



USAID
FROM THE AMERICAN PEOPLE

INTERVIEWS REGARDING IMPACT ASSESSMENT OF PRIVATE SECTOR DEVELOPMENT PROJECTS: A SYNTHESIS

JULY 2008

This publication was produced for review by the United States Agency for International Development. It was prepared by DAI.

INTERVIEWS REGARDING IMPACT ASSESSMENT OF PRIVATE SECTOR DEVELOPMENT PROJECTS: A SYNTHESIS

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

By Donald R. Snodgrass,
with comments from Jeanne Downing and Gary Woller

CONTENTS

- INTRODUCTION 2
- DEMAND FOR IMPACT ASSESSMENT 2
- FUNDING FOR IMPACT ASSESSMENT 4
- IA METHODOLOGIES 5
- IMPLEMENTATION OF IA..... 7
- DISSEMINATION AND USE OF FINDINGS 7
- TRAINING 7
- INSTITUTIONALIZATION 8
- SUMMARY OF MOST IMPORTANT POINTS 8
- APPENDIX: LIST OF SOURCES..... 10

INTRODUCTION

This paper synthesizes information and opinions gathered during a series of interviews carried out by the Private Sector Development Impact Assessment Initiative (PSD-IAI) with producers and users of impact assessments of private sector development projects. The interviews were conducted by PSD-IAI team members between March 21 and April 17, 2008. The interviewers were Jeanne Downing (JD), Zan Northrip (ZN), Jennefer Sebstad (JS), Don Snodgrass (DS), and Gary Woller (GW). The interviewees are listed in the Appendix.

The following topics are covered in this synthesis:

- Demand for impact assessment
- Funding
- Methodological issues
- Implementation of impact assessments
- Dissemination of results
- Training
- Institutionalization
- Summary of most important points

DEMAND FOR IMPACT ASSESSMENT

Within USAID, interest in evaluation generally and impact assessment specifically has waxed and waned over the years. Little good IA has been done since the Center for Development Information and Evaluation (CDIE) was abolished in the 1990s.¹ Some recent Administrators and the current Coordinator of U.S. Foreign Assistance have expressed strong verbal support, but this has yet to be backed up by commitments in terms of finance, staffing, or organizational reform.

Within the World Bank, which does far more IA than USAID, some leaders push IA while others (including some regional VPs) drag their feet, considering it a diversion from operations. The implementation staff usually does not propose evaluation soon enough, so it has to come from the top. The last Chief Economist was a champion of IA and urged the Bank to take a more systematic approach to IA and compile knowledge related to lessons learned. The current Director of the Development Research Group strongly favors IA and wants to make it the primary focus of his group's work. But some say the results-oriented cultural change envisaged for the World Bank is not happening. People still get promoted on the number of projects they take to the Board. The idea was that promotion would be based on results.

The Independent Evaluation Groups (IEG) at the World Bank and IFC carry out post-project evaluations. The interviews suggest that these evaluations are taken seriously at all levels. Careers can suffer if the IEG's evaluation of a project differs from that of the project's implementers. The Independent Evaluation Office (IEO) at the IMF carries out process evaluations of its enabling environment projects. The IEGs at the World Bank and IFC and the IEO at the IMF are all looking to carry out more formal impact assessments in the future.

Most other donors have an interest in IA, as indicated by widespread participation in NONIE (the Network of Networks on Impact Evaluation, which came out of the OECD Evaluation Network). This does not necessarily mean that they have done a lot of work or have mapped out a clear direction for their future work. We heard that the head of DFID visited the Bank recently and asked the research director what he thought DFID should be doing. MCC requires an evaluation plan as part of each of its country compacts.

Outside pressures are pushing bilateral and multilateral donors to do more IA. There is a strong international consensus (expressed, for example, in the 2005 Paris Declaration on aid effectiveness) that more needs to be done to show results from spending on foreign assistance and increase aid effectiveness. In the US, Congress, while generally supporting increases in at least some categories of aid spending, has demanded proof of results. Several Hill staffers are said to be interested in improving IA. Action by USAID to provide such proof might help curb the demand for senseless indicators. The up-coming political transition offers an opportunity to improve performance in this area.

¹ Honorable exceptions exist in the Office of Global Health and the Democracy and Governance Group in EGAT.

The Center for Global Development (CGD) has lobbied for more and better evaluation. CGD recently launched an initiative called 3IE (International Initiative for Impact Evaluation). The Poverty Action Lab at MIT (J-PAL) is another source of encouragement for IA – as long as it takes the form of randomized controlled trials (RCTs). Leaders of J-PAL claim that the RCT methodology can revolutionize development economics by yielding results analogous to those achieved in medicine through randomized drug trials.

Within development organizations, demand for IA often comes from leaders or central offices but is resisted by operational personnel including project managers, who tend to regard project monitoring as more important than IA and who often prefer anecdotes about project successes to objective analysis. Although some argue that IA must meet the needs and time horizons of project managers, the fact is that IA is more useful to those involved with aid strategy and design of future projects than it is to managers of the projects being evaluated. By the time an IA is completed (three years or more from project inception), the project may be winding down and the person who conceived and initiated it is likely to have been reassigned. Although project managers thus cannot be expected to lead the demand for IA, their cooperation, as well as that of project clients, is required for a successful study. Both groups can sabotage a study if they wish to do so. This is particularly true of RCTs, which require managers to deviate from what might otherwise be their chosen pattern of project implementation. Inducements to cooperate can be helpful. One good opportunity is to help project managers construct their project monitoring system. This work complements the design of an IA because both processes rely on causal models (program logic models, or logframes).

FUNDING FOR IMPACT ASSESSMENT

Many project managers regard IA as expensive and are reluctant to devote any significant fraction of the project budget for such studies. Some argue that a small percentage of the total program budget of a donor agency should be set aside for IA (5-10%?), but the choice of a specific percentage would be arbitrary. Thought should be given to selecting projects for evaluation. Since the benefits of IA accrue more to upper-level administrators and planners than to project managers, it is better to meet the cost of studies from a central allocation for the purpose, rather than being take it from the budgets of the projects to be evaluated. The percentage of the aid agency's total budget to be placed in such a central fund would be a strategic decision, derived from an overall strategy for IA and traded off against other demands on resources. It would also be limited by the capacity to carry out studies. USAID might carry out 5-10 well-focused studies per year.

In the World Bank a lot of money is available for IA from various sources. The studies are expensive, so more than one source must usually be drawn on; entrepreneurship is required. IA does not take a large bloc of staff time in the Bank: the staff of the Development Research Group is only 1% of the World Bank's total staff.

IA METHODOLOGIES

There is a vast literature on IA methodology and lively controversy about the most appropriate methods to use in impact assessment. A useful guideline for navigating this thicket of controversy is the principle that the questions to be answered should dictate the methodology, not the reverse.

The most prominent area of controversy is over the role of RCTs. Advocates of randomization, including many academics led by J-PAL, tout it as the gold standard – the highest achievable level of methodological rigor – and decry other methods as non-rigorous. For example, the head of one of the three major IA efforts at the World Bank,² says that she does not understand the meaning of “quasi-experimental” -- in her view, there are only experimental and non-experimental methods.³

It is generally agreed that a valid IA requires a plausible counterfactual. That is, the study must not only measure changes in the project’s impact variables but must also be able to compare what actually happened with a reasonable estimate of what would have happened in the project’s absence. For the stricter methodologists, this requires that selection bias – the tendency for project participants to differ systematically in relevant ways (that may or may not be observable) from members of a control or comparison group – be eliminated. In a RCT study, project participants are chosen at random from a population of eligible persons, households, or firms. This procedure makes it possible to compare results for participants with those of a control group chosen from the same population with confidence that other relevant characteristics are randomly distributed within the population and do not bear on the comparison. Since not all relevant characteristics are observable, other methodologies cannot meet this standard of eliminating selection bias but can only hope to minimize it through various techniques.

Objections to RCT include arguments that it is impractical (requiring program participants to be selected so as to facilitate the study, rather than by programming criteria) and/or unethical (providing program benefits to some people and excluding others at random). It is also commonly believed that the technique is applicable to simpler interventions (analogous to taking a pill to prevent or cure a disease) but not appropriate for complex and multi-dimensional development programs. Defenders of RCT point out that there are various ways to randomize, which may help get around objections.⁴

The quasi-experimental approach permits program participants to be selected by means defined by the logic of the program itself, rather than by the research design. It creates a counterfactual by selecting a treatment group more or less at random from among program participants and constructing a comparison group that is as similar as possible to the participant group in relevant characteristics other than program participation. This approach requires that a baseline survey be taken as soon as possible after the program’s inception, but it is less invasive than RCT. Since some relevant characteristics are unobservable, however, selection bias cannot be eliminated and the credibility of the counterfactual is thus open to debate.

² These are DIME (which works on rural infrastructure, human development, and urban upgrading), human resources, and the Africa Evaluation Initiative.

³ AL further states that [unspecified] work done at the Bank shows large and unpredictable discrepancies between experimental and other measures of impact for given interventions.

⁴ These include randomization at different levels (district, group, individual), randomization just above and just below a cutoff line for program eligibility, encouragement designs, and randomized phase-ins.

Qualitative research plays different roles in different approaches to IA. RCT practitioners use interviews and focus group discussions primarily to frame hypotheses but rely almost entirely on quantitative analysis to produce their results. Other researchers favor a more even balance between quantitative and qualitative analysis and regard qualitative techniques as important complements to quantitative analysis, which can be invaluable in explaining why the quantitative results came out the way they did.

Views of interviewees about methodology varied widely. In addition to J-PAL, which as an organization does RCTs exclusively,⁵ some interviewees at the World Bank and those at the Millennium Challenge Corporation (MCC) and the Coalition for Evidence Based Policy expressed strong preferences for RCT. IRIS favors mixed methods and uses RCT if it can work closely with the implementing organization. It has found many problems with quasi-experimental methods, particularly in properly defining treatment and control groups. Others were skeptical or critical of RCT. A notable view was that RCT overemphasizes the internal validity of the analysis relative to its external relevance (general applicability). One other interviewee was skeptical of all classic IA methodologies because of spillovers and the presence of other actors, believing that non-random sampling errors drown out random sampling errors in any case.

Despite this divergence of views, even interviewees who strongly favored RCTs recognized that this methodology is not applicable in all circumstances and expressed considerable pragmatism about the need to identify the most appropriate methodology for each particular application. Growing out of this realization is the point that different methodological approaches need to be developed for different types of program. What is appropriate for a health project, say, may not be the right approach for assessing the impact of a private sector development project or a rural roads project.

The interviews identified several key methodological issues:

- The importance of the counterfactual, as discussed above.
- The difficulty of getting pre-treatment baselines.
- The existence of varying forms, degrees, and durations of treatment, which can make it difficult to identify program participants.
- Entry into and departure from the treatment group over the course of the study.
- Selection of a control group.
- Spillover effects, which are good for projects (spillover benefits can be several times as great as direct benefits) but complicate IA studies. This issue leads to skepticism about the applicability of classic IA to private sector projects, which often change over time.
- The presence of other actors who influence impacts, as well as external shocks.
- The timing of the follow-up survey. When the IA is funded by the project, this may be determined by the life of the project and dictate a timing that poorly reflects project benefits. Measured project benefits may be sensitive to the timing of the follow-up survey.

⁵ As individuals, members of J-PAL also do other kinds of studies.

IMPLEMENTATION OF IA

Planning and data collection must start early, especially for RCT. World Bank task managers often put it off too long.

Internal vs. external evaluators: On the one hand, cooperation with donors and implementers is needed; on the other, expertise and objectivity are also needed and may not be available from inside. A veteran USAID official thinks serious evaluation must be internal; contractors do what USAID wants. [But USAID should want serious evaluation.]

To carry out a successful randomized evaluation, the counterpart must really want it. If they don't, then they can sabotage the study whenever they want. Agreement with a donor is not enough.

Local organizations must carry out some data collection. The selection, training, and supervision of these organizations are important matters for the success of the study. Close supervision, not only from leaders of the local organization but also from the principal researchers, is required to ensure that high-quality data are collected. This raises the cost of the study, but saving money by hiring under-qualified enumerators and supervisors is a false economy.

Program staff should be involved in planning the study but should not be involved in carrying it out.

DISSEMINATION AND USE OF FINDINGS

IA results are of no value until they are disseminated and read by people who can use them to improve aid programming. Some say that getting people to read the studies is more important than the methodology. On-line dissemination is an important tool for dissemination. The contest that PSD-IAI held some time ago to choose the best example of IA was also helpful.

TRAINING

General training of USAID staff in the fundamentals of IA would help create basic understanding, teach people how to write reasonable SOWs for IAs, and enable them to participate as educated investors in IA. Causal models would be a good starting point for this kind of training, since they are a common denominator of programming, monitoring, and impact assessment.

A second type of training is engagement with the people (USAID Missions and contractors) for whom we do IAs. Working through causal models with them in detail can yield large mutual benefits.

Finally, higher level and longer training sessions could equip some USAID staff to actually carry our IAs

INSTITUTIONALIZATION

USAID could increase its credibility enormously with good evaluation. The Hill would love it if USAID said it wanted to know what things go wrong and why. The November-April period will be a unique opportunity to propose a new evaluation process at USAID. We will have to wait and see who wins the election, who is on the transition team, and who runs EGAT. Perhaps USAID will establish a high-level evaluation unit, possibly in M. Will there be a social science advisory group?

Does USAID need a central office to promote IA or should it simply raise consciousness and let a thousand flowers bloom? If the main benefit of more and better IA would accrue to high-level managers and programmers, it follows that a central office is needed to direct the effort in line with their perceived needs. If USAID chooses this approach, it will need a plan or process to bring it about. The capacity for doing IAs would probably be no more than a few large-scale IAs per year, depending on the types of studies undertaken, done in combination with other monitoring and evaluation activities. Key program areas might be selected, as the World Bank has done. Further discussion within USAID is needed. It may prove possible to move ahead, even before the transition to the new administration in Washington.

USAID needs external evaluations funded from the top. Contractors for program implementation are beholden to the Missions and are thus not objective. USAID needs to bring back the central evaluation function and bureaus need to take it seriously.

SUMMARY OF MOST IMPORTANT POINTS

- Interest in evaluation generally and IA specifically has increased recently, responding to demands for evidence on the results of foreign assistance.
- IA is most valuable for top administrators and strategic planners; other officials, including project managers and implementers, should be induced to cooperate.
- IA must be built into project designs from the start.
- IA is best funded from a central source.

- Several IA methodologies are available, including variants of the RCT and quasi-experimental models. The best methodology to use in any particular study is the one most likely to provide answers to the questions being asked. External validity (general applicability) is as important as internal validity. The most appropriate approaches are likely to differ across sectors of activity.
- IAs are best designed by independent professionals in consultation with project personnel. The studies should be carried out by the independent professionals and local research collaborators, who should be carefully selected, trained, and supervised.
- IA results are of no value until they are disseminated and read by people who can use them to improve aid programming.
- USAID needs two kinds of training: general training for the relatively large number of staff who will contract and supervise IAs and in-depth training for a much smaller number to equip them to do IA.
- USAID is considering ways to expand and improve its evaluation activities.

APPENDIX: LIST OF SOURCES

Carola Alvarez, Chief, Strategy Development Division, Inter-American Development Bank. Interviewed April 10, 2008 by GW.

David Anderson, Project Leader, Coalition for Evidence-Based Policy. Interviewed April 15, 2008 by JD and DS.

Elizabeth Ashbourne, Senior Operations Officer, Results Secretariat, Operations Policy and Country Services, World Bank. Interviewed April 16, 2008 by JD and DS.

Janet Ballantyne, USAID. Interviewed April 15, 2008 by JD and DS.

Amitava Banerjee, Independent Evaluation Group, International Finance Corporation. Interviewed April 11, 2008 by GW.

David Black (and Mark Billera), Strategic Planning and Research Advisor, DCHA/DG, USAID. Interviewed April 17, 2008 by JD and DS.

John Boyle, Senior Environmental Specialist, Africa Region, World Bank. Interviewed April 17, 2008 by JD and DS.

Asli Demirguc-Kunt, Senior Research Manager, Finance and Private Sector Development Research Group, World Bank. Interviewed April 17, 2008 by JD and DS.

Arianna Legovini, Lead Monitoring and Evaluation Specialist, Africa Impact Evaluation Initiative (AIM), Africa Region Results and Learning, World Bank. Interviewed April 17, 2008 by JD and DS.

Henry Panlibuton, AFE. Interviewed April 15, 2008 by JD and DS.

Bhavani Pathak, Strategic Policy & Evaluation Advisor, USAID, PAICO. Interviewed April 17, 2008 by JD and DS.

Glenn Rogers, USAID. Interviewed April 15, 2008 by JD and DS.

Dana Rutherford and Geetha Nagarajan, IRIS Center. Interviewed by April 7, 2008 by GW.

Donna R. Stauffer, Office of the Director of U.S. Foreign Assistance, U.S. Department of State + Krishna Kumar, Senior Social Scientist, Office of the Director of U.S. Foreign Assistance. Interviewed April 16, 2008 by JD and DS.

John Hicklin, Deputy Director of Independent Evaluation, International Monetary Fund. Interviewed April 10, 2008 by ZN and GW.

James L. Walker, Senior Economic Advisor, Bureau for Asia and Near East, USAID. Interviewed April 7, 2008 by GW.

Melinda Crowley, Senior Performance Management Analyst, Office of Strategic and Performance Planning, U.S. Department of State. Interviewed April 10, 2008 by ZN and GW.

Hap Carr, Manager of Evaluation Division; Franck Wiebe, Chief Economist; and Delia Welsh, Evaluation Manager (not correct title), Millennium Challenge Corporation. Interviewed April 10, 2008 by ZN and GW.

Martin Ravallion, Director, Development Research Group, World Bank. Interviewed April 17, 2008 by JD and DS.

Mark Sundberg, Manager, Independent Evaluation Group, World Bank + James Sackey, Lead Economist. Interviewed April 10, 2008 by GW.

Pedro Olinto, Senior Economist, Development Economics and Office of Chief Economist, Development Impact Evaluation Initiative (DIME), World Bank. + Niall Keleher. Interviewed April 16, 2008 by JD and DS.

Rachel Glennerster, Executive Director, Abdul Latif Jameel Poverty Action Lab, MIT. Interviewed March 21, 2008 by JS and DS.

Ruth Levine, Vice President, Center for Global Development. Interviewed April 16, 2008 by JD and DS.

Robert D. van den Berg, Director, Evaluation Office, Global Environmental Facility. Interviewed April 15, 2008 by DS.

Teunis van Rheenen, Coordinator for Partnerships and Impact Assessment, Secretary to the Board of Trustees, Director General's Office, IFPRI. Interviewed April 16, 2008 by JD and DS.