IFC International Finance Corporation World Bank Group

MONITOR

Results Measurement for Advisory Services

Innovations in Impact Evaluation in IFC

The mission of the International Finance Corporation (IFC) is to promote sustainable private sector investment in developing countries, helping to reduce poverty and improve people's lives. Evaluating the impact of IFC projects is therefore extremely important and also extremely challenging. To this end, IFC is developing and implementing practical designs for program evaluation across all IFC advisory services by drawing on best-practices and leading-edge expertise.

The gold standard for scientific evaluation is a randomized experiment in which some subjects under study (the *treatment group*) are randomly selected to participate in a program and non-participants are randomly assigned to the *control group*. There are sound methodological reasons to adopt this study design when it is practical and ethical to do so, but not all development assistance projects fit this mold. Nevertheless, institutions like IFC need to conduct rigorous program evaluations to assess progress towards institutional goals and continuously improve the quality of advisory services provided. IFC's Results Measurement effort is moving beyond conventional "before-vs.-after" analyses to develop and implement a diverse portfolio of experimental and quasi-experimental evaluation designs, which address the issue of attribution, and are consistent with widely-accepted best-practices.

While experimental and quasi-experimental methods have typically been implemented in the areas of health and education, IFC is among the first to apply these methodologies in the realm of private sector development. Particular attention is given to (1) evaluating pilot projects prior to roll-out and replication, and (2) projects that require testing several approaches to identify which is most effective. The goal of these evaluations is to learn as much as possible and then close the information loop by channeling important lessons back to project staff and improving project design and implementation.

This *Monitor Note* discusses the key challenges and some examples of innovative approaches to measuring impacts in private sector development projects in IFC.

Addressing Different Forms of Bias: The Biggest Challenge

The bane of any evaluation is bias, which can occur for a wide host of reasons and take many different forms. To mitigate the risk of bias from internal sources, IFC frequently hires external consultants, NGOs or universities to collect data and analyze results. Evaluation processes are devised to reduce the likelihood of conscious or unconscious cherry-picking to produce desired results. For example, in an evaluation in which farmers were randomly selected for assistance (such as the *Vinnyitsa Dairy* project, Ukraine,discussed below), the consultant hired to analyze the data was blinded to which was the treatment group and which was the control group until after all modeling decisions were made. As such, these decisions were based purely on methodological considerations, entirely independent of an interest in producing a particular result.

Another pernicious form of bias is *selection bias,* which occurs when participants and non-participants are systematically different in some way that could have produced different results across the groups *even in the absence of the program*. For example, if the evaluation is a study of a training program (such as *Business Edge,* discussed below) and participants self-selected into the program, it is quite possible that the more motivated and informed participants would have done better anyway and any superior post-program performance observed among participants was really driven by their superior pre-program characteristics.

Threats posed by selection bias are always cause for concern, especially when participation is not randomly assigned. These threats may be mitigated by making special assumptions (backed by evidence whenever possible) about the units under investigation in the given

study. For example, matching is a method for program evaluation that obviates the threat of selection bias vis-à-vis the construction of a control group of non-participants who match participants on all relevant characteristics (such as the *Romania Dialysis* project, discussed below). In practice, it is rarely possible to construct a perfectly matched control group, or even to measure all possible relevant characteristics. Nevertheless, matching is widely regarded as a rigorous methodology when evidence is available to show that treatment and control groups are similar enough to produce a close approximation to the perfect matches.

Experimental and Quasi-Experimental Design

In addition to matching, there are many different types of quasi-experimental designs, and not all are created equal. Some require stronger assumptions than others. For every quasi-experimental design there is a particular type of evidence that can bolster the credibility of the result. It is often possible to be able to prove that one's results still stand even if there is some selection bias.

Whatever the methodology applied, best practice in program evaluation requires a clear statement of assumptions and the use of coherent frameworks for analysis. Whenever possible, IFC goes beyond simple before-after comparisons, because this approach implicitly assumes no changes would have been observed in the absence of the intervention. This assumption is almost always highly implausible unless there is hard evidence to support it.

A year ago, IFC's Results Measurement Group began building a portfolio of evaluations consistent with these principles. Evaluation designs are reviewed by experts from academic institutions such as MIT, Harvard, Chicago, Warwick, and are implemented jointly with local partners, NGOs, universities, consulting firms and our colleagues within the World Bank Group. Below, we highlight IFC's innovative evaluation activities across all advisory business lines and, where available, discuss results and lessons learned from our experience.

Business Enabling Environment

Business Enabling Environment (BEE) programs are designed to improve the overall environment for doing business in client countries. Activities include identification and diagnostic assessments of business environment problems, improving legal and institutional frameworks, facilitating trade, and establishing commercial dispute resolution mechanisms. Approaches to evaluating Business Enabling Environment Programs range from simple before-after comparisons to more involved quasi-experimental control-group methods, encouragement (instrumental variable) designs, and randomization.

Alternate Dispute Resolution, Balkans and Pakistan

IFC launched alternate-dispute-resolution (ADR) initiatives in the Balkans in 1993 as an alternative to the regular court system, which is costly, slow, and lacks the public's confidence. These initiatives assisted with legal reforms, public education, establishing sustainable mediation centers, and training expert mediators. An external independent ex-post project review demonstrated ADR's overall effectiveness: more than 80% of ADR firms resolved their cases within 3-5 years through mediation, whereas less than 50% of non-ADR firms resolved their case within the same time frame. The mediation process itself typically took less than 15-days, and more than 80% of litigants who selected ADR reported that they were highly satisfied and would readily use ADR again.

Success in the Balkans prompted the project's replication in *Karachi, Pakistan,* and this time the evaluation was designed prior to implementation. Following the classic *encouragement* (also known as instrumental variable) design, firms are being randomly selected to receive encouragement (in this case, information and training) to participate in mediation. To estimate project impact on litigation time and cost, the values of these indicators will be compared across groups with and without encouragement, including those participating and not-participating in ADR. Even though the "treatment" (ADR) has not been randomized (because this would be highly impractical), randomizing encouragement to pursue ADR is sufficient to mitigate the threat of selection bias. And because the encouragement is randomized, we will also learn about the impacts of IFC's information and training activities.

Business Registration Simplification, Peru and Brazil

In Lima, Peru, IFC designed a reform program to reduce the time and cost required for obtaining a municipal license. Following the reform, an independent before-after analysis demonstrated that time required to obtain a license fell from 160 days to 3 days, and the total cost of registration fell by more than 60%. The overall impact was an increase in registrations by 260% from the prior year. In the wake of the Lima results, IFC's Latin America Facility chose to rollout a similar Business Registration Simplification Program in Sao Paolo, Brazil to ascertain the impact of reforms on promoting the formalization of the informal sector. Unlike the Lima evaluation design, which was a before-after analysis, in Brazil the IFC will undertake a *randomized* design. Before-after approaches cannot identify treatment effects without extremely strong assumptions, and randomization renders these assumptions are no longer necessary. Firm-level surveys will be implemented before and after in each treatment and control Sao Paolo sub-prefecture.

Infrastructure

IFC is the only multilateral institution to offer governments direct advisory services relating to private sector participation transactions. The Advisory Services Transactions Group provides support on these transactions and advice on policy and regulatory issues, optimizing commercial value while remaining sensitive to the relevant economic and political contexts. In addition, IFC also provides advisory services in the areas of health and education.

Education-Kenya Schools

IFC is implementing a School Support Program in Kenya to facilitate private schools' access-to-finance in order to promote their sustain-

able growth and improve educational services. To evaluate program impact we are partnering with the World Bank on an encouragement (instrumental variable) design, whereby a random sample of schools will be selected for direct assistance and training. By monitoring the performance of schools that do and do not receive encouragement, including those with and without financing, we will be able to estimate the effects of both the training and the financing and avoid problems posed by selection bias.

Health-Lesotho Hospitals

In Lesotho, IFC is supporting the construction of a major new hospital and evaluating two types of impacts. First, we are using a *differences-in-differences* (DID) approach to estimate the impact on revenues of firms that are providing services to the new hospital. We measure the relevant indicators before and after the intervention for participating firms and non-participating firms, and then measure the differences in time across both groups. Secondly, we will estimate the broader public health impacts of the new facility, comparing the situation on the ground before and after. This evaluation will be the second for this project; the first baseline was conducted in 2003 by experts from the Boston University School of Public Health who will continue to work with us in Lesotho.

Health-Romania Dialysis

In Romania, IFC was engaged in the privatization of dialysis services in eight hospitals located in rural and urban settings. IFC is currently in the design stage of an evaluation that will apply the matching methodology. This methodology involves matching each of the eight privatized clinics with one or more similar clinics, so that when we look at post-program outcomes we can be confident that differences between participating and non-participating clinics are driven by IFC's involvement. Importantly, this evaluation will leverage information and expertise accumulated in a prior study of this project completed in 2005. We will confirm that recommendations in that study were followed-up and appropriately addressed.

Access to Finance

IFC provides Access to Finance Advisory Services to the financial sector in order to facilitate the development of effective institutions within individual financial intermediaries and related institutions (eg., credit bureaus). This includes programs intended to improve the enabling environment for financial markets, facilitating market conditions amenable to increased future investment. More than 70% of these advisory services focus on micro, small, and medium entrepreneurs (MSMEs).

MSME Finance, Africa

IFC's MSME Finance Program is a major initiative in Africa that plans to deepen and broaden the African financial sector and raise the standards of financial services to MSMEs. The program will involve 10 to 15 financial institutions (FIs) in 5 to 7 countries, invest up to $100 million in new commitments, and provide up to $15 million in technical assistance grants to improve the FIs' performance. A *quasi-experimental* evaluation design will compare the outcomes and impacts of the program at partner banks versus non-participating banks. For one client bank that has implemented reforms sequentially one branch at a time, we will also compare the impacts in terms of increasing lending in volume and quality to MSMEs across branches.

SME Banking, Bangladesh

An evaluation of IFC's SME Banking TA programs in BRAC and Citybank in Bangladesh, a multi-component program focused on improving banking advisory services, will—in addition to comparing results across affected and unaffected branches of these banks —also compare results across matching participating/non-participating banks. This is part of a broader review of bank advisory work being carried out by IFC, which includes programs in Egypt and Riyadh and the Pacific Islands.

Environmental and Social Sustainability

The goals of the environment and social sustainability business line are to incubate innovative business initiatives delivering environmental and social benefits, to demonstrate their commercial viability, and to encourage their independent replication in the private sector in emerging markets. Monitoring and evaluation approaches have included a wide range of methodologies, from before vs. after analyses to experimental designs.

Access Bank, Nigeria

IFC's advisory services to Access Bank is helping women-owned small and medium enterprises obtain access to finance by providing training to women entrepreneurs as well as guidance to bank loan officers. The evaluation was designed at program inception and involves a *quasi-experimental* comparison of participating and non-participating entrepreneurs' sales and asset growth before and after the intervention. The group of non-participating entrepreneurs will be selected with assistance from local chambers of commerce, to closely match the group receiving the loan.

DFCU Bank, Uganda

In Uganda, an access-to-finance program supporting women entrepreneurs is being evaluated via a truly *innovative randomized experiment,* whereby women who failed to qualify within the IFC partner bank's normal credit approval process will be given a random probability of receiving a loan under a special risk-sharing agreement with IFC. The growth of sales, assets, and employees will be monitored for those that obtain loans as well as those that do not. Randomized selection ensures that the ex-post differences across recipient and non-recipient groups can be attributed to the project.

Value Addition to Firms

Value Addition to Firms programs are designed to help enterprises (particularly SMEs) improve their performance, corporate governance structure, and access to capital, thereby expanding employment and income generation in local communities.

Indonesia Seaweed

In February 2004, the Program for Eastern Indonesia SME Assistance (PENSA) launched a program intended to help seaweed farmers that offered advice on planting, harvesting, and post-harvest handling. Analysis conducted by partners at the MIT Poverty Action Lab was based on a comparison of participating vs. non-participating farmers after confirming that these groups did not significantly differ across important pre-program characteristics. Additionally, randomized experiments were implemented to determine which types of technical seaweed-farming advice were likely to be most effective in specific regions. Findings from this in-depth evaluation have provided several useful lessons for IFC's agribusiness programs, including:

- It can be expensive to obtain knowledge on the current state of farming practices in a given area and learn precisely the kinds of information that are appropriate in different locations.
- Because information obtained is not always applicable across industries or locations, it can be expensive to scale-up and expand these programs.
- There may be a hidden logic to current behavior practiced by experienced local people, such that good advice is not necessarily the best advice; techniques that excel in the laboratory do not necessarily work well on the ground

Endeavor Entrepreneurship, South Africa

The evaluation for the Endeavor Entrepreneurship-Promotion Project in South Africa employs a *matching-based evaluation* design to estimate the effects of training and direct assistance on revenue and employment. Baseline data was collected from all the Endeavor applicant prior to the program's inception, and this data was used to carefully match each participating firm to one or more others that did not participate but did share similar characteristics. Matching allows for the construction of a control group that closely resembles the treated group of Endeavor firms.

Vinnyitsa Dairy, Ukraine

In Ukraine, IFC is providing direct animal husbandry assistance and training to dairy farmers and regional milk collectors in an effort to improve the level and quality of milk production. This program is vital to the well-being of Vinnyitsa, an economically strategic and struggling dairy farming region. Here a *randomized trial* is being implemented to leverage the sequential structure of group training sessions: a random selection of farmers is receiving the assistance at the beginning of the program (the treatment group) and a random sample of other farmers will receive it at the end (the control group). By tracking output and quality for both groups from program inception through completion, we should be able to obtain an unbiased estimate of program effects.

Business Edge, Cairo

Business Edge is an SME Management Training program designed to improve the performance of SME managers by developing SME-oriented management training materials and building the capacity of SME training providers. The program was launched in Asia, and after demonstrating its success in that region it has spread rapidly, with new programs now being implemented in Yemen, Lebanon, and Egypt, where IFC is partnering with Microsoft. In all three countries we have developed designs to estimate the effects of IFC's involvement; in Egypt, the design is a *randomized trial.* Two groups of trainees will be randomly picked from within the Microsoft value chain that will receive the training in two sets of waves. Baseline information will be collected for both groups at the same time (prior to the training session for the first group, which is considered the treatment group.) This treatment group will attend the training and the control group will attend a subsequent training four to six months later. Immediately prior to the control group's training session, information will be collected about the status of both control and treatment groups. Microsoft Egypt will also be interviewed to help estimate the magnitude of effects at the top of the pyramid.

Closing the Loop

The underlying goal of all the evaluations described above is to learn as much as possible and then close the information loop by channeling important lessons back to project staff, thereby improving ongoing project design, implementation, and evaluation. For example, the evaluation of IFC's *Alternate Dispute Resolution* project, launched in the Balkans, provided insight into the types of information awareness campaigns and local partnerships critical to program success, and these factors were carefully considered when planning the next generation ADR project in Karachi. Similarly for the business registration program in Lima—experience gained from this project evaluation have been instrumental in the design of IFC's new registration program in Sao Paolo, as well as in the case of the IFC-IDA Business Registration Simplification program in Nigeria. Lessons learned from specific agribusiness programs (eg., *Indonesia Seaweed, and BILT, India*), have proven to be broadly relevant throughout this business-line and additional efforts are being made to understand the true market failures in IFC sector work. In all these cases, not only have the program design and implementation improved following evaluation and replication; the evaluations have improved with experience as well.

Only by asking sharper questions, collecting better data, and applying increasingly rigorous analysis can IFC improve the quality of its advisory services and overall development impact.

For questions or additional information please contact
Results Measurement Unit
Geeta Batra, x 34442
Alexis Diamond, x 80476.

Monitor

Monitor shares key findings from in depth reviews of advisory services programs and projects conducted by external evaluators.These reviews address the relevance, efficiency, effectiveness and sustainability of the Advisory services programs.

Results Measurement Unit of the SME department
Resultsmeasurement@ifc.org