Using Household Surveys to Build Analytic Capacity

Lorraine Blank • Margaret Grosh

This article reviews the results of efforts in five countries to build a national capacity to analyze social policy; these efforts were undertaken in conjunction with Living Standards Measurement Survey (LSMS) projects. Case studies for Bolivia, Jamaica, the Kyrgyz Republic, South Africa, and Vietnam show that building such capacity requires explicit planning, significant time and money, open access to data, and strong support from policymakers. Lessons are drawn about four aspects of building capacity—training, technical assistance, research, and recruitment. The lessons learned from these LSMS projects apply in a more general sense to other kinds of capacity-building projects.

Strengthening the effectiveness, efficiency, and responsiveness of government is a priority for developing countries as they prepare for the 21st century. The ability to monitor the impact of economic and social reforms, evaluate the outcomes of public programs and projects, and undertake policy-relevant research is an important component of good governance, and together these institutional capabilities constitute what is referred to as the capacity for policy analysis.

Policy analysis uses multiple methods of inquiry to examine the quantifiable outputs of public policies to determine if the goals and objectives of a particular policy or program are being achieved (Waterman and Wood 1993; Dunn 1994). Such analyses are an important input into the process of formulating and evaluating government policies. The field is grounded in economics but draws upon other disciplines as well. The techniques used include quantitative modeling, statistical analysis, econometrics, qualitative research, and political and institutional analysis. This article summarizes the results of efforts in five countries to build a national capacity to analyze social policies covering welfare, education, health, fertility, nutrition, labor, and employment.

Because national capacity to undertake such operations is weak in most developing countries, the current practice is to rely largely on foreign researchers (contracted
by the government through externally funded projects or directly by international development agencies). We believe that a strong case can be made for generating good policy analysis locally, whether commissioned by the government or conducted by in-country analysts. Commissioned analysis is done at the time and on the topic that the government requires, and local researchers who know the environment are more likely to take into account domestic institutional issues and may be more knowledgeable about the full range of data that could be brought to bear on an issue. True integration of the results into the policymaking process requires a sense of ownership of the data and analysis that comes only with being part of the process from beginning to end. Local researchers may be better placed to disseminate the results and provide follow-up advice on their studies. Involving national policymakers and analysts in the process of setting the agenda for policy analysis and organizing the data collection to support it ensures that local needs are met. This means that building local analytic capacity should produce more effective social policies.

Several factors have contributed to the limited analytic capacity in many developing countries. Problems with the quality of the data are extensive. Where adequate data have been collected, opportunities for their use are often missed because researchers are given limited—or no—access. And countries often have very few researchers, especially those connected with established universities and trained in advanced statistical techniques and research methodologies. Although access to computers and software may be adequate, there is likely to be little in-country experience with statistical packages. Graduate training in policy analysis is either nonexistent or lacking in rigor. Finally, but most significantly, there is no culture of using social data to address policy questions. Policymakers have little experience posing policy questions in ways that are amenable to analysis; they also are frequently unwilling to provide open access to information. As a result government officials seldom demand policy analysis (Peterson 1991; Simonpietri and Ngong 1995; Synergie 1996; Thorbecke 1996). These factors are related in a vicious circle. If data are not collected or not released to analysts, aspiring students will not hone their skills in quantitative analysis but will move into other areas of their disciplines. If policymakers do not have access to good policy studies, they do not learn to integrate them into the policy process, and they are not interested in funding data collection efforts.

To address these limitations, international development agencies, including the United Nations, the U.S. Agency for International Development (USAID), and the World Bank, have provided financial support for household and regional surveys and other efforts to collect quality data by providing technical assistance, training, and computer hardware and software to policy analysis units (both independent and in government), research centers, and training programs. Other efforts to promote capacity have involved using local researchers in externally funded research projects and training programs (Paul, Steedman, and Sutton 1989; Grindle and Hilderbrand 1995; Myers 1997), establishing regional training programs and peer review net-
works for researchers (World Bank 1990; Fine 1995), and developing evaluation systems and a demand for evaluation (World Bank 1994; Mackay 1998).

Using Living Standards Measurement Surveys to Build Analytic Capacity

The World Bank's efforts to strengthen analytic capacity in conjunction with LSMS survey projects date from 1985 (see box 1). The skills required to implement these and other large-scale surveys—and the statistical, econometric, and modeling techniques involved—are much the same as those used to analyze social sector and economic data, so the lessons learned apply in a general sense to other kinds of capacity-building efforts as well.

Initially, LSMS survey projects focused almost exclusively on data collection; little attention was paid to building analytic capacity. In the last few years, however, task managers have begun to incorporate activities to build analytic capacity into the design of LSMS projects. Some projects do not include capacity building as central elements, but we now know more about what small efforts can do and how to craft them and are beginning to have a clearer understanding of the factors and sequencing required for the larger efforts. In this article, we focus on analyzing social sector policies, and our primary concerns are how to remedy the scarcity of policy analysts and how to encourage the use of analysis in policymaking.

The Experience in Five Countries

The case study countries—the Kyrgyz Republic, South Africa, Vietnam, Bolivia, and Jamaica—are geographically dispersed, economically and politically varied, and of different sizes. Most important, they represent a wide range of levels of capacity and different degrees of involvement in capacity building. The first Kyrgyz project

Box 1. Living Standards Measurement Surveys

The World Bank has been sponsoring LSMS surveys in developing countries since 1985, often with partial or full financing from other development agencies. These multitopic household surveys, which sample 2,000 to 5,000 households, are designed to monitor welfare and analyze behavior. The questionnaires always include comprehensive measures of consumption, usually income from employment, self-employment in agriculture or household enterprises, and transfers; access to and use of social services; and a range of outcomes and behaviors associated with health, fertility, nutrition, education, migration, savings, and housing. Community questionnaires gather information on the local economic environment, the availability of services, and prices of basic goods. See Grosh and Glewwe (1998) or the LSMS Home Page, http://worldbank.org/lsms/lsmshome.html, for a more detailed description.
was not designed to build analytic capacity at all but is typical of projects where a single analytic imperative drives a survey project. The first Vietnam project was similarly focused on data collection but thoughtfully used small amounts of money for capacity building. In South Africa another single-survey effort did involve a large number of potential future analysts in the survey design and dissemination plan. In contrast, Bolivia and Jamaica were multiyear, multimillion-dollar efforts designed to build analytic capacity.

**Kyrgyz Republic**

The first Kyrgyz Multi-Purpose Survey (KMPS) was implemented in 1993 to help the government and other local institutions assess the poverty and employment situation in order to design a program of targeted social assistance. The goal was to collect national data as quickly as possible. No budget was specifically allotted for building analytic capacity.

At the time, the country's existing capacity was limited to implementing nonrandom household surveys that were designed and subsequently analyzed in Moscow. Local researchers were not trained in quantitative analysis and had no experience in analyzing large data sets, either in government or in academic institutions. Computers were used, and programmers were available but were unfamiliar with statistical programs. Policymakers did not know how to use the survey to inform policy decisions. Driven by the need to design a social safety net project as quickly as possible, the survey project did not include building local capacity for data collection or analysis. The construction of a sampling frame would, however, provide the sampling basis for future collection of survey data.

The results of the analysis were presented at a seminar in September 1994, attended by 50 senior officials from government ministries, trade unions, and universities. Participants were invited to request specific tabulations related to their individual spheres of interest, and these were prepared and distributed the following day. The presentation of these customized tabulations generated considerable excitement among the attendees, who recognized the potential contribution of such analysis to policy decisions. This experience helped pave the way for a second, multiyear project with explicit goals for building analytic capacity. This project funded four rounds of the survey—two in 1996, one in 1997, and one in 1998. A formal user committee representing the social sector, economic ministries, and Goskomstat (the Kyrgyz statistical agency) reviewed the content of each questionnaire. The Research Triangle Institute, an independent nonprofit research and development organization headquartered in North Carolina, provided technical assistance, including training in data collection and analysis. Experts worked with staff from Goskomstat and government ministries to produce poverty profiles and statistical abstracts. Selected staff from these agencies participated in short in-country
and overseas training programs in survey research, data processing and analysis, construction of consumption aggregates, poverty measures, anthropometric measurement, and sampling theory. The project cost $3.7 million, of which $545,400 was for training.

Was the program successful? Although local staff participated in the design of each survey, their contribution was limited by their lack of experience with the uses of survey data. As anticipated, they were more active in the production of the 1996, 1997, and 1998 poverty assessments and helped to develop pension models based on the 1996 survey.

Two factors have limited the development of Kyrgyz analytic capacity: first, the statistical methods that had been in place for 70 years were considered perfectly adequate by policymakers, and second, Kyrgyz analysts were suspicious of the household survey because it had been designed, processed, and analyzed in a foreign country. Despite the declaration of open access to the data, researchers were not familiar enough with the new instrument to forsake traditional data sources. Thus it was hard to get people to demand training to use that data. Moreover, policymakers have not requested much information collected from the surveys. Although efforts to stimulate demand were made immediately after the first survey, they have had little lasting impact. Observation missions to other countries are scheduled, but these efforts are coming late in the project. Finally, it is generally agreed that in the rush to implement and analyze the surveys, technical assistance has been directed more at completing analytic tasks than at transferring analytic skills—a common phenomenon in survey projects.

**South Africa**

This 1992 survey was undertaken to create a credible national database on poverty for use in monitoring poverty, analyzing policy, and designing programs. A related objective was to strengthen the capacity for these tasks among individuals who were likely to join a new majority-rule government. The African National Congress and the Congress of South African Trade Unions requested the project, which was conducted outside of government. The Labor and Development Research Unit at the University of Cape Town coordinated the project. Originally envisaged as a two-phase project, the first step focused on upgrading and filling the gaps in national poverty information and the second phase on strengthening the capacity for continuous poverty monitoring. In the end, only the first part was implemented. The project cost about $1.5 million, of which approximately $500,000 went to activities most closely linked with building analytic capacity, including workshops, regional studies, and publication of the survey abstract.

The survey questionnaire was designed in a series of workshops attended by more than 30 social scientists over a one-year period. Most of the fieldwork was carried out.
in the latter part of 1993. The data were available for the new majority-rule govern-
ment in 1994.

Although there are 22 universities in South Africa, and many researchers have
been educated abroad, few individuals have been trained in quantitative analysis.
As a result, there were few analysts, either in government or in academic institu-
tions, who were able to analyze household data sets. This was especially true among
black South Africans. Three workshops were held to train researchers in process-
ing and analysis, and training in reading the data was provided to staff members at
several government ministries in 1995 and on a demand (and fee-paying) basis
after that.

Access to data was completely open, and a very active data dissemination program
was implemented. Within a year after the survey was completed, data sets had been
distributed to 13 government agencies, 10 universities (7 of the 10 were traditionally
white), 9 independent research institutes, international aid agencies, and various
local and international individuals. By 1996 more than 36 research reports had been
produced from the survey. The findings were used to inform government decisions
regarding unemployment programs for women and pension programs for the eld-
erly; a white paper on water and sanitation; and a report on key indicators of poverty.
And the data were being used in university courses to strengthen interest and skills in
policy analysis.

As a result, the country’s capacity to implement a national household survey im-
proved in a “learning by doing” framework with technical assistance from World
Bank staff in questionnaire design, sampling techniques, fieldwork, data processing,
and preparation of the statistical abstract. The capacity to analyze the data for basic
information such as that contained in the statistical abstract was enhanced through a
process that began with debating many versions of the questionnaire in many re-
gions of the country. This experience provided an orientation for social scientists
who had not previously worked with data from an integrated household survey, and
it supported the Central Statistics Office’s efforts to undertake household surveys
that more or less duplicate the objectives of the LSMS survey.

Policy and research analysts in South Africa began talking about household data
sets and the analytic capacity necessary to use them. As a result capacity-building
efforts have been included in a three-year Mellon Foundation project that came on
stream in 1998. The goal of the project is to increase the number of social scientists,
policy analysts, and others with the skills needed to design, implement, and analyze
data. The project will fund graduate work in demography at the University of Cape
Town, including long- and short-term training in quantitative social science meth-
ods as well as the West Cape Area Study, which will be implemented with technical
assistance from the University of Michigan.² The area study will be used as a training
tool for graduate students as well as a mechanism to generate reliable data. Total
project costs will amount to $330,000.
The challenge is to respond to the demand for increased quantitative analysis by helping the government establish a more permanent system for monitoring poverty and to increase the capacity for policy research and policy analysis in relevant government agencies and in the less established and primarily black universities.

**Vietnam**

The first Vietnam Living Standards Survey was conducted in 1992–93 with funding from the United Nations Development Programme (UNDP). The goal was to generate a comprehensive data set to help guide the policy decisions being taken as part of the transition to a market-based economy. The project was jointly managed by the State Planning Committee and the General Statistics Office. The costs of implementing the survey were $516,000, including $80,000 for training.

At the time, the country's analytic capacity was limited to basic statistics, with little expertise in advanced analytic techniques. Where expertise existed, reliable raw data and computer hardware and software were absent. Capacity-building strategies in the first project focused on preparing and disseminating the published abstract; a study mission to Thailand to observe survey procedures and to look at how the data were utilized; and short-term training (two to three weeks) in the use of statistical software, the structure and contents of the data set, and policy analysis. The project supported learning by doing through preparation of the survey abstract and working papers based on the data. The Ford Foundation funded a six-week workshop in policy analysis, statistics, computer training, preparation of research papers, and an orientation to the data, followed by a two-week writing workshop on preparing research papers. The Ford Foundation provided participants in this workshop with a computer powerful enough to use the survey data, a copy of a statistical software package, and a copy of the survey data. Project-funded training targeted policy analysts and technical staff from the central government, including the General Statistics Office, the State Planning Committee, and social sector ministries; Ford Foundation training targeted researchers from universities and research centers outside central government.³ The Ford Foundation provided $70,000 for training and equipment and $50,000 to sponsor local analysis.

The first survey resulted in some positive changes in the policy analysis environment. To understand the extent of these changes, it is important to remember that the survey broke new ground in Vietnam in several ways—it was the first household survey with all the unit record data on computer, it was the first survey that included training programs for ministry staff, and it was the first time that the statistical package had been used in Vietnam.⁴ That the statistical abstract was produced, largely by Vietnamese analysts, just six months after the data were collected was a significant achievement.

The initial survey and surrounding activities created demand for further training in data analysis. A second survey was fielded in 1997–98 as part of a project to

---

*Lorraine Blank and Margaret Grosh* 215
develop capacity to collect, analyze, and disseminate social statistics data. The project is funded by the UNDP and the Swedish International Development Agency (SIDA) with total project costs of $1,782,400, of which $700,000 is earmarked for training. The project is funding study tours, short-term in-country training in statistics and data analysis for participants from 22 agencies, and short-term overseas training in basic and advanced statistics, survey design and planning, sampling theory, and poverty measurement.

A wide range of international researchers and donor agencies have requested the data from the first survey, but the Vietnamese government itself has used the data only minimally. That is partly because analytic capacity is still limited but primarily because there is as yet little demand for data-driven policy analysis. Participants in training programs are rarely called upon to use their newly acquired skills once they return to their jobs.

Training has been extensive in terms of the number of institutions reached, but only a limited number of individuals within each organization is involved, and the training is fairly superficial. In the absence of long-range planning geared to developing the skills required to implement and analyze household surveys, strategies to build these skills progressively have been lacking. Consequently, training has been offered to as many participants as possible instead of being provided in-depth to a more limited number of participants. The workshops give trainees the basic technical skills necessary for descriptive analysis, but the skills required to formulate research questions and prepare research reports have yet to be acquired. The challenges for the future are to provide long-term training to participants who have been selected on the basis of ability, rather than seniority, and to develop a strategy for ensuring an ongoing supply of policy analysts.

Bolivia

The first LSMS-type survey was conducted in Bolivia in 1989. Five annual surveys followed, at a total cost of $2.6 million, including $1.5 million for training. Training was directed primarily at implementing and processing the surveys, and many of the changes that were adopted were applied in the government's subsequent household surveys. Despite the existence of the LSMS survey data and data from other sources, there was little interest in social policy analysis. Bolivian universities did not have a tradition of quantitative policy research. Similarly, there was very little analytic capacity in the social sector ministries; for example, researchers who intended to analyze Statistical Institute data would ask for tables rather than the raw data. External consultants conducted most of the studies that were done on household survey data.

By the late 1980s, however, as social policy issues rose to prominence on the national political agenda and the work programs of the international development
agencies, the gaps in in-country analytic capacity became obvious. As a result USAID undertook a $2.4 million project to build up the country’s domestic capacity through the establishment of a social policy analysis unit known as UDAPSO. The government contributed to operating costs with funding from other USAID-supported programs. SIDA paid for five staff positions through the Bolivian civil service program. The group included a multidisciplinary team of up to 12 professionals and concentrated on education, health, poverty, income distribution, and social sector expenditures.

The Harvard Institute for International Development (HIID) provided 32 months of long-term technical assistance and almost 40 months of service of short-term consultants. These consultants advised on best practices, interpretation of results, and future work and provided feedback on the quality and relevance of the work. In-house training (in the form of two-week workshops on the use of large data sets and statistical software, basic and intermediate statistics, econometrics, and other analytic tools) was also provided. Staff participated in workshops at Harvard in more general areas such as poverty measurement, education planning, and health planning. On-the-job training occurred through participation in research projects and preparation of research reports. The World Bank also provided a month of training in econometrics.

Between 1992 and 1995, these newly trained analysts produced 5 books, 2 research monographs, and 36 working documents. LSMS data were essential to a great deal of this research. The work was well received by other government agencies and international donors. Staff members participated in policy dialogues with senior government officials, and the dissemination of research allowed analysts to develop collaborative alliances with multilateral organizations.

The policy analysis unit’s early success resulted from a combination of intangible factors, including political support from the minister of planning, leadership, and esprit de corps. When the minister and director of the unit changed, most of the professional staff left, and recruiting difficulties plagued the organization. It was disbanded in 1997. Despite this setback, the government did not abandon its capacity-building commitment but instead created a division for social policy within the Ministry of Planning based on many of the former unit’s functions and staff. This move may actually increase the impact of previous capacity-building efforts because the new division is closely linked to the Ministry of Finance, providing it with the technical analysis that underlies budget decisions. As a result, it has achieved legitimacy among the sector ministries.

To ensure a continuing supply of trained policy analysts, USAID funded a project in 1994 to establish a new master’s degree program in public policy and management at the Catholic University of Bolivia. HIID has provided technical assistance for work with faculty and staff in managing the program, teaching courses for graduate students and for public and private sector managers, and acquiring library and teaching materials.
Jamaica

The Survey of Living Conditions (SLC) was designed to collect data to monitor the social impacts of public policies and the delivery of social services. The survey was first implemented in 1988 and has been carried out annually since then. By 1992 the SLC Jamaica's capacity for data collection was carried out in a smooth and orderly fashion and was providing an exceptionally rich database. In-country analysis, however, was limited to production of the annual abstracts and some studies (of varying quality) produced by local academics; few in-country researchers were capable of undertaking sophisticated analysis.

With the assistance of the World Bank and funding from the Netherlands, a new project was developed to increase the analytic capacity of the Planning Institute, the sector ministries, and the University of the West Indies. The $3.4 million project came on stream in July 1993. About $2.8 million was devoted to building analytic capacity and the rest to conducting further rounds of the SLC. The project established a social policy analysis unit in the Planning Institute; outlined a program of research; and provided technical assistance for graduate courses in statistical analysis, research methodology, and policy analysis for senior researchers in the sector ministries and for faculty at the university, as well as graduate school curriculum development. The project set up a social indicators data bank; graduate fellowships; workshops in statistics, research methods, techniques of policy analysis, and computer processing for junior research staff in the sector ministries; and seminars to introduce policymakers to the uses of data in policymaking. It also funded computer equipment and technical assistance to the Ministry of Labor and Welfare (the only ministry not receiving assistance from other donors) and support for five rounds of the SLC.

The social policy analysis unit, known as the Policy Development Unit, was established and operational for the whole life of the project and serves as an analytic and a project implementation unit. The unit has produced policy studies on poverty and has developed a methodology for tracking poverty over time. Unit staff revised and updated the poverty line studies that the government had contracted out in 1989, and they played an important role in preparing the government's poverty strategy. The SLC abstract is now produced solely by in-country analysts. The timeliness and complexity of the abstract have increased since the project's inception.

When the Policy Development Unit was first established, staff salaries were competitive with those in the public sector. It was difficult to find qualified individuals to fill the analyst posts because the pool of trained individuals was extremely small. Salaries have eroded since then, relative to government pay scales, and are no longer competitive. The turnover of professional staff has been high. Although a significant number of trained individuals is available (as a result of the project), recruiting has been difficult. This is a reflection not only of budget constraints but also of the diminished political support for the unit.
The University of Toronto provided technical assistance to upgrade graduate education in social policy analysis at the University of the West Indies. Ten-week courses provided training in statistics, multivariate analysis, research techniques, cost-benefit and effectiveness analysis, and other techniques to graduate students, faculty, and senior researchers from the sector ministries and the Planning Institute. The series ran twice. Eighteen fellowships were awarded, and a much larger cadre of graduate students participated in the courses. The courses have been made part of the permanent curriculum, but their continued delivery is not guaranteed because none of the current faculty has the required skills to teach the quantitative courses. The university is making a serious recruitment effort, but its success is not assured.

A social indicators data bank, which houses a computer laboratory and a library of relevant social sector data sets, including SLC data, was established at the university, which now covers the recurrent costs of the data bank. The data bank is encouraging the integration of hands-on data analysis in a wider range of classes than heretofore possible and is thereby encouraging quantitative work in university training broadly.

A series of in-service workshops (averaging 30 contact hours) was held for junior staff in the line ministries on the subjects of statistics, computer literacy, research methodology, and other techniques of policy analysis. The goal was to improve the ability of these staff members to report basic statistics on services delivered in their ministry. Enrollment was large, but few people took more than one course, and it is not clear that they use the training on their jobs.

The SLC has continued to be conducted every year, with decreasing reliance on technical assistance. Recruitment of qualified computer analysts has been a problem, partly because of low public sector wages and partly because very few people are qualified to work with statistical software. Seminars for policymakers were scheduled to integrate policy research into decisionmaking but were never implemented. This diminished policymakers' interest in obtaining data-driven answers to policy questions and in ensuring staff involvement in various other parts of the project.

The project was designed with funds to support an annual agenda of social research. The initial call for proposals was issued in the first year of the project. Two years after the start of the project, no research had been funded. Twelve proposals were approved in the last three years of the project, but several of these are of more modest sophistication than originally hoped for. The poor showing on this component may reflect some shortcomings in the details of how the call for proposals and the review process were done. More likely, it indicates the difficulty (exacerbated by the failure to implement the policymaker seminars) that policymakers and government analysts have in posing policy-relevant research questions and understanding how these could be answered with quantitative data. Another factor was probably the limited experience with proposal preparation among researchers (which might have been addressed earlier in the university training courses).
In the 11 years since the first SLC, the capacity for monitoring poverty and the amount and quality of quantitative policy analysis done in Jamaica have gradually increased. As a result the nature of the public policy debate about social issues has changed somewhat. Capacity has yet to reach the level of sophistication originally envisioned, however. The government has expressed its interest in continuing to strengthen policy analysis capacity, but external funding will be required.

The Jamaican experience demonstrates that local capacity to implement an LSMS survey can be strengthened and that local analytic capacity can be built, but that it is a slow and incremental process. The challenge is to ensure that these gains are sustainable, including the ability to produce a continuous supply of trained analysts and ongoing demand for policy analysis by decisionmakers.

Lessons Learned

Experience in the five countries selected indicates that integrating efforts to build analytic capacity with data collection efforts provides an immediate, relevant, and useful data source that can be used as a learning device for analysts and policymakers. Although some of these lessons may appear obvious, experience shows that they have often been ignored and are thus worth highlighting. We break the lessons into two groups: four requirements for building capacity—explicit planning, resources, an open data access policy, and demand for policy analysis—and four tools for building capacity—training, technical assistance, participation in research activities, and the establishment of policy analysis institutes.

Requirement One: Explicit Planning

Real and sustainable capacity building is unlikely to occur in the absence of clearly articulated goals and strategies. Clarity is required from the outset regarding the extent of existing in-country policy analysis capability, the types of capacity to be built, and the plan for doing so. Planning for capacity building requires that the project planners make decisions about:

- The agency or agencies to be responsible for data collection, data processing, data analysis (initially and over time), and data dissemination.
- The aspect(s) of capacity to be strengthened. For example, capacity building may focus on data collection for production of the statistical abstract; on the capacity within sector ministries to utilize survey data, operational statistics, and other data for planning purposes; or on the capacity of public sector and university researchers and private consultants to undertake more complex analyses using more advanced techniques.
statistical techniques. Capacity building may focus on each aspect individually or on a combination of these.

• The technical assistance, training and equipment that is required.

The objectives will differ from country to country. In countries like the Kyrgyz Republic and Vietnam, where there is almost no experience with unit record data, the initial goal should be to develop a cadre of government staff able to undertake simple means and cross-tabulations, to make these results available to a larger group of government staff, and to train a limited number of researchers in more sophisticated analytic techniques. Longer-range goals would be to develop more sophisticated analytic capacity among a wider number of analysts. In countries like Jamaica and South Africa, where a tradition of descriptive analysis of policy-relevant data already exists, the goal is to build a cadre of professionals able to do complex research (hypothesis testing, causal analysis, simulations of policy alternatives) on the social sectors for use in policymaking.

Activities to build capacity for social policy analysis should not be focused solely on economists. The multisectoral nature of social policy and of LSMS data means that substantive expertise in, for example, education, health, or poverty studies is required to formulate questions and to interpret data. Further, the technical skills required to analyze LSMS and other social sector data are common to several disciplines, including economics, sociology, education, public health, and management. Faculty and graduate students from the departments of economics, government, sociology, education, and public health participate in upgrading programs in Jamaica. UDAPSO included a multidisciplinary staff, with positive results. In all of the countries reviewed, the capacity-building activities successfully included staff from several agencies with different disciplinary backgrounds.

**Requirement Two: Resources**

Building analytic capacity is a slow and incremental process that takes significant investments of time and money. At least a decade is required to achieve a level of in-country policy analysis capacity that includes testing of hypotheses, formulation of models, and simulations of policy alternatives. Thus the time frame for capacity building extends beyond the life of most externally funded projects. More than one project will be required. A minimum investment of $5 million (not including survey costs) over several years can be anticipated, when the starting point includes some existing capacity in government and in local universities for at least the production of means and frequencies.

A lesson from the countries studied is that projects that support only a single round of a survey are not sufficient to upgrade analytic capacity significantly. At best, they can teach a limited number of people enough to do competent cross-
tabulations but not enough to pose policy questions or bring to bear more sophisticated analytic techniques (as the Vietnam experience illustrated). Single-round surveys can also help to build awareness of the need for a later, more extensive project and may serve as a pilot to explore particularly sensitive issues, assess risks, and fine-tune objectives and plans (as demonstrated in all five countries). While single-round survey projects will be implemented from time to time, they can provide only very modest aid in building analytic capacity.

**Requirement Three: Open Data Access Policies**

Efforts to build analytic capacity around a survey can not even begin unless and until data are broadly accessible. Both the formal policies governing access to data and the services for disseminating data sets affect the extent to which researchers will use data. Open data access is required because the bulk of policy analysis, and therefore of activities to build it, must take place outside the central statistics agency. Good policy analysis requires detailed knowledge of current and proposed policies and programs and often of advanced statistical techniques—skills that are not the bread and butter of survey institutions. Moreover, statistical agencies do not make policy. Their ability to conduct policy analysis is further limited in that they usually have only a handful of staff engaged in analysis, and they are usually subject to pressures to plan new surveys. Actors outside the statistical agency must therefore be included in programs to build analytic capacity. This, in turn, requires an open data access policy. The capacity-building program should, however, include the staff of the statistical agencies to help them be more aware of their clients' concerns and to remain credible partners in the efforts to provide good data for policy formulation and monitoring.

The importance of access to data was somewhat hidden in the presentation of the individual case studies because, in fact, all five countries instituted open data access policies and included participants from a range of government agencies. Several of the countries organized data dissemination services; South Africa and Vietnam provided copies of the data sets to the participants in the workshops and training seminars. Open access to LSMS data does not necessarily imply open access to other data sets. The government of Vietnam, for example, maintains open access to the LSMS data but continues to deny researchers access to other data sets.

**Requirement Four: Demand for Policy Analysis**

Policy analysis should be driven by the needs of policymakers. The problem is that policymakers in developing countries frequently do not know how to frame policy questions or use the results of policy analysis. A common theme in the case studies is the need to convince government officials that data can enhance public policy decisions. Making sure from the outset that policymakers are included in the network of
activities to build analytic capacity is vital. For example, a presentation of the results of even a single round of a survey can be a useful vehicle for generating excitement regarding the potential contribution of such analysis to policy decisions. At the dissemination workshop in the Kyrgyz Republic, policymakers were invited to request custom tabulations that were conducted the very next day. In South Africa the dissemination workshops were tailored to interest policymakers in the results of work that had been done on the South African data and additional examples of work in other countries that could be replicated. In Vietnam analysts participated in the training along with their own computer technicians so that the policy analysts could pose questions for which the programmers would often generate answers. One-day workshops in each ministry could also be useful in orienting individuals to data-driven policy analysis. Disseminating the survey results, including the statistical abstract and working papers, can also stimulate the interest of program managers in data-driven analysis. But it is important to remember that one-day workshops do not generate sustained demand; continued efforts are necessary to convince decisionmakers of the gains to be expected from investing in such projects.

Ideally, there is a virtuous circle among policymakers, analysts, and data collectors. Analysts will, unbidden, produce analysis useful to understanding good policy. Policymakers will be influenced by the results and will commission work on specific issues or alternative policies. Both efforts will drive good data collection. Initiating this cycle requires breaking into the loop on both the demand (policymaker) and supply (policy analyst and data collector) sides.

**Tool One: Training**

A critical mass of trained policy analysts is required before significant changes in the policy analysis environment can be effected. Because sending large numbers of individuals abroad for training is neither feasible nor desirable, creative approaches to in-country training should incorporate a mix of short- and long-term studies.

Policy analysis encompasses a range of techniques, from simple descriptive tabulations to complex statistical analysis. Most analysts, especially those in sector ministries, need to be able to analyze a mix of operational statistics as well as budget, survey, and other data for monitoring and for planning. A smaller number, including senior research staff in government and researchers in universities and other research centers, need advanced analytic skills to answer more complex research questions about the efficiency and effectiveness of programs. Training will often be needed at both levels.

Evidence from the countries studied indicates that short-term training—probably at least six weeks in duration—can help to improve a nation’s capacity for basic analysis, including the preparation of statistical abstracts and simple two-way tables, and can be relatively effective in enabling researchers to produce useful output from
raw data. The ability to formulate research questions, to undertake more complex analysis, and to prepare research reports requires more extensive formal training, however. The inclusion of such training as a component of several projects has provided a wider perspective for assessing its effectiveness. The first lesson is that the demand for training is likely to be high. Second, selection procedures for training programs should be standardized to ensure that all participants have the necessary prerequisites for participation, including some knowledge of computers and access to a computer outside of class. Third, participation should be limited to staff whose jobs require such training. Finally, participants should be required to enroll in a comprehensive upgrading program rather than in individual courses. These mechanisms will reduce class size, permit more interactive teaching approaches, and help weed out less committed participants. Judging from experience in Jamaica and Vietnam, refresher courses in mathematics and computers may be required to prepare participants for higher-level studies.

An important lesson highlighted by the Vietnam study is that trainees need to have opportunities to interact with experienced analysts. Very few participants in this study used their newly acquired skills when they returned to work; skills are likely to atrophy when this happens.

One-shot training programs, no matter what the length or the level, do not address future needs for trained analysts. Sustainable capacity building will require strengthening graduate education to ensure the availability of ongoing training in policy analysis, including training in a range of sophisticated research and statistical techniques. One relatively easy and fruitful approach is to encourage the use of the survey data in whatever undergraduate or graduate statistical or policy analysis courses exist in national universities (through an open data access policy, well-documented data sets, and perhaps the funding of special pilot projects or a data bank). Even this requires some initial analytic and processing capacity, however. Assembling a critical mass of faculty to teach sophisticated research and statistics courses may be difficult. Upgrading existing faculty is a long-term effort. At the same time, it may be difficult to recruit already trained faculty. Clearly, fortifying or establishing graduate programs will not be feasible in every country. Regional approaches to capacity building, such as the African Economic Research Consortium, which is attempting to strengthen graduate training and research in Sub-Saharan Africa, present a viable alternative (Fine 1995). An unintended benefit of the Jamaica project is that the University of the West Indies in Jamaica is a regional institution and non-Jamaican graduate students are enrolled in the upgrading program.

Tool Two: Technical Assistance

The case studies demonstrate the importance of technical assistance for many aspects of the LSMS process. They show that the programs are most effective when the exter-
nal expert works closely with a local counterpart and that care must be taken to ensure that resident technical assistants do not end up substituting for—rather than training—local staff. This last lesson implies that extreme care must be taken in writing the terms of reference for the external experts and in arranging counterparts. It also means that short-term consultants and twinning arrangements (where an external agency serves as a mentor for a local agency over several years) may be more effective than resident technical assistants in promoting improved capacity. When foreign technical assistance is used to substitute for local staff, opportunities to improve in-country capability are lost. In Bolivia the long-term technical assistance provided by HIID seems to have been effective in building capacity because the technical assistance was designed as a teaching device rather than as a mechanism to accomplish data collection and analytic tasks.

Tool Three: Research Apprenticeships

Involvement in all stages of policy research, including identifying the research questions, preparing the proposals, implementing the research, and writing up the final report, is essential for learning and applying the skills acquired in training. The research can be designed to help refine the skills of the staff members—something like an apprenticeship. Thus the technical assistance should serve a mentoring function rather than actually carrying out the research. In such project components, it would be premature to set the research agenda before some work has been done on the demand side with policymakers and some capacity building has occurred on the supply side with analysts.

Tool Four: Policy Analysis Units

Policy analysis units are difficult to staff and even more difficult to sustain. Ample lead time is required to establish new units or to upgrade existing ones. Governments are likely to have difficulty staffing policy analysis units because civil service salaries make it almost impossible to recruit qualified and experienced analysts and because the shortage of qualified personnel means that competition will be stiff for the few people with the knowledge and experience required. In most cases, it will be necessary to upgrade the current staff, even though high turnover can be expected as experienced individuals leave for better-paying jobs outside government.

If the project design includes establishment of a policy analysis unit and if this unit is to coordinate many activities with other agencies (such as universities, the statistical agency, and sector ministries), vesting the administrative and research leadership functions in two different positions may be wise, for two reasons. First, it is difficult to find the right mix of research and administrative skills and experience in one person, and second, the time spent on routine administrative duties, fire-
fighting, and interagency diplomacy may mean that the administrator is unable to devote significant time to the substance of research.

The sustainability of such efforts frequently depends on the level of political support they enjoy. A minimal level of support will be required even to formulate a capacity-building effort, but the project itself can generate further demand for data-driven analysis. A household survey is a useful mechanism for sensitizing policymakers to the types of questions that can be answered with empirical data and to stimulate interest in quantitative analysis of policy questions.

Notes

Lorraine Blank is a consultant at the World Bank, and Margaret Grosh is a senior economist in the Human Development Network at the World Bank. This paper relies greatly on five case studies and the many people who contributed to them:

- The Bolivia case study was written by Manuel Contreras and benefited from comments by Jere Behrman, Christian Darras, Vicente Fretes-Libils, George Gray-Molina, Menno Pradhan, Juan Carlos Requena, Eugene Szepesy, Rosa Talavera, and Nico van Niekerk.
- The Jamaica case study was written by Lorraine Blank, Margaret Grosh, and Pauline Knight, with comments provided by Pat Anderson, Paul Glewwe, P. B. K. Murthy, Jacques van der Gaag, and Colin Williams.
- The Kyrgyz Republic study was written by Raylynn Oliver, with comments provided by Jane Falkingham, Valentina Lomakova, Michael Mills, and Alexey Proskuryakov.
- The South Africa case study was written by Francis Wilson and Dudley Horner, with comments by Harold Alderman, Ann Duncan, Stephen Klassen, Carlo del Ninno, and Neeta Sirir.
- The Vietnam case was written by Lisa Drummond, with comments by Sara Bales, David Dollar, Paul Glewwe, Dominque Haughton, Jonathon Haughton, Jennie Litvack, Adam McCarthy, Nick Prescott, Jens Wandel, and Diep Vuong.
- The synthesis paper benefited from comments by Ann Duncan, Paul Glewwe, Jonathon Haughton, Emmanuel Jimenez, Jennie Litvack, John Newman, Raylynn Oliver, Laura Rawlings, Neeta Sirir, and participants in an informal seminar in the World Bank. It was edited by Fiona Mackintosh. The Development Research Group of the World Bank provided funding for the consultancies involved in all aspects of the study.

1. The survey was carried out under the direction of researchers from the University of North Carolina, Paragon Research International, and the Institute of Sociology of the Russian Academy of Sciences.
2. Modeled on the Detroit Area Study.
3. In Vietnam universities are teaching institutions, mainly at the undergraduate level, and normally do not have research programs.
4. In previous surveys data were collected and aggregated at the district level. Aggregate numbers were forwarded to the provincial level, where they were again aggregated and sent to Hanoi, where national-level aggregations were compiled. There was no file that contained the original information for each household.

References

The word "processed" describes informally reproduced works that may not be commonly available through library systems.


