Measuring Effectiveness to Improve Effectiveness

The methods used by USAID to assess and evaluate economic growth programs are not providing adequate information for judging the effectiveness of various activities. More systematic and rigorous approaches are needed to improve program performance and provide a stronger basis for organizational learning to promote growth and prosperity.

How does USAID measure the effectiveness and impact of programs to promote economic growth? How *should* USAID do this? These issues are under active discussion within the Agency, in recognition that existing systems are not providing adequate information to determine which programs have worked or how well they have worked in helping partner countries reach a sustainable path to rapid and broad-based growth.[1] This type of information can and should be the foundation for institutional learning to improve the allocation of development assistance funds and strengthen the design of future programs.

To complement other briefing notes in this series that present case-study evidence on the effectiveness of USAID's economic growth (EG) programs, this paper provides a summary and critique of the methods and measures that are presently used for this purpose. To balance the critique, the paper also considers how the system can be improved, drawing on examples from the Millennium Challenge Corporation (MCC) and the World Bank. The discussion focuses on EG programs but much of the analysis applies as well to other foreign assistance program objectives (National Research Council 2008, Center for Global Development 2006).

USAID'S CURRENT ASSESSMENT METHODS

USAID's field activities, including EG programs, are developed within a Country Assistance Strategy (CAS) that defines goals and objectives of U.S. foreign assistance in the country. Each field activity is supposed to incorporate a monitoring plan and a results framework linking project inputs to outputs, intermediate results (IRs), and assistance objectives (AOs). This approach to "managing for results" sets out measurable performance targets and generates systematic information on progress in achieving them. Some performance targets are defined using standard indicators from the U.S. Government's Foreign Assistance ("F") Framework (2006), while others are specific to a particular project or program. Many projects also produce short and often anecdotal success stories to convey achievements succinctly to a broader audience.

Beyond the reporting on project–level results, USAID undertakes annual or semiannual portfolio reviews to check performance against the intended results. For a subset of projects, particularly where the review finds a need for additional information, USAID conducts independent evaluations. USAID's Automated Directives System (ADS) defines an evaluation as "the systematic collection of information about the characteris-

This briefing note is part of a series produced for the EGAT Bureau at USAID as a contribution to the recurrent debate on development priorities. Each note in the Programming for Growth series examines a topic relating to the value and effectiveness of USAID's economic growth programs. All the titles in the series are listed on the last page of this note.

tics and outcomes of AO projects, or activities" (ADS 203.3.6, 2008). The ADS also offers broad guidelines on evaluation methods but states that USAID has "no standardized methodology." In recent years, most EG evaluations have been based on information from interviews and rapid or participatory appraisal techniques. They typically focus on project management, client satisfaction, and the achievement of outputs and IRs relative to targets in the monitoring plan.

Very few results reports or evaluation studies attempt to gauge the development impact of projects or activities, relative to changes that likely would have occurred even in the absence of USAID's intervention (the counterfactual). In instances where this has been done, the analysis has rarely applied a sound methodology. (See Briefing Note 8 on impact evaluation.) Consequently, the current system leaves key questions unanswered: To what extent have observed improvements been a result of the project? Would they have happened anyway? What was the actual impact in terms of benefits for the host country? Did the benefits of the project justify the cost? Could the project have delivered better outcomes per dollar spent?

USAID'S CURRENT MEASUREMENT TOOLS

The quality of USAID's monitoring and evaluation system depends not just on the methods employed, but also on the indicators used to gauge the effectiveness of aid interventions. Output indicators are generally straightforward measures of progress in carrying out project activities. The indicators used to track intermediate results are very diverse, reflecting the broad spectrum of EG activities. Some IR indicators provide useful but narrow information on outcomes, while others are highly indicative of deeper development outcomes (see Briefing Note 9, Intermediate Results).

The "F" Framework attempts to establish more systematic monitoring and reporting by mandating the use of standardized performance indicators, which are compiled through the Foreign Assistance Coordination and Tracking System (FACTS). USAID's Annual Performance Report then summarizes global progress in terms of the extent to which targets are achieved for key performance measures.

"F" Indicators

The "F" Framework identifies nearly 190 indicators for the eight program areas comprising the strategic goal of Promoting Economic Growth and Prosperity.[2] This large number of indicators stems from the inherent complexity of the development process, in that different countries require different types of interventions to overcome key constraints to broad-based growth (see Briefing Note 4 on Reforming Policies and Institutions to Foster Growth). Indeed, even this list excludes many important types of results. Hence, the need to report on "F" indicators can drive programs toward activities that may not make the most sense for the host country.

The "F" indicators focus on outputs and low-level IRs, as benchmarks of progress towards achieving higher-level outcomes. Box 1 presents some illustrative examples. The framework does not measure broader outcomes, which are influenced by many factors other than the project intervention. (In practice, even lower-level results can be affected by external factors, though project reports tend to assume that these results are attributable to USAID.) Hence, the framework does not provide a meaningful yardstick for comparing

BOX 1

Examples of Standardized Indicators for Economic Growth Programs

Illustrative examples of standardized performance indicators for EG programs at the output level include

- **The number of participants (or person-days) of training in various areas;**
- **The number of firms receiving capacity building assistance to export;**
- **The number of diagnostics conducted in various areas; and**
- **The number of policies, regulations, or procedures analyzed.**

At the results level, the standardized indicators include

- **The number of investment measures made consistent with international agreements;**
- **The reduction in costs to trade goods across borders as a result of USG assistance;**
- **The number of on-site bank examinations undertaken with USG assistance; and**
- **The number of alternative dispute resolution mechanisms put in place.**

"apples and oranges" across projects or programs, nor an overall gauge of the effectiveness of the EG portfolio.

The "F" indicators are even less useful for comparisons across strategic objectives. Aside from the obvious lack of comparability, the indicators used for economic growth programs tend to have less immediate appeal (yes, even for economists) than many of the indicators used for strategic goals involving direct benefits for individuals—for example, the number of girls completing secondary education, the number of infants vaccinated, or the number of individuals with AIDS receiving antiretroviral treatment as a result of USAID support. And yet EG programs may still be delivering vital (though indirect) benefits to the same target groups while simultaneously helping countries move toward gaining the capacity to finance high-quality education and health services themselves.[3] If the focus on incommensurate indicators is indeed a handicap in conveying the value of EG programs, then it is all the more important for USAID to develop better measures of effectiveness for these programs.

Key Performance Indicators for the Annual Performance Report

As noted above, the Foreign Assistance Framework mandates annual performance reports (APR) to promote greater transparency and accountability to the American people in the use of aid funds. At the time of this writing, the APR for FY2008 is the latest one available to the public. This report clearly summarizes a selection of performance results; as noted, however, in an external review of the APR, "the performance metrics leave considerable room for improvement" (Mercatus Center 2009, 4).

Box 2 lists the indicators used in the 2008 APR to summarize USAID performance in the eight program areas that fall under the Strategic Goal of Promoting Economic Growth and Prosperity. These indicators focus on program outcomes rather than immediate outputs. The report combines data on each indicator from USAID projects in numerous countries and compares the results to a consolidated target for the year.

The overall APR finding for the EG objective is that USAID "did not meet" the target for 8 of 12 indicators.[4] There are such serious problems with the indicators, however, that this summary statistic holds little meaning. A few examples illustrate this point:

- ***Macroeconomic Foundations.*** For lack of timely data, the designated macroeconomic indicator—budget deficit as percentage of GDP—was omitted from the overall perfor-

BOX 2

Economic Growth Indicators in the Annual Performance Report for FY2008

USAID's Annual Performance Report for FY2008 uses the following indicators, by program area, to summarize Agency performance under the Strategic Goal on Promoting Economic Growth and Prosperity:

Macroeconomic Foundation for Growth

- Three-year average in the fiscal deficit as a percentage of GDP

Trade and Investment

- Time necessary to comply with all procedures required to export and import goods

Financial Sector

- Credit to private sector as a percentage of GDP.

Infrastructure

- Number of people with increased access to modern energy services as a result of USG assistance
- Number of people with access to cellular service as a result of USG assistance
- Number of people with increased access to Internet service as a result of USG assistance
- Number of people benefiting from USG-sponsored transportation infrastructure projects

Agriculture

- Number of rural households benefiting directly from USG interventions in agriculture
- Percentage change in value of international exports of targeted agricultural commodities due to USG assistance

Private Sector Competitiveness

- Number of commercial laws put into place with USG assistance that fall in the 11 core legal categories for a healthy business environment

Economic Opportunity

- Percent of USG-assisted microfinance institutions that have reached operational sustainability

Environment

- Quality of greenhouse gas emissions reduced or sequestered as a result of USG assistance
- Number of hectares under improved natural resource or biodiversity management as a result of USG assistance

mance score. This is remarkable, given that the macroeconomic program area accounted for one-eighth of total EG spending for FY2008 (USAID 2009, 40). Even if timely data were available, however, this would be a poor indicator of the effectiveness of USAID's interventions. Well-governed poor countries often have large overall budget deficits that are fully consistent with prudent fiscal management and macroeconomic stability, because they benefit from large amounts of concessional financing relative to GDP. Conversely, a low deficit may not signal good fiscal performance. Instead it may indicate that the country has not qualified for much donor financing because of a lack of progress on reforms; or it may reflect cutbacks in spending on infrastructure, health, and education in lieu of measures to mobilize more revenue to pay for essential services. In addition, many of USAID's interventions in this program area, such as improvements in the allocation of expenditures, have effects that would not be visible in this indicator.

- ***Trade and Investment.*** The World Bank's time-to-trade (T2T) indicator is an important gauge of country performance on trade facilitation. But in the 2008 APR the indicator refers to just seven countries, including Afghanistan, where the T2T score deteriorated. With such a small sample, this one case of retrogression caused the overall result (34 days) to fall short of the target (33 days) by a hair.
- ***Financial Sector.*** Credit to the private sector as a percentage of GDP is an excellent indicator of financial sector development. But the level of this indicator is determined mainly by a country's economic structure and income level. Aid effectiveness should be measured instead by the change in this ratio, preferably calculated relative to any prior trend. Also, a single-year measure for this indicator can reflect transitory business cycle effects, including the depressing effect of the global financial crisis. In short, the indicator, as defined, says little or nothing about the results of USAID support in this important area of intervention.
- ***Private Sector Competitiveness.*** The number of new laws established with USAID support in core categories for the business environment is a highly relevant indicator in some settings. But it applies to only a limited range of USAID activities in this program area. Many improvements in the business environment involve reforming regulations and building institutional capacity. And many of USAID's programs for private sector competitiveness focus on helping industry groups or participants in a value chain to improve market linkages or access new technologies. None of these interventions, among others, are captured in a count of new laws. Worse, the prominence of this indicator in the reporting system may push some programs into working on legal reforms even though this is not a priority; if so, the effect can be to distort the program design.

Three related technicalities bear mention. First, the APR provides no information on how the targets are determined and validated. If the targets are problematic, then so are the reported hit-or-miss results. Second, the report does not explain how (or whether) the targets and results are weighted across countries or projects. Finally, the use of a "yes-or-no" standard—counting indicators for which "target missed" versus "target met or exceeded"—oversimplifies the picture by overshadowing information on how close the results were to the targets. This is especially apparent for the T2T indicator: as noted above, USAID missed the target by just one day, yet in the summary chart this shows up simply as a "miss."

The fundamental problem, however, is that none of the metrics used in the APR provide a common yardstick for determining the value of various programs or comparing performance across strategic goals and program areas. This problem cannot be solved by trying to redefine the choice of program-specific indicators. Some type of common yardstick is needed for comparing different types of apples and oranges.

APPLYING A COMMON YARDSTICK

One broadly accepted approach for applying a common yardstick to different types of aid projects or activities is to estimate the respective economic costs and benefits and then express the net benefits in terms of the rate of return on the investment. This analysis can be compared across competing uses of budget resources. In practice, it is often difficult to place an economic value on program benefits. In such cases a common alternative is to apply cost-effectiveness analysis, which compares the cost of various approaches for achieving a given objective—such as improving girls' enrollment rates or modernizing a country's budget management system. This metric, however, cannot be compared across programs with different objectives.

USAID's ADS 202: Economic Analysis of Assistance Activities recommends the methodologies outlined in the previous paragraph as a nonmandatory approach for "determining whether an activity is a worthwhile investment for the country, i.e., whether the results from an activity are sufficiently valuable as to warrant the expenditure of scarce resources." The ADS also points out that the analysis can help activity designers identify the least-cost approach to achieving programmatic objectives.

In recent years these techniques have rarely been used for USAID's EG programs, and the required skills are not widely available in the field. There are, however, discussions about revitalizing the approach as part of the Agency's effort to become a more effective learning organization, and technical training has recently been provided in this area (Harberger 2008a and 2008b).

In any case, ADS 202 refers to an ex ante analysis at the stage of project design. It does not address the need for more systematic and informative ex post evaluations to determine what has actually worked and what has not.

The MCC Approach

In the interests of "putting results in the forefront" (Weibe 2008), the MCC requires partner countries to conduct an ex ante economic analysis of each major component of a proposed compact (the MCC term for a program agreement) to demonstrate at the outset that expected results justify the proposed investments. The analysis focuses on estimating the expected rate of return on the investment of MCC funds for each major program element. Elements with a rate of return below a specified threshold—often about 10 percent— are generally returned to the sender for revision.

The ex ante analysis is only as good as the assumptions used to estimate costs and benefits, but at least they are explicit and subject to scrutiny. In many cases, the analysis has had an immediate payoff in showing where changes are needed in the program design to improve the benefits to the host country from MCC's investment. This ex ante economic analysis is also used to structure the monitoring plan for each compact in order to track progress as envisioned in the justification for the investment. For the sake of transparency and accountability, the MCC website provides public access to the calculations and assumptions for each compact.[5]

The MCC also requires an independent ex post evaluation of the economic impact of each compact, using rigorous methods wherever possible. Given the five-year horizon for MCC compacts and the fact that the MCC was established only in 2004, it will take time for the organization to compile a body of impact evaluations. But the approach is a model effort in seeking to apply a common yardstick for assessing different programs in different contexts—before, during, and after implementation.

It must be noted that the MCC's rigorous approach to economic analysis and impact assessment tends to concentrate resources on activities for which an economic analysis can be applied most readily, such as infrastructure investments. The approach is also technically demanding and costly. These costs are warranted when designing compact programs amounting to hundreds of millions of dollars, but they probably cannot be replicated with comparable sophistication for smaller or more urgent EG projects run by USAID. And as noted above, many USAID interventions involve benefits that cannot easily be quantified. Even relatively low-cost approaches, however, can provide valuable information for improving project design.

The World Bank Approach

The World Bank has a long track record of evaluating projects and programs through what is now called the Independent Evaluation Group (IEG). Like the MCC, the bank regularly conducts an ex ante economic analysis as a yardstick for ensuring that expected net benefits are positive. This analysis often influences the design of the activities.

After projects close, the bank conducts a full ex post economic analysis on a subset of activities. According to the IEG's Annual Review for 2009 (39), the Bank estimated an economic rate of return (ERR) for only about 20 percent of the projects closing in 2001, compared to 70 percent in 1970. One reason for the decline in applying this yardstick is that the bank is involved in many activities where the economic benefits are difficult to quantify.

For this reason, the IEG has other yardsticks that are used in all ex post evaluations to establish at least qualitative comparability. The IEG rates the overall performance of each project on a scale of 1 (highly satisfactory) to 6 (highly unsatisfactory), on the basis of an analysis of efficiency, efficacy, and relevance of the activities. To be specific, "satisfactory" projects are those that "fully deliver on their development objectives, appear likely to be sustained, and generate clear benefits in an efficient manner."

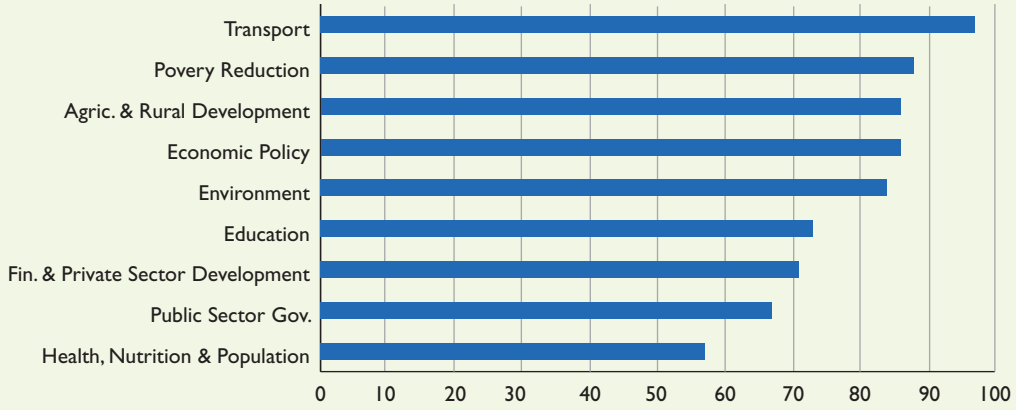
On this criterion, the IEG rated 81 percent of the projects that closed in 2008 satisfactory or better, and 19 percent unsatisfactory. Figure 1 provides a summary of the IEG ratings for 2006 through 2008, by sector (as designated by the bank). Close to 100 percent of the transportation projects were rated satisfactory or better. This is not a surprise considering that transport projects benefit from a well-tested methodology for ex ante economic analysis, which evidently has been successful in screening out bad designs. It is interesting to note that Poverty Reduction and Economic Policy projects have high success rates

by this yardstick, whereas projects aimed at financial sector and private sector development, as well as public sector governance, tend to have a lower success rate.

The IEG evaluations are similar to those conducted by USAID in that they combine subjective judgments and data analysis and focus mainly on results rather than outcomes or impacts. Unlike USAID, however, the IEG is much more systematic in evaluating projects that are closing and in standardizing the results. The bank's approach is also very different from USAID's contractor performance ratings, in that the IEG ratings are based on an independent review.

FIGURE 1

World Bank Project Ratings, Selected Sectors, 2006–2008



Source: World Bank, Independent Evaluation Group, Annual Review of Development Effectiveness, 2009

The IEG's Annual Review for 2009 cites three major lessons for improving the quality of the evaluations, all of which are relevant to USAID (http://www.worldbank.org/ieg/arde09/):

- First, many projects produce low-quality performance data.
- Second, there is often insufficient attention to data collection at the start of the project to provide a baseline for comparison.
- Third, the IEG calls for more frequent use of rigorous impact evaluations as well as cost-benefit analysis to enhance learning and improve the quality of future projects.

IMPROVING THE SYSTEM TO IMPROVE THE PROGRAMS

USAID's current systems for assessing, monitoring, and evaluating economic growth programs generate valuable information, especially on project outputs and intermediate results. Indeed, other briefing notes in this series present ample evidence from the existing systems to show that many EG programs have been highly successful in delivering strong development outcomes.

There is broad agreement, however, that the existing methods and metrics do not provide a solid foundation for institutional learning to improve overall program effectiveness. Furthermore, current performance reports have not succeeded in providing systematic information for judging the effectiveness of activities or comparing effectiveness across program areas. Indeed, it is not possible technically to use a handful of activity-specific indicators to represent broad program achievements.

The MCC and World Bank approaches discussed above suggest ways in which USAID's approach can be improved. At the design stage, even low-cost applications of standard tools for economic analysis would help to maximize prospective benefits by weeding out projects or project elements that appear to be less effective in terms of likely costs and benefits. At the close-out stage, more frequent and systematic use of independent evaluations—including rigorous impact evaluations, where possible—would provide much better information about how well various interventions have worked and why.

Finally, it is important to recognize that getting better information on project effectiveness—the focus of this note—is only half the story. As emphasized by a recent report for USAID on building knowledge to improve democracy-assistance programs, USAID also needs to strengthen systems for organizational learning in order to take full advantage of the information that is generated (National Research Council 2008). The bottom line from doing it right should be better overall program per-

formance and hence more effective use of the aid dollar for promoting growth and prosperity.

REFERENCES

Center for Global Development. 2006. When Will We Ever Learn: Improving Lives Through Impact Evaluation. May.

Commission on Helping to Enhance the Livelihood of People around the Globe. 2007. *Beyond Assistance: The HELP Commission Report on Foreign Assistance Reform.* December.

Harberger, Arnold. 2008a. Introduction to Cost-Benefit Analysis Part I: Profiles, Foreign Exchange and Capital. University of California, Los Angeles. Prepared for USAID.

Harberger, Arnold. 2008b. Introduction to Cost-Benefit Analysis Part II: Labor Market Issues. University of California, Los Angeles. Prepared for USAID.

Mercatus Center, George Mason University. 2009. ***FY 2008 Annual*** *Performance Report Scorecard: U.S. Agency for International Development.*

Modernizing Foreign Assistance Network. 2008. New Day, New Way: U.S. Foreign Assistance for the 21st Century.

National Research Council. Improving Democracy Assistance: building Knowledge through Evaluations and Research. 2008.

USAID. 2009. Fiscal Year 2008 Annual Performance Report.

USAID. The Automated Directives System, ADS Chapter 203, Assessing and Learning. 2008; and ADS 202, Economic Analysis of Assistance Activities. No date.

Wiebe, Frank. 2008. Aid Effectiveness: Putting Results at the Forefront: MCC's "New Institutional Approach". Millennium Challenge Corporation. October.

World Bank. 2009. Evaluation of Cost-Benefit Analysis in Bank Projects: Approach Paper.

World Bank. Independent Evaluation Group. 2009. "Annual Review of Development Effectiveness: Achieving Sustainable Development."

NOTES

1 See HELP Commission (2007) and Modernizing Foreign Assistance Network (2008), among other critiques of the present system. The need to strengthen evaluation systems is also a central tenet of the 2005 Paris Declaration on Aid Effectiveness.

2 The eight priority program areas are Macroeconomic Foundation, Trade and Investment, Financial Sector, Infrastructure, Agriculture, Private Sector Competitiveness, Economic Opportunity, and Environment.

3 See Programming for Growth Briefing Notes 1 and 2, on the critical importance of growth and the relationship between growth, poverty and well-being.

4 The APR summary of performance results excludes the macroeconomic foundation indicator, because time constraints precluded obtaining data beyond the period 2004 to 2006 for use in the report.

5 See: http://www.mcc.gov/mcc/results/planning/index.shtml.

The Programming for Growth Series

Overview

Briefing Notes on Basic Issues

1. Economic growth as a goal of U.S. Foreign Assistance
2. The critical role of economic growth
3. Growth, poverty, and well-being
4. Reforming policies and institutions to foster economic growth
5. Measuring effectiveness to improve effectiveness

Briefing Notes on USAID's Economic Growth Programs, with Case Studies

6. Linking people and food: The role of economic growth programs in achieving food security
7. Postconflict programming for growth
8. Economic impact
9. Intermediate results
10. Benefits over the long term

BRIEFING NOTES CAN BE READ IN ANY SEQUENCE.

CountryCOMPASS
charting pathways to ***growth***

For more information, visit **www.countrycompass.com** or contact

C. Stuart Callison, USAID Activities Manager, CAS II Project
TEL: +1-202-712-4959
E-MAIL: ccallison@usaid.gov

The Programming for Growth briefing note series was produced by Nathan Associates Inc for review by USAID under the Country Analytic Support Project (CAS II).