.
6. It will use analytical methods that are appropriate, in that they match the type of data collected.
7. If a quasi-experimental methodology is used, it will include data on both treatment and control group members before impact could have occurred (baseline).

Two types of impact assessment potentially satisfy these minimum acceptable methodological standards: experimental and quasi-experimental. The experimental method requires that MSEs selected to participate or not participate in the program be drawn at random from a population that meets the relevant selection criteria. Those randomly selected to participate in the program become members of the treatment group for purposes of the impact assessment, while those randomly selected not to participate become members of the control group.

Quasi-experimental methodologies were developed to deal with the messy world of field research, where it is not always practical, ethical, or even possible to assign firms to treatment and control groups on a random basis. In contrast to experimental methods, quasi-experimental methods do not randomly assign units to treatment or control groups but instead compare groups that already exist. Treatment group members are selected via random sampling of known program participants, while control group members are selected via random sampling of known non-participants who have characteristics similar to those of the treatment group.¹³

In addition to a survey that satisfies the requirements of either the experimental or the quasi-experimental approach, a credible impact assessment should use a mixture of quantitative and qualitative research methods. Quantitative methodologies include, principally, household and firm level surveys and analysis of secondary data. Qualitative methodologies include, principally, in-depth key informant interviews, focus group discussions, case studies, and a range of participatory assessment methods. Qualitative research methods can supplement quantitative methods in important ways by providing deeper insight into the reasons for impacts measured in the quantitative research and can also detect impacts that may have been missed in the quantitative study.¹⁴

¹³ For more on minimally acceptable methodological standards, experimental and quasi-experimental methods, and methodological challenges in conducting impact assessments, see the forthcoming Impact Assessment Primer Series article by Lucy Creevey, “Methodological Issues in Conducting Impact Assessments of Private Sector Development Programs.”

¹⁴ For more on quantitative and qualitative impact assessment methodologies, see the forthcoming Impact Assessment Primer Series article by Lucy Creevey, “Collecting and Using Data for Impact Assessment.”

WHAT ARE METHODOLOGICAL CHALLENGES OF DOING IMPACT ASSESSMENTS?

Several methodological challenges arise when standard procedures for doing impact assessment are applied to new generation PSD programs. Among the more important challenges are:

- Selecting valid control groups
- Assessing impact at the industry, value chain, cluster, or market levels
- Controlling for spillover impacts
- Accommodating panel attrition

Selecting valid control groups. In the quasi-experimental approach, identifying a valid control group is necessary to create a counterfactual and establish attribution. To be valid, the control group must be located at a site similar to the site in which the treatment group is located (that is, similar in relevant characteristics that may include population density, economic development, geography, climate, infrastructure, market access, and soil quality) and it must possess personal characteristics similar to the treatment group, both observable (e.g., age, gender, economic status, social status, education, sector, and experience) and unobservable (e.g., entrepreneurship, risk seeking/aversion, attitudes, and values).

In the experimental approach to impact assessment, the treatment and control groups are both selected randomly from a population of firms or individuals who possess traits that would qualify them for participation in the program. In the quasi-experimental approach, the treatment group has already been selected by program management and a control group must be chosen to match it as closely as possible in all relevant characteristics. Matching the treatment and control groups according to these criteria can be difficult, particularly in terms of unobservable characteristics, and it requires close consultation among study designers, field researchers, and program staff.

Failure to match treatment and control groups appropriately creates *selection bias*. In principle, selection bias can either inflate or deflate measured impact, but the more common case in quasi-experimental studies is that impact is overstated because program participants have advantages (related to location or personal characteristics) that would have led to higher performance on impact variables, even if they had not participated in the program. Some degree of selection bias is likely to be present in any quasi-experimental study. The aim is to keep it small enough that it does not invalidate the assessment findings.

Assessing impact at the industry, value chain, cluster, or market levels. Assessing impacts at the industry, value chain, cluster, or market levels raises particular challenges. At these levels, the main problem is the absence of a plausible control group for establishing attribution. One could compare a value chain, say, with another value chain in the same country (e.g., mangos in Kenya vs. coffee in Kenya) or with the same value chain in a different country (e.g., mangos in Kenya vs. mangos in Uganda), but many other influences would enter in, making it difficult to attribute observed differences to program activities.

While such a comparison might be of some use, impact at the industry, value chain, cluster, or market level is more likely to be assessed through information gathered from industry participants and experts using qualitative methods. The views and insights of key informants at these levels can be triangulated to get a coherent picture of the structure of the markets concerned and changes over time that may plausibly be attributed to program activities. This kind of evidence, however, is unlikely to be conclusive on the question of attribution.

Controlling for spillover impacts. Program activities at the firm and household levels often involve the dissemination of information through training, advice, and other forms of learning. In such cases, there is likely to be spillover impacts as program participants pass useful information and practices on to their friends, relatives, and

neighbors. Spillover impacts can also be negative. An example would be if business formation and growth by program clients siphoned sales away from competitors' businesses.

Spillover impacts make impact assessment more difficult, because they blur the distinction between program participants and non-participants. The result is either systematic underestimation or overestimation of true program impacts, depending on whether the spillover impacts are positive or negative. One approach to limit spillover impacts is to locate control sites physically distant from treatment sites, but doing so risks introducing other differences (e.g., climate, soil conditions, access to markets, and infrastructure development) that create selection bias. Another approach to account for spillover impacts is to conduct interviews, focus group discussions, or other qualitative methods with key informants who presumably possess knowledge about potential program spillovers, both positive and negative.

Accommodating panel attrition. Panel attrition refers to a problem that arises in quasi-experimental impact assessments when the treatment and control groups who responded to the baseline survey must be contacted for participation in the follow-up survey. Panel attrition occurs for several reasons: respondents move, die, are taken sick, decide not to participate, or cannot be located. Excessive panel attrition hinders the analysis of research findings by leaving too few observations to allow for meaningful statistical analysis. To accommodate panel attrition, over-sampling is required in the baseline survey.

WHEN SHOULD IMPACT ASSESSMENTS BE DONE?

Ideally, impact assessment is built into the program design from the start. For a quasi-experimental study, a baseline assessment should be conducted as soon as program participants can be identified, or as soon thereafter as possible. The key is to establish the participants' condition before they have been significantly affected by program activities and compare their status to that of the control group. At least one follow-up assessment should be made at an interval of two or more years. If seasonal differences are significant, the follow-up survey should be conducted during the same season as the baseline survey. Analysis of data collected in the baseline and follow-up surveys, along with qualitative research results, form the basis on which the program's impact is measured.

WHAT DOES IMPACT ASSESSMENT COST?

Impact assessment can be carried out at various levels of sophistication with correspondingly different price tags. A credible quasi-experimental impact assessment that uses a mixture of quantitative and qualitative methods could cost up to and exceeding \$100,000. An experimental study could be cheaper because a baseline survey would not be needed. Such expenditure would clearly not be justified for every program, but it is a small amount relative to the total cost of multimillion dollar programs and is also well worth it when applied to innovative programs with a large potential for learning that can be used to improve future program designs. In some cases, significant cost savings may be realized without compromising the validity of the impact assessment by putting more emphasis on qualitative assessment and making greater use of lower-cost junior and/or local investigators. Because of their cost, impact assessments should be used strategically to answer important programming and policy related questions or in conjunction with innovative, expensive, controversial, or other programs with large potentials for generalized learning.

WHAT ARE THE PSD IMPACT ASSESSMENT INITIATIVE AND THE IMPACT ASSESSMENT PRIMER SERIES?

This primer on impact assessment is the first in the Impact Assessment Primer Series to be produced by the PSD Impact Assessment Initiative. The PSD Impact Assessment Initiative is an activity funded by USAID under the Accelerated Microenterprise Advancement Project (AMAP) to promote impact assessments of new generation PSD programs. It accomplishes this objective in four ways.

1. Building a conceptual model that improves understanding of the impacts of new generation PSD programs.
2. Developing and testing rigorous methodologies for measuring the impact of new generation PSD programs at four levels: (a) participating enterprises; (b) associated households; (c) product markets; (d) support service markets.
3. Producing insights about the most effective types of new generation PSD interventions and how they work through implementation of high quality impact assessments and desk research.
4. Providing USAID, USAID missions, and other donors, policy makers, and practitioner organizations with realistic options for assessing the impact of new generation PSD programs and supplying methodological and other guidance on how to conduct credible impact assessments.

The Impact Assessment Primer Series represents partial fulfillment of Item 4 above. The Primer Series is targeted primarily to persons within USAID who have responsibility for or who are otherwise interested in assessing the impact of PSD programs. It should also be of interest to other donors, policy makers, and practitioner organizations who wish to promote private sector development. The Impact Assessment Primer Series will address the range of planning, methodological, logistical, budgetary, and other issues that arise in conducting impact assessments of new generation PSD programs. Topics currently scheduled to be covered in the Primer Series include:

- Developing a causal model for PSD programs.
- Creating a research plan and selecting an impact assessment approach.
- Making critical decisions and their implications for planning and implementing an impact assessment.
- Identifying and addressing methodological challenges in impact assessment.
- Selecting appropriate impact indicators.
- Planning and budgeting for impact assessment.
- Collecting and using data.
- Analyzing and interpreting results.
- Monitoring PSD programs and its relationship to impact assessment.
- Identifying learning resources for impact assessment.

Suitable methodological approaches to assessing the impact of new generation PSD programs are being worked out under the PSD Impact Assessment Initiative. These approaches reflect the vast literature on impact assessment in general as well as the experience currently being gained through the Initiative's test case applications of quasi-experimental impact assessments of PSD programs in Kenya, Brazil, India, and Zambia. Once the final results of those studies are in and have been analyzed together with the results of research conducted by others (the World Bank, IFC, DFID, OECD, UN, etc.), significant learning about what works in PSD programming will have occurred. Yet gaps in our knowledge and questions will inevitably remain. As these issues emerge, further topics will be added to the Impact Assessment Primer Series to fill the gaps and resolve ongoing controversies, contradictions, and ambiguities.

U.S. Agency for International Development

1300 Pennsylvania Avenue, NW

Washington, DC 20523

Tel: (202) 712-0000

Fax: (202) 216-3524

www.usaid.gov