

**PARIS21 Task Force on Improved Statistical Support for Monitoring
Development Goals**

**Household Surveys and the
Millennium Development Goals**

Juan Muñoz, Sistemas Integrales, Chile

Kinnon Scott, Development Economics Research Group, World Bank

Do not cite without permission from authors

This work does not reflect the official position of Sistemas Integrales or the World Bank. The authors would like to thank Neil Fantom and Diane Steele of the World Bank for helpful comments.

Contents

Introduction

Statistics Systems and Household Surveys

 National Statistical Systems

 Inter-linkages of Statistical System

 The Role of Household Surveys in the National Statistical System

Main Surveys and Data for the MDGs

 What Makes Household Surveys Different

 The Role of Household Surveys in Meeting the MDGs

Household Surveys: What is and What Could Be

 The Current State of Household Survey Systems

 What Would be an Ideal System

Generating an Improved National Household Survey System

 Planning and Process

 A Proto-type Model for a Proto-type Country

 Other Ingredients

Conclusion

References

Appendix 1: Common Household Surveys

Appendix 2: Millennium Development Goals, Targets, and Indicators and Their Measurement

Table 1. Main Purpose of the Household Survey Models Discussed in this Report

Table 2. Main Distinctive Methodological Features of the Household Survey Models

Table 3. Topics Covered by the Different Survey Models

Table 4. Typical Inputs Required by the Household Survey Models

Table 5. Household Surveys and the MDG Indicators

Table 6. Data Requirements of MDG Indicators from Household Surveys and Surveys Meeting these Needs

Table 7. Monitoring the MDG Indicators with a Household Survey Program

Figure 1. Logical Framework: Data Needs

Introduction

The international discussion of the need for more and better data for policymaking has reached an almost unprecedented level. International agreements on the Millennium Development Goals (MDGs) and country-level Poverty Reduction Strategies (PRSPs) have focused attention on the statistical systems that must produce the data to measure progress on the MDGs and PRSPs. Unfortunately, however, in many countries of the world, there has been a serious under-investment in the types of statistics and statistical systems that could contribute to our understanding of, and ability to, affect welfare levels. [In part this is a function of the public goods nature of statistics; the problem is worse in poor countries than in richer ones. The MDG- and PRSP-generated demand for data is specifically for detailed data in and on the poorer countries of the world---precisely where the deficiencies are the greatest. This combination of substantial new demand and limited supply has led to the current emphasis among the international community and countries on improving statistics.

Demand for data is not new, however. Societies and governments have kept detailed financial and economic records for centuries. Censuses and other household level data collection efforts have also evolved over the years, going from simple counting exercises to complex programs designed to determine the causes of individual and household behavior and the effect of government policies on the population's choices and welfare. Indeed, the current system of statistics in the world represents the result of centuries of actions, research, and methodological improvements in the world's ability to understand the interaction of various forces, be they economic or social, on the welfare of individuals, households, and countries. Any improvements in data that can be made today must take into account the systems that have been developed, the users of these systems, and both the traditional data needs and the new needs that arise.

The purpose of this report is to assess the role of household surveys in providing data for monitoring the MDGs and how this might be strengthened. Household surveys are important: over half of the MDGs can be measured using household survey data. However, while the context for this work is that of the current MDG-driven demand for data, a broader view of the national statistical system as a whole and its users is needed. Failure to recognize the place of household surveys within such a system may mean that short-run progress does not translate into long-term improvements and sustainability of statistics. Thus the first part of this report looks at the statistical systems that exist at the national level and the role that household surveys play in the production of policy-relevant data production. The second part reviews the main types of surveys that are currently implemented in many countries, looking at their purposes, uses, and key characteristics, as well as their aptness for meeting the data needs of the MDGs. In the third part we describe an optimal program of household surveys that could provide data for the MDGs along with other necessary data. and contrast this with the reality in most countries in the developing world, making recommendations for how to move from the current system to a more effective one. The report concludes with a discussion of the role of the international community in improving the household survey system: what actions could be reasonably taken and what the value added of such actions might be.

Statistical Systems and Household Surveys

Measuring and monitoring the MDGs represents only one type of demand for data. But a national statistical system has a range of purposes, goals, and requirements. In order to evaluate the role and utility of household surveys in producing information relevant to the MDGs, one has to start with an understanding of the context in which household surveys are designed and implemented. In other words, one needs to understand the interrelation of the various statistical activities that make up a country's statistical system and the fact that the value-added of each activity depends, to large extent, on the effective implementation of other activities.

National Statistical Systems

A national statistical system encompasses the production, dissemination, and use of data in a country. To give some idea of the complexity and magnitude of this undertaking, in the United Kingdom, 45 different government institutions are defined as contributors to the national statistical system; in Sweden, 25 institutions are officially designated as contributors, in 22 areas of statistics; and in the United States, over 70 agencies are involved. These contributors include line ministries, other government agencies, and economic actors, including the national statistical office (NSO) and through it, households and individuals in the country.

The bulk of the data generated within government is what is often lumped under the heading of administrative data. Administrative data, in their simplest form, are inventories and accounting exercises. These provide fundamental information on the inputs---financial, human, and otherwise---of the government programs, the activities carried out, and the basic goods and services provided. Such data are collected by the line ministries of a country as part and parcel of the programs and policies implemented by the ministries. In a well-designed system, such data are generated almost automatically as a by-product of the government's everyday activities. In other systems, specific data collection tasks are added to the ongoing activities of government agencies. What is important to note is that data generated in this form reflect the existing programs of the government. Thus in some manner, they are static forms of data production and are largely internal to the government.

Administrative data serve several functions. At the most basic level, they allow the government to monitor the use of resources and the goods and services into which these resources are converted. They also permit the analysis of the efficiencies and cost-effectiveness of government activities. While, in theory, the production of such data is fairly straightforward---each government agency has some variation of a management information system in place---the sheer volume of the data to be produced and the number of actors in the government ensures that the effective collection and use of such data are often quite complicated.

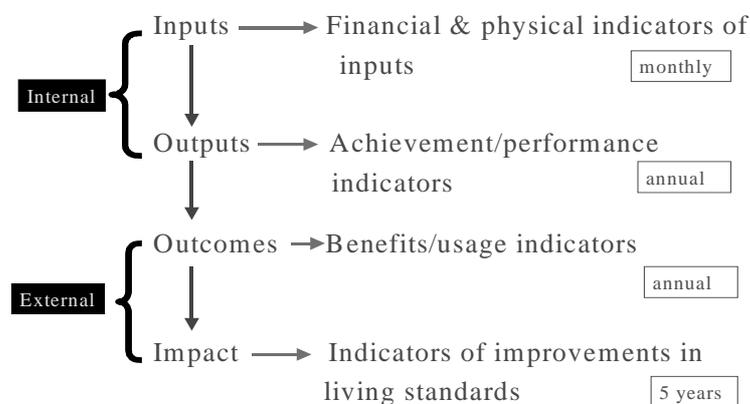
The other realm of statistics needed by governments to actually determine the effectiveness of policies and interventions is that coming from surveys of households, enterprises, agricultural production units, and individuals. Often produced by the NSO, a key player in the national statistical system, these data allow governments and the rest of society to measure the benefits produced by government action and to determine the extent to which policies are resulting in the achievements of societies' overall goals.

Unlike administrative data, data on the benefits and impacts of policies are external to the government: measurement takes place outside government agencies. The distribution and magnitude of benefits and impacts can be understood only by collecting data directly from those the government programs are designed to affect. This report focuses primarily on this latter data collection effort: surveys of households.

Inter-linkages of Statistical Systems

More important than the sheer magnitude of data produced (or needed) and used, is the fact that the various forms of data represent, not separate exercises, but a linked system---one that requires all of its parts to function properly if it is to provide governments with the data necessary for understanding the past and the present, and helping to guide the future. Think of a country as a large, enormously complicated project whose overall goal is the improvement of the welfare of the country's population. Just as a project can be conceptualized, and the data needed to measure its success identified using a logical framework, so too can the data needs of a country. For a country to achieve its goal of increasing the population's welfare, it needs data on inputs and outputs---the first two levels described in figure 1. But to know the extent to which the project is providing benefits and ultimately reaching its goal, one needs to know about outcomes and impacts. The first two levels of data are contained within administrative data. The last two depend on data from those affected by the "project": the citizenry of the country.

Figure 1. Logical Framework: Data Needs



The two types of data are not substitutes for each other, but instead form the two halves of a coherent national statistical system. On the one hand, it is worth little for governments to track spending and services produced if they have no idea if the services produced are being used, by whom they are being used, and what the short-term and medium-term effects of such services are. On the other hand, if one only knew the number and type of households using a specific good or services, a lack of administrative data would impede efforts to increase the efficiency of public spending, thus lowering the impact successful programs could have by limiting coverage and the like.

It is imperative to take into account the interrelated nature of a national statistical system. Efforts to improve national statistical systems, be they national or international, need to

recognize this. Over-emphasis on improving one type of data may lead to an actual worsening of other data collection efforts, with the subsequent result that the investment actually results in a less effective or more fragile national statistical system. This is a key point made in the evaluations of two major international programs designed to improve the production and use of data: one in Africa (Social Dimensions of Adjustment, SDA) and the other in Latin America (Improving the Surveys of Living Conditions, MECOVI).¹

An additional concern, and one that is especially valid in the current urgent discussion of how to produce MDG indicators, is that there is a trade-off between the short term and the medium and long term. The SDA evaluation report argues that, “some tradeoffs exist...between immediate data needs and the long-term objectives of institution building for data collection. Apparent short-term success may thus mask long-term social costs...The attraction of foreign resources... may result in diversion of national efforts, without leading to actual sustainable strengthening of capacities. It might therefore not be surprising if the statistical survey capacities in some participating SDA African countries turn out to be worse than before the program’s work load” (Aboyade et al, 1993).

In short, an emphasis on current data needs creates a risk that the mistakes of the past will be repeated. If resources are focused only to meet the short-term needs, or attention placed only on one part of the national statistical system, it is not clear that there will be an overall improvement in statistics. In fact, the argument can be made that it was exactly this unbalanced approach in the past that helped contribute to the current lack of data and the limited capacity of national statistical systems to respond to new data needs that are arising.

The Role of Household Surveys in the Statistical System

It is hard to overestimate the importance of household surveys for a national statistical system. There are two fundamental benefits of a functioning household survey program as part of the overall system of data collection and analysis. First, household surveys are the only source of specific types of data required by countries for effective policymaking. Second, the active nature of household surveys compared to alternative sources of data such as administrative records, provides the only viable alternative for improved policymaking in the short run. In this section we briefly describe these two, unique benefits of household surveys within the context of a national statistical system.

Unique Source of Data

Fundamentally, without household surveys a country is incapable of measuring and---more importantly---understanding critical facets of its situation and that of its population. Such a lack would severely constrain effective policymaking. Take two simple examples, poverty and unemployment. Information on the wealth of the country, either as a whole or per capita, can come from national accounts, with census data. But only household surveys can reveal the number of households living in poverty and the depth of such poverty. For unemployment, data from employment offices can specify how many people are registered for unemployment benefits and programs. But these same administrative data cannot show how many people are unemployed at any given time: registration is as much a function of the effectiveness of the

¹ For the SDA evaluation, see Aboyade et al (1993). For the MECOVI evaluation, see Ryten (2000).

employment service to help people find jobs and the generosity of benefits as it is a function of the overall unemployment level in the country.²

As importantly, or perhaps more so, household surveys are the only tool that allows governments and researchers to identify the causes of observed outcomes, and thus make possible the design of policies to change and improve outcomes. To illustrate this point, consider the difference in information that the Ministry of Health can provide in the area of malnutrition and what can be obtained from a well-designed household survey. Assume for the moment that a country has universal health coverage and all children under five years of age regularly visit medical personnel for checkups. Even in this perfect, albeit rare, case where the Ministry of Health can provide concrete data on the number of malnourished children in the country, household survey data are still a necessity. The causes of malnutrition are many, including poverty, lack of knowledge of good nutrition, and lack of access to clean water. While the Ministry of Health administrative data can specify the percentage of children suffering from malnutrition, they provide no, or very limited, insight into the causes of this malnutrition rate. In contrast, a household survey allows the researcher or policy analyst to determine the extent to which the various potential causes of malnutrition are actually responsible for the observed outcome. Armed with this information, policies to lower malnutrition rates can be formulated---policies that will have a much greater potential for cutting the rate and doing so in an economically efficient manner.

In broader terms, it is the ability of household surveys to complement administrative data that is key. Governments design policies for specific purposes, such as increasing school enrollment, improving access to markets, and increasing the participation of women in the economy, for example. Programs are universal or targeted to specific groups. Budgets are committed, programs aimed at achieving the policy goals are developed, and specific activities are carried out. But policy goals are reached not simply because the government wills them to be so, but through the interaction of the programs and household or individual behavior. Thus only through an understanding of how households behave, their constraints, and opportunities is it possible to determine which programs to implement and the characteristics these programs need to have to have an impact, and what the true impact of the programs has been. There is no substitute for household surveys in evaluating the impact of government policies. Given the magnitude of public spending, and finite resources countries have at their command, having a tool such as household surveys that allows one to increase the efficiency and effectiveness of public policy is not a luxury---as some seem to argue---but a basic need.

Household Surveys as Active Data Collection

Economic and social situations are changeable: often understanding today's situation requires data that differ from those required in the past. In short, data needs are fluid, not fixed. It is argued in this report that household surveys are the only source of data that can realistically be expected to provide a reliable flow of new data within a reasonable period---precisely because they are an active form of data collection. In other words, as the needs for new data arise, existing household surveys can be adapted to collect these new data or, in rarer cases, new survey instruments and methodologies can be developed and implemented. In this way, the time

² As much as one-third of all unemployed in the United States never register as such, the U.S. Bureau of Labor Statistics estimates (BLS 1994 and work in Central Asia shows the large gap between "registered unemployed" and unemployment rates, based on the ILO definition (Klugman and Scott 1997).

lapse between the moment of identifying the need for data and having them available to policymakers can be relatively short: on average, one to two years, and if urgent, more quickly.

In contrast, improving the quality and quantity of administrative data is a much broader proposition. Administrative data stem from either explicit data collection exercises related to ongoing activities of the line ministries, or are a by-product of routine activities at these government agencies. As such, they are embedded in the current programs, procedures, and activities of the government. Improving or changing the nature of these data often implies the need for structural reforms in the line ministries. For quality purposes, this could imply a need for improving the management information systems, improving record keeping, or training. But this is only a small part of the puzzle: what is often needed is completely new data. And such data will emerge only from changes in the programs and policies of the government. Given the complex nature of designing reforms, achieving consensus on them, and then actually implementing them, relying on changes in administrative data to provide data responsive to newly identified data needs is unrealistic. With the exception of changes at the margin, transforming administrative data to meet new data requirements is a five- to ten-year undertaking. And, even with such changes, administrative records may well still fall short of meeting demand as they typically cover only public institutions and activities; data on the private sector is often outside the purview of administrative data systems of the government.

The fact that, in practice, household survey data often become a substitute for administrative data underlines this point. Administrative data are often of limited scope and quality: the time and other costs involved in their improvement are so great as render them of only limited usefulness in the short run. Household surveys, as an active data collection method, are simply more capable of being responsive to new and changing data demands.

Main Surveys and Data for the MDGs

What Makes Household Surveys Different?

Many household surveys that are currently implemented in the developing countries can, and do, provide data for measuring and monitoring the MDGs. However, none of these surveys was developed solely for this purpose. Instead, each survey has been designed to provide to fill specific data and policy needs within the national statistical system. The end goal of each survey affects all facets of the survey, including its suitability for providing data for the MDGs. Only by understanding the differences among the current surveys and the key reasons for these differences can one begin to outline a practical program of household surveys that would meet the needs of the national statistical system and the MDGs. In this section we provide a comparative description of the major household survey models currently in place and then assess each for its ability to provide data relevant to the MDGs.

Aims and Purpose

The first and most important distinction between different household surveys refers to their overall aims and purpose. Some surveys simply strive to *measure one or more indicators at one point in time*. If repeated at regular intervals, these surveys have the potential of *monitoring progress* toward specific goals, by tracking the evolution of these indicators. A more ambitious

intention is to *understand and explain the reasons behind the observed measures*, to help governments achieve these goals efficiently.

The last objective is particularly relevant to this discussion, which focuses on the role of household surveys in monitoring a score of MDG indicators. Centering the spotlight on such a clear and explicit objective is of course extremely helpful, but this should not conceal the much wider responsibility of household surveys as a part of the policy planning process. Tracking the evolution of indicators is only part of the endeavor. Welfare is a multi-dimensional phenomenon, and without a proper understanding of the synergies underlying government actions and household behavior, monitoring can become a sterile exercise. It may be even dangerous, since indicators are only symptoms, and remedial policies targeted on changing the symptoms rather than the root causes can occasionally be counter-productive, in medicine as well as in social policy.³

Another hazard relates to the way statistical agencies look at household surveys. Twenty years ago, most statisticians explicitly considered surveys as measuring tools, exclusively. Conveying to them the enlarged notion of surveys in policy planning has been a long and difficult process that is still ongoing. The pursuit of indicators, if blindly given an overpowering status, also entails the risk of stepping backward in the survey design progress.

Table 1 summarizes the purpose of the different household survey models discussed in this report. Additional information on each survey discussed here can be found in appendix 1.

³ There are many examples, but one should suffice to illustrate the problem. Assume that Indicator 4 (the prevalence of underweight children under five years of age) is found to be 20 percent. How can it be reduced? There is a wide variety of possible actions, including improving the access to clean water (so that children do not lose weight from diarrhea), implementing maternal education campaigns, distributing oral re-hydration tablets, distributing food directly to households with children, or subsidizing certain food items, to name a few. Which of these measures would have more short-term impact? Which is more likely to be cost-effective in the long term? How can leakage be avoided by effectively targeting the action to underweight children? What are the risks of creating other problems (such as child obesity) as a result? The answer to these questions requires a lot more household-based information than merely weighing and assessing the age of children, which is all that is needed to produce the indicator.

Table 1. Main Purpose of the Household Survey Models Discussed in this Report

Survey Model	Sponsors	Main Purposes
Censuses	UNFPA UNDP	<ul style="list-style-type: none"> • Measure basic demographic and housing indicators for the country and for sub-national units every 10 years • Develop sample frames for all other household surveys
Core Welfare Indicator Questionnaires (CWIQ)	UNDP WB Africa Region	<ul style="list-style-type: none"> • Measure and monitor a limited range of welfare indicators
Demographic and Health Surveys (DHS)	USAID Macro Inter.	<ul style="list-style-type: none"> • Measure and monitor demographic and health indicators
Income/Expenditure Surveys (IES)	Central banks IMF	<ul style="list-style-type: none"> • Define the Consumer Price Indexes basket of commodities • Provide direct measures of household consumption for the system of national accounts • Define poverty lines
Integrated, Multi-Topic Surveys (LSMS/IS/FLS)*	World Bank RAND	<ul style="list-style-type: none"> • Measure and monitor all relevant welfare indicators (demographic, health, education, occupation, income, expenditure and consumption) • Define poverty lines and establish poverty profiles • Explain and model the factors underlying poverty, to guide policy programming and analysis
Labor Force Surveys (LFS)	ILO	<ul style="list-style-type: none"> • Measure and monitor employment and unemployment indicators
Multiple Indicator Cluster Surveys (MICS)	UNICEF	<ul style="list-style-type: none"> • Measure and monitor health and education indicators
Reproductive Age Mortality Survey (RAMOS)	USAID	<ul style="list-style-type: none"> • Measure the maternal mortality ratio
World Health Surveys (WHS)	WHO	<ul style="list-style-type: none"> • Measure health indicators and the responsiveness of the health system

* Living Standards Measurement Study Surveys, Integrated Surveys, Family Life Surveys

Methodological Differences

The different objectives result in important methodological distinctions between the various survey models.

Sample size and geographic disaggregation. Surveys with limited objectives, such as Labor Force Surveys (LFS) or Core Welfare Indicator Surveys (CWIQ), use shorter, less detailed questionnaires and pose fewer exigencies in terms of the selection, training, and supervision of the field staff. Their simplicity gives them the potential of visiting large samples of households in order to provide estimates for various subnational domains, as well as for the country as a whole.

More ambitious efforts such as Income and Expenditure Surveys (IES), Demographic and Health Surveys (DHS), and the Living Standards Measurement Study Surveys (LSMS) family of multi-topic, integrated surveys are much more delicate operations. These surveys rely on complex questionnaires that can be successfully administered only with high-quality field staff. Thus a greater emphasis is given to the careful selection, training, and supervision of the field staff.

Given this, these surveys are difficult to implement on very large samples and are unable to provide results for numerous subnational units.

Frequency of data collection and reporting. Surveys should ideally be repeated at regular intervals, defined primarily by analytical considerations, or—in the context of monitoring MDG indicators—by the dynamism of the indicators and by the government's ability to react to the observed changes.

Some indicators, such as the rates of unemployment or coverage of immunization, reflect economic conjunctures or short-term government actions that can be expected to change rapidly and may deserve to be observed frequently. Surveys such as the LFS are thus generally expected to be permanent rather than sporadic. In contrast, other indicators, such as access to electricity and housing condition, will change slowly and are measured adequately by the decennial census.

In between these two extremes, certain aspects of household welfare, such as the amount and composition of household budgets, and derived indicators such as poverty measures, may or may not change quickly. Poverty rates in many countries have remained static over five or more years. In these cases, surveys such as the DHS, IES, and LSMS are adequately fielded every five years. But economic crises or external shock may actually shift structural aspects of the economy or the society in the short run (consider the case of the East Asian “tigers” in the 1998 crisis or Argentina in 2002--03). Thus it is not always clear that the appropriate periodicity of implementation may need to be longer.

In practice, however, the frequency with which surveys are conducted often does not depend on analytic considerations, but on financial ones. Statistical tradition is also an important factor. For instance, in most transitional countries, income and expenditure surveys are still conducted as permanent operations, basically because that is the way they were defined at the time of the command economy.

Period of field operations. The various survey models also differ on the intended period of field operations. Most surveys, ideally, should be conducted over a 12-month period, to avoid seasonal biases—or even to measure seasonal changes explicitly, as is sometimes the case of the LFS, which is often supposed to produce quarterly estimates of unemployment.

Besides seasonality, another reason for spreading the period of data collection over several months or a year is to capitalize on the heavy investment in human resources that all surveys need to make. It is cheaper and easier to select and train a small number of good interviewers and supervisors and put them to work for many months than to train a larger number and use them for a shorter period. Longer periods of data collection also permit the implementation of more reliable quality control measures. When the period is too short, problems are often detected after the survey is finished and no corrective actions are possible.

However, some survey designers may occasionally consider that seasonal biases and substandard quality control measures are the price to pay for producing quick figures, and will try to concentrate fieldwork in a few weeks or a few months. For some data, such as unemployment rates, where a set reference week is used, the period of data collection must be quite short to avoid recall problems.

Duration of the interviews and number of visits to each household. The duration of the interview depends, of course, on the size of the questionnaire, but this is not the only factor.

In income and expenditure surveys, the very nature of the data collection technique imposes a series of visits to each household. These surveys generally ask households to use diaries to record frequent expenditures such as food. In market economies, diaries are kept for a period ranging from one week to one month, during which the interviewer is supposed to visit each household each day or every other day to check the diaries and inquire about non-food expenses with retrospective questionnaires. In former command economies, diaries are kept for much longer periods (sometimes for more than a year), and the interviewer visits each household once or twice a month.

Even when multiple visits are not required by the data collection technique, they may be imposed by other reasons. Most questionnaires contain modules that record data about individual household members. With rare exceptions (such as for small children,) these modules should ideally be administered directly to the persons themselves rather than to so-called *proxy* respondents, meaning that the interviewer should visit each household as many times as needed in order to find and interview each respondent individually.

In practice, statistical agencies face many constraints and perverse incentives that prohibit the application of personal interviewing. With the exception of the LSMS and IES, household surveys are often designed, budgeted, and conducted on the basis of single visits, and the incidence of proxy answering is far larger than it should be. Actual practice is unfortunately far from best practice in this sense.

Table 2 summarizes the main methodological distinctions between the survey models.

Table 2. Main Distinctive Methodological Features of the Household Survey Models

Survey model	Typical sampling parameters			Frequency of data collection and reporting		Period of data collection	Number of visits to each household	Duration of the interview
	Sample size (households)	Households per cluster	Geographic desegregation	Recommended	Typical			
Censuses	All households in the country		Any level	10 years	10 years in most countries	1 day to 1 month	1	Less than an hour
Income / Expenditure Surveys (IES)	2,000-20,000	10-20	3-10 regions Urban/rural ¹	5 years	5-10 years (permanent in some countries)	12 months	5-10	1-2 hours per visit
Labor Force Surveys (LFS)	5,000-50,000	10-40	5-20 regions Urban/rural	Quarterly	1-5 years (permanent in some countries)	3 months	1	30 minutes per active hh member
Integrated, Multi-Topic Surveys (LSMS/IS/FLS)	2,000-5,000	10-20	3-8 regions Urban/rural	5 years	10 years	12 months	2 or more	1-2 hours per visit
Demographic and Health Surveys (DHS)	5,000-20,000	15-30	5-20 regions Urban/rural	5 years	5-10 years	3-4 months	1	2-4 hours
Reproductive Age Mortality Surveys (RAMOS)	---	All hhs where a woman of reproductive age died	---	---	Once ²	---	1	2 hours
World Health Surveys (WHS)	2,000-3,000 (5,000 persons 18 years or older)	10-20	<3 regions Urban/rural	---	Once	3-4 months	1	1 hour
Multiple Indicator Cluster Surveys (MICS)	4,000-5,000	20-40	<5 regions Urban/rural	5 years	Twice	3 months or less	1	1 hour
Core Welfare Indicator Questionnaires (CWIQ)	5,000-20,000	20-30	5-20 regions Urban/rural	1 year	Once or twice	1 month	1	Less than 1 hour

¹ In many countries, these surveys are only administered in urban areas and thus provide no data for rural areas.

² This is a fairly new survey methodology, which explains why it has been done only once in any country.

Quality Control Mechanisms

After an evolution of almost a century, the quantitative aspects of household survey design now stand on the relatively solid theoretical grounds provided by scientific sampling theory. Other aspects of the design, such as the definition of the kind of data that needs to be collected and the search for adequate measuring techniques, are also beginning to take shape.

Progress on the qualitative and managerial aspects of survey implementation, unfortunately, has not proceeded at the same pace. Ensuring the quality of the collected data remains one of the weakest and most underestimated aspects of the endeavor.

All serious survey models recognize the importance of field supervision, and their field manuals provide useful guidelines in this regard. However, few of them put these measures into practice fully. In fact, survey budgets and schedules---plagued by inadequate funding and unrealistic deadlines---often provide all sorts of incentives for *not* supervising the fieldwork properly. The standard practice of inserting a “data cleaning” phase between fieldwork and analysis in the survey project schedule is indeed living proof that many survey practitioners appear to believe that there is no way to avoid getting dirty data from the field.

Both the DHS and the LSMS survey models consider explicit, built-in measures to ensure that field supervision is in fact implemented (in other words, to put in practice the Napoleonic principle of supervising the supervisors). In many cases, the LSMS has been able to raise the data quality standards to a level that has made data cleaning subsequent to fieldwork unnecessary. The key has been the integration of computer-based quality controls to field operations.⁴ The application of Computer-Assisted Personal Interviewing (CAPI) techniques may potentially further increase the benefits of integrating computer-based quality controls to survey fieldwork in the future.

Survey Contents

In the context of monitoring the MDG indicators, the intended content of the different survey models becomes a useful criterion for comparison. Table 3 shows the topics covered by the different survey models.

Table 3. Topics Covered by the Different Survey Models

	Demographics	Housing	Education	Health	Fertility	Anthropometrics	Occupation	Wage Income	Income from self-employment	Farm income	Other Income	Food expenses	Food consumption	Non-food expenses	Assets	Data on school attended	Data on clinic attended	Community-level data
Censuses	Basic*	Basic	Basic	---	---	---	Basic	---	---	---	---	---	---	---	---	---	---	---
Income / Expenditure Surveys (IES)	Basic	Yes	Basic	??	??	??	Basic	Yes	Basic	Basic	Yes	Yes (by diaries)	Sometimes	Yes	Often	---	---	---
Labor Force Surveys (LFS)	Basic	Basic	Basic	---	---	---	Yes	Sometimes	Basic	---	---	---	---	---	---	---	---	---
Integrated, Multi-Topic Surveys (LSMS/IS/FLS)	Yes	Yes	Yes	Yes	Yes	Often	Yes	Yes	Yes	Yes	Yes	Yes (by recall)	Yes	Yes	Yes	Often	Often	Yes
Demographic and Health Surveys (DHS)	Yes	Yes	Basic	Yes	Yes	Sometimes	Basic	---	---	---	---	---	---	---	Often	Often	??	??
Reproductive Age Mortality Surveys (RAMOS)	Basic	Basic	Basic	Yes	Yes	---	---	---	---	---	---	---	---	---	---	---	---	Yes
World Health Surveys (WHS)	Basic	Yes	Basic	Yes	Yes	Self-reported	Basic	---	---	---	---	Basic	---	Basic	Yes	---	---	---
Multiple Indicator Cluster Surveys (MICS)	Basic	Yes	Basic	Yes	Yes	Yes	---	---	---	---	---	---	---	---	---	---	---	---
Core Welfare Indicator Questionnaires (CWIQ)	Basic	Basic	Basic	Basic	Basic	Sometimes	---	---	---	---	---	---	---	Basic	Basic	---	---	---

* Basic here means that only a limited set of variables are collected on this topic- the 'basics', "yes" means that the topic is treated fully or with more detail.

Most household surveys, whatever their major subject of interest might be, collect some basic demographic data, as well as basic data on the education level and main occupation of the respondents. Most of the MDG indicators, however, require a more in-depth approach that only

⁴ Data entry and consistency controls are applied on a household-by-household basis as a part of field operations, so that errors and inconsistencies are solved by means of eventual re-visits to the households.

some of the models can provide in their current format. For instance, indicator 4 (the prevalence of underweight children under-five years of age) can be obtained only from surveys that weigh and measure children, and indicators 1 to 3 (the poverty-related indicators,) require a comprehensive measure of household consumption that only a few of the survey models can provide.

The ability of each survey model to become a platform for monitoring specific MDG indicators is explored in more detail in the third and fourth sections, below.

Resources

Survey models also differ in the amount and quality of the resources needed for their design, implementation and analysis. These can be broadly grouped into six classes: field staff wages and allowances; transportation; data entry; central management; technical assistance; and equipment, infrastructure, and other resources.

The absolute and relative incidences of these inputs in any given survey project are heavily dependent on local circumstances. They also depend on the sample size, but to a much lesser extent than might be thought, given the weight of fixed costs such as technical assistance. In spite of these difficulties, table 4 contains an effort to compare the survey models in monetary terms, to provide a general idea of the orders of magnitude involved.

The first columns of table 4 give a notion of the inputs required to conduct a 5,000-household survey of each kind, followed by the total cost and the cost per household under these conditions, assuming the unit values indicated (which are of course country-specific, but are assumed to be constant here for comparability.) The last two columns refer the unit cost to a sample size that is more typical of each specific survey model.

Table 4. Typical Inputs Required by the Household Survey Models

	Technical assistance (design)	Technical assistance (analysis)	Central management	Interviewers	Supervisors	Data-entry operators	Field-staff training	Transportation	Other	Total cost	Cost per household	Typical sample size	Cost for typical sample size
	(for a 5,000-household sample)												
	Person-months @ \$25k	Person-months @ \$25	Person-months @ \$700	Person-months @ \$400	Person-months @ \$500	Person-months @ \$400	Person-months @ \$400	Car-months @ \$1k	US\$	US\$	US\$ /hh	Households	US\$
Censuses	---	---	---	---	---	---	---	---	---	---	\$5	---	---
Income / Expenditure Surveys (IES)	6	6	90	540	120	90	60	90	\$150k	\$939k	\$188	5,000	\$939k
Labor Force Surveys (LFS)	2	2	20	60	6	10	15	20	\$50k	\$221k	\$44	10,000	\$442k
Integrated, Multi-Topic Surveys (LSMS/IS/FLS)	6	8	90	360	120	120	50	120	\$150k	\$955k	\$191	5,000	\$955k
Demographic and Health Surveys (DHS)	10	12	60	200	40	40	70	40	\$150k	\$926k	\$185	10,000	\$1,852k
Reproductive Age Mortality Surveys (RAMOS)	---	---	---	---	---	---	---	---	---	---	\$100	---	---
World Health Surveys (WHS)	2	4	20	180	20	30	30	60	\$75k	\$405k	\$81	2,000	\$162k
Multiple Indicator Cluster Surveys (MICS)	4	4	20	60	12	10	15	30	\$50k	\$334k	\$67	5,000	\$334k
Core Welfare Indicator Questionnaires (CWIQ)	3	3	20	30	3	3	15	20	\$50k	\$255k	\$51	10,000	\$510k

Some of the technical features of the survey models are so different that their inputs cannot be compared without caution. The most obvious case is that of censuses. Considering them as household surveys for these purposes is obviously a bit artificial; they are included in table 4 only to show the typical figure of \$5 per household (\$1 per person) that is often quoted for these operations.

The RAMOS surveys are also very different from the rest of the surveys, because they do not visit a nationally representative sample of households, but only those where women of reproductive age have died. The \$100 per household figure shown on table 4 is a purely notional estimate.

The surveys also differ in the way they approach data entry and data editing. Most of the surveys consider data entry as the process of transferring the data from the questionnaires into computer files, which will be later “cleaned” through office editing. Data entry is generally done with human operators. The CWIQ surveys use scanners instead of human operators to these effects, which explains its small operator/interviewer ratio.

Conversely, the operator/interviewer ratio of the LSMS surveys is high because it uses a very different approach of integrating computer-based quality controls to field operations. The data are entered into computers concurrently with the interviews, their quality is checked while the interviewers are in the field, and inconsistencies are solved by actually re-visiting the households.

The technical assistance requirements of the various survey models are also very different, and very country-specific (except perhaps for the DHS, which follows very standardized protocols, even in this respect.) Fielding a complex survey such as an LSMS or an IES will generally require at least 12 months of technical assistance. Depending on the skills available at the country's NSO, simpler surveys may require one half of that, or even less. The unit cost of \$25,000 per person-month used in table 4 is supposed to include international travel and international subsistence costs.

One cost item *not* shown in table 4 is the administrative effort of the international sponsor. This cost is accounted for explicitly in the DHS surveys (as a 35 percent overhead in the agreements between USAID and the consulting firm Macro International), but it is only implicit in the other survey models.

The Role of Household Surveys in Monitoring the MDGs

Household survey data are of critical importance for monitoring progress on the MDGs. Data for more than half of the 48 MDG indicators can come from household survey data. And, perhaps most importantly, ten of the indicators cannot be calculated at all in the absence of household survey data.. (See table 5 for details. A complete list of the MDG goals, targets and indicators can be found in appendix 2.) Without household survey data, it is not possible to construct any of the key poverty indicators related to the goal of eradicating extreme poverty and hunger, nor could indicators concerning the share of the population with access to improved sources of water, condom use rate, and the prevalence of malaria prevention activities be estimated or monitored.

Sixteen of the indicators can be constructed using either household survey data or other sources of data such as administrative records and vital statistics. In many countries, these administrative systems are weak or non-existent. Where this is the case, and where administrative data do not exist, the dependence of a country on data generated by household surveys will be even greater.

Table 5. Household Surveys and the MDG Indicators

- Indicator can *only* be measured with household survey data
-) Indicator can be measured with household survey data or other data source
- Indicator *cannot* be measured with household survey data

GOAL	Indicators	How measure?
1	1. Proportion of population below \$1 (PPP) per day	●
	2. Poverty gap ratio [incidence x depth of poverty]	●
	3. Share of poorest quintile in national consumption	●
	4. Prevalence of underweight children under-five years of age)
	5. Proportion of population below minimum level of dietary energy consumption)
2	6. Net enrolment ratio in primary education)
	7. Proportion of pupils starting grade 1 who reach grade 5)
	8. Literacy rate of 15-24 year-olds	●
3	9. Ratios of girls to boys in primary, secondary, and tertiary education)
	10. Ratio of literate females to males of 15-24 year-olds	●
	11. Share of women in wage employment in the non-agricultural sector)
	12. Proportion of seats held by women in national parliament	---
4	13. Under-five mortality rate)
	14. Infant mortality rate)
	15. Proportion of 1 year-old children immunized against measles)
5	16. Maternal mortality ratio)
	17. Proportion of births attended by skilled health personnel)
6	18. HIV prevalence among 15-24 year old pregnant women)-
	19. Condom use rate of the contraceptive prevalence rate	●
	20. Number of children orphaned by HIV/AIDS	●
	21. Prevalence and death rates associated with malaria	---
	22. Proportion of population in malaria risk areas using effective malaria prevention and treatment measures)
	23. Prevalence and death rates associated with tuberculosis	---
24. Proportion of tuberculosis cases detected and cured under directly observed treatment short course (DOTS)	---	
7	25. Proportion of land area covered by forest	---
	26. Ratio of area protected to maintain biological diversity to surface area	---
	27. Energy use (kg oil equivalent) per \$1 GDP (PPP)	---
	28. Carbon dioxide emissions (per capita) and consumption of ozone-depleting CFCs (ODP tons)	---
	29. Proportion of population using solid fuels	●
	30. Proportion of population with sustainable access to an improved water source, urban and rural)
	31. Proportion of urban population with access to improved sanitation)
	32. Proportion of households with access to secure tenure (owned or rented))

GOAL	Indicators	How measure?
8	33. Net ODA, total and to LDCs, as percentage of OECD/DAC donors' gross national income	---
	34. Proportion of total bilateral, sector-allocable ODA of OECD/DAC donors to basic social services	---
	35. Proportion of bilateral ODA of OECD/DAC donors that is untied	---
	36. ODA received in landlocked countries as proportion of their GNIs	---
	37. ODA received in small island developing states as proportion of their GNIs	---
	38. Proportion of total developed country imports from developing countries and LDCs, admitted free of duties	---
	39. Average tariffs imposed by developed countries on agri. products, textiles, clothing from developing countries	---
	40. Agricultural support estimate for OECD countries as percentage of their GDP	---
	41. Proportion of ODA provided to help build trade capacity	---
	42. Total number of countries that have reached their HIPC decision points and their HIPC completion points	---
	43. Debt relief committed under HIPC initiative, US\$	---
	44. Debt service as a percentage of exports of goods and services	---
	45. Unemployment rate of 15-24 year-olds, each sex and total	●
	46. Proportion of population with access to affordable essential drugs on a sustainable basis	---
	47. Telephone lines and cellular subscribers per 100 population)
	48. Personal computers in use per 100 population and Internet users per 100 population	●

Importantly, there is not one, single household survey that can meet all of the data needs involved in monitoring progress on the MDGs. Nor is it entirely feasible to imagine that one survey could be created to meet the MDG needs: the data needs are too broad and too heterogeneous. In practical terms, the data requirements of the different indicators lead to contradictory requirements in terms of survey methodology.

On the one hand, to capture data on rare events, a survey must be administered to a larger sample of households. Without such a large sample, one would not expect to capture enough cases of the outcome of interest, and the sampling errors or confidence intervals around the estimate would be so large as to nullify the value of the indicator. On the other hand, the construction of other indicators, such as poverty, has substantial and complex data requirements per household interviewed: at minimum, total consumption requires detailed data on household expenditures on food and non-food items, the value of all home production of such items, the value of gifts, donations of such items received by the household from both public and private sources, and the use value of housing as well as of durable goods.⁵ Because of the complex nature of the indicator in question, the potential for high levels of non-sampling error is high. Such error can lead to biases in the final indicator that is constructed--biases that can neither be measured nor rectified. Thus it is imperative to focus on decreasing non-sampling error in such survey efforts. But non-sampling error is inversely correlated with sampling error: as sample sizes increase, the level of non-sampling error rises. So a survey designed to produce data for money-metric measures of welfare, as needed for monitoring progress on goal 1 for the MDGs, would need to be administered to the smallest sample possible that can provide reasonable levels of precision.

In short, for several of the MDG indicators, a key requirement of the survey providing the data would be that a large sample of households be interviewed. For other MDG indicators, exactly the opposite is the case. Thus it is difficult to imagine one survey that could be used to meet all the household data needs of the MDG indicators: one would expect that a minimum of two surveys at the household level would be required. Only in the case where the indicators requiring large sample surveys could be constructed using alternative data sources (vital statistics, administrative records, and the like) could it be plausible that one survey could fill the data needs of the MDG indicators. And that survey would have to be one with primarily a poverty focus that also includes modules on the other indicators that can be constructed only from household surveys.⁶

It is possible, in the interim period—that is, after fielding a survey that allows for the adequate measurement of welfare levels—to carry out large sample surveys that, in addition to capturing data on other key indicators, also collects data on the key correlates of poverty. One could use these key correlates of poverty to estimate the poverty rate instead of measuring it directly. Clearly this method does not preclude the implementation of a survey specifically designed for poverty measurement. First, without the first poverty measurement survey, there would be no

⁵ The data requirements are equally onerous for constructing total income. Detailed information would be needed on labor income from wage employment, independent professional employment, income from businesses run by household members and from informal sector activities, the income derived from agricultural activities, home production of goods and services, non-labor income from rents, capital, and public and private transfers.

⁶ It is important to keep in mind that monitoring the MDGs is not the only goal of household surveys: decisions on the appropriate surveys will need to take into account the other, equally important, needs for household data that exist.

way to identify the relevant poverty correlates for the larger sample survey. (There is no evidence that correlates of poverty can be extrapolated across countries.) Second, the correlates of poverty change over time, so a poverty measurement survey would be needed with a certain frequency. There is no evidence to suggest that one can simply set the poverty correlates in year one and assume that these will be valid years later.⁷

Fortunately for those responsible for ensuring that the data needed for MDG indicators are produced, there are a variety of surveys that are currently being implemented, or that exist in the survey methodologist's arsenal, that can provide the data needed to construct the 50 percent of the MDG indicators that can be measured using survey data. For many indicators, more than one survey can provide the information needed. Table 6 provides a review of the data needed and the surveys that exist to provide such data. What is needed is for each country to review the package of surveys it is implementing. Countries need to determine if this is the optimal package for providing MDGs data and if there are changes that are needed to ensure that all the required data can be collected and/or if there are ways to decrease costs by avoiding duplication across surveys.⁸

⁷ An empirical study done by Ravallion (1996) using very aggregated measures highly correlated to poverty was not successful in predicting poverty two or three years later. Further work by Lokshin and Umapathi in the research group of the World Bank is being carried out.

⁸ Note that duplication across surveys can actually be very important, as it provides a way of cross-checking the data collected, identifying problems with certain data collection techniques, and/or validating the results.

Table 6. Data Requirements of MDG Indicators from Household Surveys and Surveys Meeting these Needs

% = Measures indicator

& = Could measure with some changes

☹ = Does not measure indicator

Indicators	Requirements of the survey to ensure adequate measurement	Is additional information needed	IES	LFS	CWIQ	DHS	MICS	LSMS/ IS/ FLS	RAMOS
1. Proportion of population below \$1 (PPP) per day	Total Consumption of households	PPP conversion factor	&	☹	☹	☹	☹	%	☹
2. Poverty gap ratio [incidence x depth of poverty]	Total consumption of households	PPP conversion factor	&	☹	☹	☹	☹	%	☹
3. Share of poorest quintile in national consumption	Total consumption of households	PPP conversion factor	&	☹	☹	☹	☹	%	☹
4. Prevalence of underweight children under-five years of age	Weight/height data on children < five, large sample	Anthrop. reference tables	☹	☹	☹	%	%	%	☹
5. % of pop. below min. level of dietary energy consumption	Total food consumption of households (quantities)	Caloric conversion tables	&	☹	☹	☹	☹	%	☹
6. Net enrollment ratio in primary education	Age and current enrollment in primary school	—	☹	☹	%	%	%	%	☹
7. Proportion of pupils starting grade 1 who reach grade 5	Past and current primary school enrollment	—	☹	☹	%	%	%	%	☹
8. Literacy rate of 15-24 year-olds	Ability to read and ability to write	—	☹	☹	%	&	%	%	☹
9. Ratios of girls/ boys in primary, secondary, tertiary education	School enrollment by level and gender	—	☹	☹	%	%	%	%	☹
10. Ratio of literate females to males, 15-24 year-olds	Ability to read and ability to write, gender, age	—	☹	☹	%	%	%	%	☹
11. Share of women in wage employ. in the non-agr. sector	Gender, wage-employment, sector of employment	—	&	%	&	&	☹	%	☹
13. Under-five mortality rate	Births, deaths, large sample	—	☹	☹	☹	%	%	&	☹
14. Infant mortality rate	Births, deaths, large sample	—	☹	☹	☹	%	%	&	☹
15. Proportion of 1 yr.-old children immunized against measles	Age, immunization history	—	☹	☹	%	%	%	%	☹
16. Maternal mortality ratio	Deaths, large sample	—	☹	☹	&	%	%	&	%
17. Proportion of births attended by skilled health personnel	Births, attendance by different personnel	—	☹	☹	&	%	%	%	☹
18. HIV prevalence among 15-24 year old pregnant women	Blood testing of individuals, large sample	—	☹	☹	☹	&	☹	☹	☹
19. Condom use rate of the contraceptive prevalence rate	Contraceptive use, specific question on condoms	—	☹	☹	☹	%	☹	%	☹
20. Number of children orphaned by HIV/AIDS	Parental deaths, extra-household sample	Information on adult deaths	☹	☹	☹	&	☹	&	&
22. % of pop. using effective malaria prevention/treatment	Questions on bed nets and treatment of bed nets	Data from health system	☹	☹	&	%	☹	&	☹
29. Proportion of population using solid fuels	Use of various solid fuels by households	—	&	☹	%	&	☹	%	☹
30. % of pop. w/ sustain. access to an improved water source	Access to water by source and hhd location	—	&	☹	%	%	%	%	☹
31. % of urban population with access to improved sanitation	Sanitation services /types used	—	&	☹	%	%	%	%	☹
32. Proportion of households with access to secure tenure	Home and land ownership, type of title	—	&	☹	&	&	☹	%	☹

Indicators	Requirements of the survey to ensure adequate measurement	Is additional information needed	IES	LFS	CWIQ	DHS	MICS	LSMS/ IS/ FLS	RAMOS
45. Unemployment rate of 15-24 year-olds, each sex and total	Age, gender, current employment, job search	—	&	%	%	&	ə	%	ə
47. Telephone lines, cellular subscribers per 100 population	Land line and cellular phones in household	—	&	ə	&	&	ə	&	ə
48. Personal computers in use and internet users/ 100 pop.	Ownership of computer equipment, use of internet	—	&	ə	&	&	ə	&	ə

Household Surveys: What Is and What Could Be

The Current State of Household Survey Systems

Although the instruments and methodologies to collect MDG indicators and other data from households exist, such surveys are not always implemented in ways that allow the monitoring of the MDGs or detailed analysis of household behavior. The current state of household survey data collection in many countries is far from adequate. Surveys are done rarely and used for policymaking even less frequently. Some causes of this lack of use are the very low quality of the data, a lack of comparability over time, and, in extreme cases, complete inaccessibility to the data by policymakers.

It is difficult to determine the direction of causality in explaining the paucity of quality household surveys and data in the developing world.⁹ The investment by many countries in this area of statistics is often quite low: the NSOs are often the “poor relation” among government institutions. Budgets are unrealistically small, wages are inadequate, the institution itself is held in low esteem (often seen to be vulnerable to political pressure, or not producing anything of value), staff rotation is high, and skills of existing staff are low. To make matters worse, in countries where the national accounts are the responsibility of the NSO, the lion’s share of the already limited budget is dedicated to these activities. All these factors lead to an institution that is essentially incapable, or capable only with heroic efforts, of producing quality data for policymakers.¹⁰

An untenably low level of resources is the beginning of a vicious cycle. Under-performing statistical agencies produce no outputs that policymakers value. Thus these same policymakers have little incentive to increase the budgets of the NSOs. And, worse, as the policymakers themselves have never had access to solid, reliable household data, they have little understanding of its potential value to themselves and have few skills to actually use any data that might be produced. Somehow, in the rich countries of the world, this vicious cycle has been broken: while the NSOs and census bureaus of many Western European and North American countries may complain about limited budgets, they actually have substantial resources at their beck and call.

Regardless of the causes, the net result is that there has been a severe under-investment in household surveys---their methodologies, implementation, and analysis---in developing countries. A variety of large and small efforts to rectify this situation have been undertaken. The United Nations Household Survey Capability Program, started in the late 1970s, tried to address the problems associated with quality of surveys.¹¹ The Living Standards Measurement Study of the World Bank, started in the early 1980s, addresses issues of relevance as well as timeliness of household survey data, as did the Social Dimensions of Adjustment program (UNDP, African Development Bank, and the World Bank). The joint Economic Commission for

⁹ Quality, in terms of data, encompasses the concepts of timeliness, relevance, accuracy, and cost-effectiveness.

¹⁰ It would be useful for all those involved in statistics and statistical capacity building to understand why the statistical agencies become so ill-favored and household surveys so under-funded. We have several hypotheses but no real answer; we leave this topic for another paper.

¹¹ See United Nations (1984) for an overview.

Latin America, Inter-American Development Bank, and the World Bank's Improving Surveys of Living Conditions (MECOVI) program has focused on improving the quality and infrastructure of surveys in Latin American countries. The Paris21 initiative is also aimed at increasing attention to the need for quality survey data both through advocacy and specific support under the Trust Fund for Statistical Capacity Building. All these programs have channeled (or are still channeling) significant financial resources to developing countries' statistical offices, as well technical assistance across the range of activities encompassed by household surveys.

In spite of the resources that have been invested, the current donor discussion on the lack of data for monitoring the MDGs and for PRSPs demonstrates that the effect of such programs has not been enough to produce functioning survey capabilities in many countries. As noted, the presence of these programs may actually have hindered the development of sustainable household survey systems. Statistical offices that rely on external funding from donors are not able to ensure that the right mix of surveys is carried out or that the continuity of surveys is adequate. Donor attention on any specific country has been ad hoc; as such, it has had limited effects on the long-term sustainability of household surveys and the capacity to design, implement, and analyze them.

Before completely condemning donor interventions, it is important to recognize that no one really knows what the counterfactual would have been without such interventions. In many countries, the lack of data might have been significantly worse than it is today and both governments' and the international community's understanding of the forces driving poverty substantially less. In addition, there are examples of countries where clear progress has been made and where the reliability and frequency of household surveys is adequate or better than adequate.

What Would be an Ideal System

The ultimate goal of a statistical system, including the household survey component of such a system, is to be able to respond to the evolving data needs of the government and society. In other words, the goal of the system is not to produce data X or survey Y, but instead to have the capacity to: determine data needs as or before they arise; identify the appropriate instruments to collect such data, be they administrative data collection efforts or household surveys; implement these tools in a timely manner; and analyze new data so that valuable inputs for the policymaking process are generated. The trick, of course, is how to achieve this flexible and responsive statistical system, especially in the area of household surveys.

There are three key needs to have such a flexible and responsive system. Two are well-known: maintaining close links between data users and data producers, and making basic resources available. One is often implicit, but should be addressed explicitly: maintaining a continuous program of survey activities. Only if proper attention is paid to all three of these facets can there be any degree of certainty that a household surveys system will be able to provide the data needed by government in general and, in particular, the data relevant for the MDGs.

Linking Data Users and Producers

Much has been written in recent years on the need to increase the links between the users and the producers of data. When trying to operationalize the Fundamental Principles of Official Statistics, Willem de Vries (1998) argues under “relevance,” that there need to be mechanisms in place that allow user and producer interaction, as well as mechanisms to assess user satisfaction with the statistical system. In the design of the MECOVI program, the emphasis on user-producer linkage was incorporated into all aspects of the program, from creating data user groups in country to interact with the data producers to ensuring a balance of users and producers in the training courses. The evaluation of the Peruvian MECOVI program underlines the importance this focus had on the success of the program (Ryten 2000b).¹²

Providing Adequate Resources

Obviously, without resources, NSOs and other parts of government cannot carry out the data collection efforts required. The discussion of the cost of data collection is less strident less with administrative data, in part because a good system has the data collection effort so integrated into the work program that its costs are marginal or, at least, unobserved. In terms of household surveys, however, the costs are very observable: the oft-heard complaint of governments (and donors) is how expensive these are. It is important to put this discussion in context. A 10 percent mistargeting of a program costing \$10 million means a misuse of \$1 million every year. A household survey done once that costs \$1 million could save an amount equal to its cost every year. A major impact evaluation of a large Social Investment Fund in Nicaragua cost only 1 percent of the funds that had been spent on the project and led to substantial reallocation of resources (World Bank 2002). In other words, relative to the benefits, the costs of surveys are often trivial: a point that needs to be made by NSOs and others engaged in the budget discussions.

Creating Continuous Survey Programs

A system of household surveys has particular needs if the system is to be responsive and capable of evolving with changing data needs. A fundamental need is that there be a program of surveys: one that has a continuous element to it. This does not necessarily mean that a permanent, never-ending survey be in place---although in some cases this might be useful. More broadly, it suggests that the part of the NSO responsible for surveys have a continuous work program of designing, implementing, and, to some extent, analyzing surveys. In poor countries with weak capacity, this may mean one well-done survey implemented every two years. In larger, more substantial statistical offices, this may entail a series of three or four surveys that might even, at some point, be in the field concurrently. Or it might mean one survey that is continuously under implementation, with the design of the next round and the analysis of the previous round being done simultaneously.

One can be reasonably agnostic as to which is the preferred option, but not on the issue of continuity itself. There are five major impacts that a continuous system of surveys can have on

¹² See also Scott (2003) and Scott, Steele, and Temesgen (forthcoming) for a detailed discussion of the benefits of linking producers and users.

the quality of the statistical system, its ability to be responsive to new needs, and its sustainability. First and foremost is the impact that such a program will have on capacity building within the NSO. Improving survey quality and staff skills is an iterative process: one learns by doing as much as by theory. The worst-case scenario in a statistical office is when a new survey is done and time and effort invested in training the range of staff and contractors in the methodology, but there is no program in place to replicate the survey. In five years time, another source of funding is identified and a repeat survey planned. But at this point, there is no one left in the survey unit or even the NSO who was involved in the first survey: each survey then becomes an exercise in starting from scratch. In this situation, only the bare minimum of skills can be generated and there is no scope for building upon previous training, applying experiences and lessons learned to the new survey. In addition, there exist few incentives to staff and managers to invest in specialized skills in survey methodology and analysis: resources can be used to more effect elsewhere.

A second reason for focusing on the continuous program is simply one of economies of scale. There are savings to be made in terms of capacity building, but also in the purchase of equipment and software, and in the creation of master sample frames, among others. Investments in one survey can be spread over several surveys, thus lowering the unit cost of each. In short, a continuous system better meets the cost-effectiveness goal of the “Fundamental Principles of Official Statistics.”¹³

Third, a continuous survey program makes it possible to provide policymakers with a plan of data availability, including both the type of data to be collected (topics, variables, indicators) and the time when the data will be available. By so doing, such a program can actually increase the use of data. There will be a move away from previously used but inadequate sources of information (unemployment registries in favor of survey data, for example) as policymakers realize that they can rely on the household data to be available at certain times and to be of comparable content and quality. For the recalcitrant policymakers or politicians, pressure to use the new or better data will arise, as others are aware that better data exist. In either case, not only is the demand for data increased, but new incentives exist to acquire the skills needed to use the data appropriately.

A fourth benefit of having a continuous system in place is that it affords the government and the statistical office a defense against outside pressures for ad hoc surveys. If such a system were in place, the potential for it be hijacked by the latest donor request for a survey or the latest fad is limited. This is not to say that special surveys are not needed, of course. This is the whole point of a responsive and flexible system: it can undertake new surveys when the need exists. But the need declines: much of the outside demand for data may well be met by the continuous survey system, thus reducing the pressure to “drop everything” and start a new survey from scratch. Moreover, with minor amendments, and at marginal cost, the existing survey may be capable of meeting additional data demands.

In the case where a new survey is truly needed, or a major change in methods is indicated, a NSO with an ongoing, continuous data collection exercise in place can actually be more

¹³ This list was adopted by the Economic Commission for Europe in 1992 and later endorsed by the United Nations (see de Vries 1998).

responsive and flexible to new data demands as they arise. Staff are better trained, are more experienced, and thus can recognize when new data demands require new surveys or methods. The needs for external assistance to actually plan and implement such changes should be lower. The infrastructure is in place and up-to-date, (be this the sample frame, rosters of interviewers, or even financial procurement systems) so that responding to new data demands, even if these require a new survey, is simply more feasible and less costly.

Generating an Improved National Household Survey System

How can countries move toward a more responsive and flexible survey system that provides the data needed for today—for all users, not just for MDG monitoring---and that also has the capacity to produce data for the future? Certainly there is no one-size-fits-all solution, nor is there one package of surveys that one can recommend be implemented. Instead, the emphasis needs to be on the process—short-, medium-, and long- term—by which a useful stream of data can be produced for policymakers.

Planning and Process

This process can be best started at the moment in which a global strategic plan for the national statistics system is being developed. As noted, household surveys are just one component of the statistical system: designing the system of surveys in isolation or separate from the global planning exercise is almost guaranteed to lower the value of the surveys produced as well as the coherence of the national system.

It is not clear, however, that the developers of national statistical plans always recognize the role of surveys in the statistical system. Often strategic plans are left in the hands of those whose expertise---and interest---is in the area of macroeconomic and financial statistics. The result is that micro data needs are underplayed; there is often the quality of an after-thought to the household surveys sections of master plans.

Thus the first suggestion for increasing the value of household surveys is to ensure that planning of these is an integral part of the master plan. This will, most likely, require forming teams comprised of macro and micro specialists to draw up the master plan. For countries receiving technical assistance, this may entail more cost, as two experts---one knowledgeable in macro data, and the other of micro data---may be needed instead of one person. It seems a small cost for the potential benefit.

In the process of drawing up a master plan for statistics, an important first step is identifying the data needs and the clients or users of data. This may be a difficult exercise, as one would, ideally, like to identify not just actual users but those who would be users if adequate data were forthcoming. Additionally, some assessment of the alternative data sources needs to be made. Such questions as the following need to be addressed: What alternative data sources exist? What is the quality of these sources? Can these be improved and how? What are the relative costs of producing specific data by different means? Are there no alternatives?

This process is liable to be lengthy and require substantial levels of consultation, discussion, and negotiation. To the extent that this results in a viable plan, it is worth the time and effort. The process itself can have the additional benefits of engaging (and educating) users, of allowing (or forcing) the macro and micro statisticians in the country to learn more about the other types of, and needs for, data in a national system, and some of the tradeoffs involved. In short this strategic plan process, if done well and seriously, can help to create a vision of a national statistical system among the disparate actors benefiting from or responsible for statistical data.

Once the inventory of data demand exists, a clear prioritization of the needs must be undertaken: no country has the resources to meet all data needs. Each country's situation will decide the priority data needs. For middle-income countries, the MDG goals may already have been met. In such cases, it would be hard to justify devoting scarce resources each year to demonstrate what is already known. For PRSP countries, demonstrating that poverty alleviation is actually occurring has critical implications for future international funding: thus the indicators listed in the PRSP and the goals contained in it will take priority over other data concerns. And, while some types of data may be hugely important to some economies, in others, where there are few or weak policy levers to affect the observed outcomes, it may make more sense to focus attention on areas where policy levers do exist.

A Prototype Model for a Prototype Country

There are no prototype countries, of course. However, in the same manner that this study took the risk of describing the "typical" survey of each of the currently available models, in this section is presented a notion of how a system of household surveys might look for a hypothetical country that – as an element of its statistical plan – decides that it needs to provide policymakers the following minimum inputs from household surveys:

- Continued monitoring of the labor markets
- Regular information on the demand for social services and the performance of the social service delivery systems, for monitoring and design purposes
- Poverty measures, poverty profiles, poverty maps, and other elements required by the PRSPs, and
- Monitoring the country's progress toward the MDGs.

It is assumed that this prototypical country also wants to develop, strengthen, and maintain the capability of its NSO to implement household surveys and disseminate the information stemming from them.

Under such circumstances, a possible program of household surveys could be composed of the following minimum elements:

- A permanent labor force survey (LFS), capable of providing quarterly employment/unemployment figures for the whole country and annual estimates for its major sub-national units
- A regular program of Living Standard Measurement Study Surveys (LSMSs), conducted every four years

- A regular program of Demographic and Health Surveys (DHSs), also conducted every four years, in the intermediate years.

Such a system would have the capability of monitoring all 27 household-based MDG indicators, as shown in table 7. It is important to emphasize, however, that monitoring of the MDGs would not be the only output—perhaps not even the most important output—of the proposed system. Setting the focus on the monitoring of the MDGs is essential, but household surveys have a broader responsibility in the policy planning process. As argued before, simply observing the evolution of indicators without a proper understanding of the underlying synergies can become sterile—even dangerous. From an operational standpoint, it is also important that the NSO looks to surveys as more than mere measuring tools, as is still often the case in many countries.

Table 7. Monitoring the MDG Indicators with a Household Survey Program

Household-based MDG indicators (numbers correspond to international system)	Survey and Year												
	LFS 2005	LFS LSMS 2006	LFS 2007	LFS DHS 2008	LFS 2009	LFS LSMS 2010	LFS 2011	LFS DHS 2012	LFS 2013	LFS LSMS 2014	LFS 2015	LFS DHS 2016	
1. Proportion of population below \$1 (PPP) per day		▶				▶				▶			
2. Poverty gap ratio [incidence x depth of poverty]		▶				▶				▶			
3. Share of poorest quintile in national consumption		▶				▶				▶			
4. Prevalence of underweight children under-five years of age		▶		▶		▶		▶		▶		▶	
5. % of pop. below min. level of dietary energy consumption		▶				▶				▶			
6. Net enrollment ratio in primary education		▶		▶		▶		▶		▶		▶	
7. Proportion of pupils starting grade 1 who reach grade 5		▶		▶		▶		▶		▶		▶	
8. Literacy rate of 15-24 year-olds		▶		▶		▶		▶		▶		▶	
9. Ratios of girls/ boys in primary, secondary, tertiary education		▶		▶		▶		▶		▶		▶	
10. Ratio of literate females to males of 15-24 year-olds		▶		▶		▶		▶		▶		▶	
11. Share of women in wage employ. in the non-agr. sector	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	
13. Under-five mortality rate				▶				▶				▶	
14. Infant mortality rate				▶				▶				▶	
15. Proportion of 1 yr.-old children immunized against measles		▶		▶		▶		▶		▶		▶	
16. Maternal mortality ratio				▶				▶				▶	
17. Proportion of births attended by skilled health personnel		▶		▶		▶		▶		▶		▶	
18. HIV prevalence among 15-24 year old pregnant women				▶				▶				▶	
19. Condom use rate of the contraceptive prevalence rate		▶		▶		▶		▶		▶		▶	
20. Number of children orphaned by HIV/AIDS				▶				▶				▶	
22. % of pop. using effective malaria prevention/treatment		▶		▶		▶		▶		▶		▶	
29. Proportion of population using solid fuels		▶		▶		▶		▶		▶		▶	
30. % of pop. w/ sustain. access to an improved water source		▶		▶		▶		▶		▶		▶	
31. % of urban population with access to improved sanitation		▶		▶		▶		▶		▶		▶	
32. Proportion of households with access to secure tenure		▶				▶				▶			
45. Unemployment rate of 15-24 year-olds, each sex and total	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	
47. Telephone lines, cellular subscribers per 100 population		▶				▶				▶			
48. Personal computers in use and internet users/100 pop.		▶				▶				▶			

As can be seen from table 7, thirteen of the MDG indicators (those that can be obtained from both the LSMS and the DHS,) would be observed every two years. Two of the indicators (those obtained from the LFS) would be observed every year. The remaining twelve indicators would be observed every four years (seven of them from the LSMS, and five from the DHS).

Monitoring some of the indicators would require some adjustments to the prototypical LSMS or DHS questionnaires. Most of the fixes are relatively minor, but a few of them would require careful thinking and experimentation:

- Indicator 18 (HIV prevalence among 15--24-year-old pregnant women,) requires blood testing. This is something that the standard DHS does not do, but have experimented with recently.
- Indicator 16 (Maternal Mortality Ratio) could be measured by the DHS through the “indirect sisterhood method,” or directly by either the DHS or the LSMS, if procedures based on the RAMOS methodology were added to their community modules.
- Indicator 20 (number of children orphaned by HIV/AIDS) could also be measured by the DHS, with indirect methods that would have to be developed.¹⁴

A Permanent LFS as a Part of the Household Survey Program

One relevant question is why the LFS is proposed as the permanent component of a coordinated household survey program. One reason is that an LFS is important in its own right, since the permanent observation of the labor markets is a necessity for most countries. This built-in demand, usually by the Ministry of Labor or Economics, ensures a higher level of ownership within the country for the survey: there is explicit demand and clear users who will defend and support the survey. This, obviously has clear implications for sustainability, a key concern.

Second, the nature of the labor markets requires frequent measurement and monitoring. Labor force variables, unlike many other indicators found in the MDGs and elsewhere, change very rapidly. Annual data is inadequate and data aggregated over a year even less useful for labor market and economic policy. Thus having a continuous process of collecting labor force information is needed. For other variables, having indicators every two to four years is more than sufficient: change is slow and no variation will be seen within one year or even between years.

Third, a labor force survey has the advantage of being relatively simple. It is a “light” survey in the sense that it has a short questionnaire and can be administered rather quickly and inexpensively. It also tends to have large samples (in order to obtain needed levels of precision of labor statistics). All these features make it an ideal vehicle: it provides crucial data in an affordable format, it does not represent such an onerous burden on the statistical office that it cannot be maintained, and it allows a greater level of disaggregation over other, more complex surveys.¹⁵

¹⁴ To the extent that orphaned children remain in households, either the DHS or the LSMS is adequate to capture data on this phenomenon. However, in the case where a significant proportion of orphans end up in institutional care or homeless, the current surveys will underestimate the magnitude of the problem: additional techniques will need to be developed.

¹⁵ There are, of course, other light surveys, such as the CWIQ and even the MICS. The need for labor force variables at relatively short intervals and the inherent ownership of such surveys in a country make LFS preferable to other light surveys. And, if the prototype system outlined above is used, then all of the indicators that would be captured by these other, light surveys will be provided biannually with the DHS and LSMS.

Given its structure, in addition to providing permanent employment/unemployment figures, the LFS can also become a platform for launching sporadic, special-purpose investigations at a marginal cost, thus endowing the NSO with a capability to quickly respond to unforeseen demands. In the simplest scenario, such demand might be met by simply adding a few questions to the ongoing survey to meet the immediate, narrow need: evaluating the impact of an immunization campaign, for example. In the case of a more substantial data need, when the unexpected request is too complex to be answered by simply adding a small module to the LFS questionnaire, the large size of the LFS sample would permit conducting special-purpose operations on a sub-sample of the households visited by the LFS.

A word of warning is merited here, however. This ability of the LFS to help in the mounting of low-cost temporary survey operations should not be abused to the extreme of overloading a survey model that by nature needs to stay lean. We are, in particular, not recommending that the LFS ever become a permanent, multi-topic survey. The observation of complex household realities is the job of instruments designed to that effect, such as the LSMS.

Features of a Household Survey Program

A coordinated program of household surveys, extending to a horizon of 10 years or more, provides the opportunity to put into practice certain desirable features that have proven difficult to implement in most developing countries, where surveys are typically conducted as sporadic and isolated exercises. One key issue in this regard is the comparability of the indicators estimated from different surveys, or from different rounds of the same survey. In the context of a system of household surveys, it is important that all surveys use the same operational definitions and methods to measure concepts such as literacy, unemployment, and gross and net incomes. If one of the explicit objectives of the system is monitoring the MDG indicators, this is of course essential; only comparable measures can be useful to this effect.

Second, the selection of the households to be visited can be based on a continually updated master sample frame: a common set of primary sampling units used by most surveys. This can bring about economies of scale in sampling design and in the implementation of the household listing operations. It also has the potential of tracking demographic dynamics (particularly rural-to-urban migrations) more closely in-between censuses. Along this same vein, the LFS sampling design should consider some form of panel to improve the precision with which the changes (rather than the levels) of employment/unemployment figures are measured.

Variations on the Prototype

As said before, the proposed prototype is only a basic paradigm that may be adequate for a variety of countries, but certainly not for all them. Although the recurrent cost of a permanent LFS is much lower than that shown in table 4 (which includes the fixed costs associated to its design), some of the poorest and least statistically developed countries may not be able to afford conducting even such a low-cost operation on a permanent basis. The model could in these cases be reduced to an alternating sequence of LSMSs and DHSs that would provide labor (and other) statistics every two years, or to an even more infrequent sequence of an integrated LSMS/DHS

instrument. Fortunately or unfortunately, these countries are also less likely to have labor markets that deserve permanent observation, so that a more sporadic view may be sufficient.

Other countries may consider that the consumption measures provided by the LSMS are insufficient to support the needs of their system of national accounts or the development of CPI weights and may want to add an IES survey to the program, or replace some of the LSMSs by IESs with additional welfare modules.

Other Ingredients

Clearly there are other ingredients required for strengthening a household survey system and improving quality and sustainability. Taking advantage of technological advances, such as Computer Assisted Personal Interviewing (CAPI) or the use of Global Positioning Systems, using the web for monitoring and harmonization of interviewers work (as done in Peru), can lead to cost saving and improve data quality. Expanding and emphasizing training of all personnel involved in surveys, from the staff of the NSOs to field staff to analysts in ministries, would be a vital ingredient to the household survey system. These and other methods should form part of the effort to improve the ability to generate and use quality household-level data.

Another area where room exists to improve household surveys is that of donor coordination. The international roundtable meetings in Marrakech in 2004 identified six actions needed to improve the capability of countries and the international community to monitor the MDGs, as well as provide inputs to other development goals.¹⁶ One action was the development of an International Household Survey Network. Such a network could play an important role to moving toward a better system of statistics in developing countries. While an IHSN would not resolve all existing problems, it is worth noting some areas (the list is not comprehensive) where this network has potential to facilitate better statistics and household surveys.

The first task that could be undertaken (contracted) by an IHSN is to draw up specific recommendations for incorporating household surveys into countries' national statistical plans. Several key activities that need to form part of this plan in the area of surveys have been identified in this study; there are others. A clear set of guidelines, procedures, and issues to be addressed would be an excellent tool for all persons engaged in drawing up strategic plans for statistics.

A second useful function for the IHSN is that of a clearing house. Maintaining information on what has been done and what surveys are planned in an easily accessible format could help bilateral and multilateral organizations identify gaps and determine where their resources and attention could best be utilized. It would have the benefit of avoiding duplication of effort, allow donors to take advantage of economies of scale, and save the time and effort needed, on a country-by-country basis, to determine what is being done and by whom. While, on the surface, a simple task, in actuality it may require a fair amount of perseverance and creativity to be able to develop and keep up-to-date such an information system; those with the information may have limited incentives to be pro-active in providing the data to the IHSN.

¹⁶ Second International Roundtable on Managing for Development Results, February 2004.

The clearing house function can be expanded beyond simply inventorying existing and planned surveys to cover training and consultants. Significant resources are invested in training in statistics, but it is not clear how demand for training is identified by the countries. Perhaps the guidelines for developing the household survey component for the strategic plan could incorporate identifying training needs, or other mechanisms could be developed. There is potential to take further advantage of economies of scale by having more multi-country training programs (this is done now, but there is scope for expansion). To the extent that the IHSN could collect demand and match it with existing or planned training, and funding sources for participants, the efficiency of the training systems in place could be increased. The same type of mechanism could also be employed to link consultants from one developing country to another: the South-South approach that has so much to offer in skill transfer and capacity building.

Finally, the IHSN could further the advocacy role of the PARIS21 by providing support to countries in the Consultative Group Meetings. Such support could be to help present the strategy and the program of national household surveys, show a specific country's plan compares to other similar countries, and identify where the most important gaps are, and the like. In other words, provide a context for the discussion of a country's statistical system.

Conclusion

The current demand for data fostered by the Millennium Development Goals is a window of opportunity to improve the quality and sustainability of data in developing countries. There is a clearly articulated demand for data at the international and country level. There is a sense of urgency that is serving to move data improvement to the top of the priority list. And because of the international commitment to the MDGs, there is the potential for significant investments to take place.

The trick is to avoid wasting this opportunity. As has been argued in the report, efforts to improve the system of data collection and use in a country need to keep in mind four key points. First, the focus must be on the entire system of data production and use, not just one isolated element of it. Second, monitoring the MDGs is not the *raison d'être* of the statistical system. Supporting the production of MDG-relevant data should not be done at the expense of producing other valuable and necessary data for governments and policymakers. Third, one-off efforts tend to undermine efforts to build sustainable capacity. The urgency for MDG data should not distract countries and the development community from the medium- and long-term goals. Finally, planning and coordination, are necessary. Ad hoc solutions have no staying power. In short, while there are no guaranteed routes to success, there are certainly enough lessons in the field to teach us how to avoid the largest pitfalls. If countries and the international donor community fail to learn these lessons, we can look forward to an on-going, if not eternal, discussion about the weaknesses of the data without actually improving it.

Appendix 1. Common Household Surveys

Censuses

For the purposes of this discussion, censuses can be considered as household surveys with unique characteristics:

- They strive to observe all households in the country, rather than a sample of them. This gives them the capability of providing results for all subgroups of the population, particularly for small area units.
- They are seen by most governments as essential operations (often as the NSO's reason of being), which generally gives them an assured 10-year periodicity.
- Field operations are conducted during a very short period (often one day, or very few days, and seldom more than one month).
- They resort to the use of a large numbers of interviewers, which makes quality controls difficult.
- They collect data on a limited number of indicators, using short questionnaires.

Besides their role as household surveys with such special features, censuses are an essential element of a country's household survey system, because they are needed to develop sample frames for all other household surveys.

Income and Expenditure Surveys (IES)

Income and expenditure surveys were initially conceived with two principal objectives. The first is to find the shares of different commodities in the budget of households, to define the composition of the baskets used by Consumer Price Indexes. The second is to provide direct measures of household consumption for the system of national accounts. The definition of poverty lines can also be considered a third major goal of income and expenditure surveys.

Household expenses are registered with diaries and/or recall questionnaires. Diaries record each individual purchase made by the household during a specific observation period. Recall questionnaires inquire about expenditures made by the household in the past. IESs in various countries differ in the extent to which they resort to the use of each kind of instrument. During the second half of the past century, surveys in command economies relied exclusively on diaries kept by each household for at least one year, and often for much longer. Surveys in market economies generally use diaries for shorter periods (typically from 1 week to 1 month,) to record food and other frequent expenses, and recall questionnaires to record expenditures on other commodities.

IESs are generally fielded for 12 months, to take into account the seasonal variations of consumption patterns. Sample sizes range from 2,000 to 20,000 households, depending on the required level of geographical desegregation. Stratification tends to favor the observation of urban areas, to the extent of restricting the scope of the survey solely to urban areas in some countries.

In command economies, the IESs used to be a permanent effort of the national statistical agencies. In contrast, developing countries conduct IESs as sporadic exercises, seldom more often than every five years or so. This stems in part from the nature of the data collected, which is considered to reflect structural features of the economy that do not change quickly enough to deserve closer monitoring, and in part because IESs are difficult and expensive operations.

Labor Force Surveys (LFS)

The main objective of the Labor Force Surveys (LFSs) is to measure the levels of employment and unemployment. LFS often also collect data on wage incomes and may provide qualitative information on non-wage activities as well.

The core LFS questionnaires and definitions were developed by the International Labour Organization (ILO) around 1970, and they are generally applied with little variation. This permits reliable comparisons of the labor force indicators across countries and between different rounds of the survey in the same country.

Each round of the LFS is typically fielded over a short period (one to three months,) and in each round each household is visited only once. In principle, a basic questionnaire is applied to each active household member in a personal, relatively brief interview. In practice, however, the LFS is often fielded under tough constraints in terms of field staff and timing, and interviewers end up accepting proxy answers very often.

There is considerable variation on what constitutes a survey round. In some countries, the LFS is a permanent effort of the NSO, which releases employment/unemployment figures on a quarterly basis (sometimes, on a monthly basis, using the “moving quarter” averaging technique.) In other countries, the LFS is a much more sporadic survey, conducted every three to five years.

When the LFS is a permanent survey, the sampling design generally considers some form of paneling. Certain households are re-visited in two consecutive quarters and/or in the same quarter of two consecutive years. The aim is to increase the precision with which the variations of unemployment figures are estimated.

Living Standards Measurement Surveys (LSMS)

The Living Standards Measurement Study (LSMS) survey instrument was developed in the 1980s as a means to help policymakers understand the determinants of observed social and economic outcomes and, thus their ability to design effective programs and policies. The LSMS is designed to provide a comprehensive picture of household welfare and the factors that affect it. The LSMS provide analysts with data that allows them to assess: welfare levels and distribution; the links between welfare and the characteristics of the poor; the levels of access to and use of social services; the impact of government programs; and the causes of observed social outcomes.

The LSMS typically incorporates data collection at the individual, household, and community level. The survey instruments are designed in close consultation with policymakers in the

country to ensure the relevance of the data to be collected. The resulting questionnaires, especially the one at the household/individual level, tend to be quite complex. To ensure high quality data, the samples are kept small (2,000--5,000 households) to minimize non-sampling error). Substantial efforts and resources are devoted to data quality (such as month-long interviewer training, use of direct informants, concurrent data entry with in-field corrections taking place at the households, low supervisor-interviewer ratios). A final aspect of ensuring data quality is to have countries agree to open-access policies of their data sets.¹⁷

Demographic and Health Surveys (DHS)

The DHS survey model follows a long tradition of demographic research dating back to the 1970--84 World Fertility Survey (WFS) program. Its original intention was to produce internationally comparable measures of fertility, mortality, contraceptive use, maternal and child health, and other demographic indicators, but its goals have broadened to include urgent topics such as HIV/AIDS, STDs, anthropometrics, and child malnutrition; and frequent requests of the participating countries for data on access to health services and intra-familial violence.

A typical 5,000-household DHS requires about eight months of preparation (including training,) and is fielded for three months, using a staff of 50 field workers grouped into 10 teams. Preliminary results are generally produced a few weeks after the end of field operations, and full analytic reports are available six months later.

The DHSs have a well-deserved reputation for technical excellence in questionnaire and sampling design—including the routine publication of sampling errors for all relevant indicators. Data management is very standardized. It includes a well-specified set of consistency controls to be applied in the data entry and cleansing phases and it always concludes with a well-documented public-access database.

DHSs have been conducted since 1985 by Macro International, a U.S.-based firm, under successive five-year contracts with USAID.

Reproductive Age Mortality Surveys (RAMOS)

The main purpose of the Reproductive Age Mortality (RAMOS) Surveys is to evaluate the magnitude of maternal mortality and other causes of death among women of reproductive age (WRAs.)

The RAMOS survey methodology has been developed as a part of a USAID-sponsored project. It operates in two phases. In the first phase, all deaths occurring in a community within a given period are recorded from a variety of sources, including routine death registrations (if available), medical records in community health facilities, interviews with health care or social services providers, and monitoring of burials in the community graveyard. In the second phase, WRA deaths are investigated to determine the cause of each death and its relation to pregnancy. Data

¹⁷ Although the World Bank does not own the LSMS data sets, the Bank has permission to directly disseminate over half of the data sets (30 percent are available on the website for downloading). In the remaining cases, governments must approve individual requests: permission is denied in only about 10 percent of the cases.

sources at this stage include medical records and coroners' reports, interviews of health care providers, and interviews of surviving family members (verbal autopsy).

Although the verbal autopsy entails conducting household-level interviews, the RAMOS surveys are not national household surveys in the same sense as the other surveys reviewed here, because the households visited are not a probability sample of all households in the country. The surveys deserve to be considered in this discussion, however, because the RAMOS approach is considered to provide the best possible estimation of MDG indicator 16 (the Maternal Mortality Ratio, MMR,) and because it can be potentially integrated into the fieldwork of national household surveys that routinely collect community-level information, such as the LSMS or the DHS.

World Health Surveys (HWS)

The WHS is a relatively recent arrival on the scene of international household survey models. It was conceived by the WHO with the objective of providing country-level information on the health states of populations; risk factors (such as tobacco, alcohol, and pollution) and their association with health states; responsiveness of health systems; coverage, access, and utilization of key health services (such as immunization, treatment of childhood illness, STD, and HIV/AIDS); and health care expenditures.

The WHS sampling design and questionnaires are standardized. They are submitted to participating countries along with prescriptive instructions for their application. The survey has been conducted in about 70 countries but no country reports are available at this moment.

Core Welfare Indicator Questionnaire (CWIQ) surveys

The Core Welfare Monitoring Survey (CWIQ) was developed by the World Bank, with the close collaboration of the UNDP, UNICEF, and the ILO. The CWIQ is a household survey designed to measure changes in key social indicators for different population groups: specifically indicators of access, utilization, and satisfaction with core social and economic services.

The CWIQ incorporates a number of features designed to help national statistical offices produce timelier and more reliable statistical data for monitoring national programs. These include: a large sample of households; a simple questionnaire with multiple choice questions for easy and rapid data collection; strong emphasis on high-quality fieldwork; the use of optical scanners to speed up data entry; pre-programmed validation procedures to ensure high built-in data quality levels; automated standardized outputs; and distribution of a CD-ROM with all survey documentation and survey data available for further analysis.

Appendix 2.

Millennium Development Goals, Targets, and Indicators and Their Measurement

Goals And targets	Indicators	How currently measured ¹⁸
Goal 1. Eradicate extreme poverty and hunger		
Target 1. Halve, between 1990 and 2015, the proportion of people whose income is less than one dollar a day	1. Proportion of population below \$1 (PPP) per day	Household survey
	2. Poverty gap ratio [incidence x depth of poverty]	Household survey
	3. Share of poorest quintile in national consumption	Household survey
Target 2. Halve, between 1990 and 2015, the proportion of people who suffer from hunger	4. Prevalence of underweight children under-five years of age	Household survey
	5. Proportion of population below minimum level of dietary energy consumption	FAO food balance sheet, Household survey, models
Goal 2. Achieve universal primary education		
Target 3. Ensure that, by 2015, children everywhere, boys and girls alike, will be able to complete a full course of primary schooling	6. Net enrollment ratio in primary education	Administrative records
	7. Proportion of pupils starting grade 1 who reach grade 5	Administrative data
	8. Literacy rate of 15-24 year-olds	Household survey, model based
Goal 3. Promote gender equality and empower women		
Target 4. Eliminate gender disparity in primary and secondary education preferably by 2005 and to all levels of education no later than 2015	9. Ratios of girls to boys in primary, secondary, and tertiary education	Administrative data
	10. Ratio of literate females to males of 15-24 year-olds	Household survey, model based
	11. Share of women in wage employment in the non-agricultural sector	Household survey, administrative records, official estimates
	12. Proportion of seats held by women in national parliament	Records of national parliaments
Goal 4. Reduce child mortality		
Target 5. Reduce by two-thirds, between 1990 and 2015, the under-five mortality rate	13. Under-five mortality rate	Registry of vital statistics, household survey, model based estimates
	14. Infant mortality rate	Registry of vital statistics, household survey, model based estimates
	15. Proportion of 1 year-old children immunized against measles	Administrative data, household survey
5. Improve maternal health		
Target 6. Reduce by three-quarters, between 1990 and 2015, the maternal mortality ratio	16. Maternal mortality ratio	Vital statistics, RAMOS, household surveys, model based estimates
	17. Proportion of births attended by skilled health personnel	Household survey, administrative data
Goal 6: Combat HIV/AIDS, malaria and other diseases		
Target 7. Have halted by 2015 and begun to reverse the spread of HIV/AIDS	18. HIV prevalence among 15-24 year old pregnant women	Routine sentinel surveillance
	19. Condom use rate of the contraceptive prevalence rate	Household survey
	20. Number of children orphaned by HIV/AIDS	-----
Target 8. Have halted by 2015 and begun to reverse the incidence of malaria and other major diseases	21. Prevalence and death rates associated with malaria	Predictive model, variety of data sources
	22. Proportion of population in malaria risk areas using effective malaria prevention and treatment measures	Household survey for prevention, administrative for treatment
	23. Prevalence and death rates associated with tuberculosis	Model based estimates, administrative data
	24. Proportion of tuberculosis cases detected and cured under directly observed treatment short course (DOTS)	Administrative data
Goal 7. Improve environmental sustainability		
Target 9. Integrate the principles of sustainable development into country policies and programs and reverse the loss of environmental resources	25. Proportion of land area covered by forest	Nat'l forest inventories, satellite images
	26. Ratio of area protected to maintain biological diversity to surface area	Administrative data
	27. Energy use (kg oil equivalent) per \$1 GDP (PPP)	National energy balance sheet

¹⁸ Information on how the indicator is presently measured is taken from Carraro et al, (2004).

Goals And targets	Indicators	How currently measured ¹⁸	
	28. Carbon dioxide emissions (per capita) and consumption of ozone-depleting CFCs (ODP tons)	Administrative Data	
	29. Proportion of population using solid fuels	Household survey	
Target 10. Halve, by 2015, the proportion of people without sustainable access to safe drinking water	30. Proportion of population with sustainable access to an improved water source, urban and rural	Household Survey, or administrative records	
Target 11. By 2020, to have achieved a significant improvement in the lives of at least 100 million slum dwellers	31. Proportion of urban population with access to improved sanitation	Household Survey or administrative data	
	32. Proportion of households with access to secure tenure (owned or rented)	Household Survey	
Goal 8. Develop a global partnership for development			
Target 12. Develop further an open, rule-based, predictable, non-discriminatory trading and financial system Target 13. Address the special needs of the least developed countries Target 14. Address the special needs of landlocked countries and small island developing States Target 15. Deal comprehensively with the debt problems of developing countries through national and international measures in order to make debt sustainable in the long term	33. Net ODA, total and to LDCs, as percentage of OECD/DAC donors' gross national income	Estimates provided by countries	
	34. Proportion of total bilateral, sector-allocable ODA of OECD/DAC donors to basic social services (basic education, primary health care, nutrition, safe water and sanitation)	Estimates provided by countries	
	35. Proportion of bilateral ODA of OECD/DAC donors that is untied	Estimates provided by countries	
	36. ODA received in landlocked countries as proportion of their GNIs	Estimates provided by countries	
	37. ODA received in small island developing states as proportion of their GNIs	Estimates provided by countries	
	38. Proportion of total developed country imports (by value and excluding arms) from developing countries and LDCs, admitted free of duties	NA	
	39. Average tariffs imposed by developed countries on agricultural products and textiles and clothing from developing countries	NA	
	40. Agricultural support estimate for OECD countries as percentage of their GDP	Estimates provided by countries	
	41. Proportion of ODA provided to help build trade capacity	NA	
	42. Total number of countries that have reached their HIPC decision points and number that have reached their HIPC completion points (cumulative)	IMF reports	
	43. Debt relief committed under HIPC initiative, US\$	IMF reports	
	44. Debt service as a percentage of exports of goods and services	IMF and WB estimates	
	Target 16. In cooperation with developing countries, develop and implement strategies for decent and productive work for youth	45. Unemployment rate of 15-24 year-olds, each sex and total	Household Surveys, administrative records (latter cannot provide complete picture of ILO definition of unemployment)
	Target 17. In cooperation with pharmaceutical companies, provide access to affordable, essential drugs in developing countries	46. Proportion of population with access to affordable essential drugs on a sustainable basis	Interviews with country experts
Target 18. In cooperation with the private sector, make available the benefits of new technologies, especially information and communications	47. Telephone lines and cellular subscribers per 100 population	Estimates provided by countries, annual ITU questionnaire	
	48. Personal computers in use per 100 population and Internet users per 100 population	Estimates provided by countries, annual ITU questionnaire	

References

- Aboyade, Ojetunje, Ran. A. Yotopoulos, Hans F. Illy, and Meja Naur. 1993. "The Social Dimensions of Adjustment: A General Assessment." Consultant report to the Social Dimensions of Adjustment (SDA) Steering Committee, Washington, D.C.
- Carraro, Ludovico, Salman Kahn, Simon Hunt, Georgina Rawle, Matt Robinson, Manos Antoninis, and Laura Street. 2003. "Monitoring the Millennium Development Goals: Current Weaknesses and Possible Improvements." Consultant Report to Paris21. Oxford Policy Management, London.
- De Vries, Willem F. M. 1998. "How are We Doing?: Performance Indicators for National Statistical Systems." *Netherlands Official Statistics* 13 (Spring) :5--13.
- Klugman, Jeni, and Kinnon Scott. 1997. "Measuring Labor Market Status in Kazakhstan." In Jane Falkingham, Jeni Klugman, Sheila Marnie, and John Micklewright, eds., *Household Welfare in Central Asia*. London: Macmillan Press LTD. [Ltd.?)
- Ravallion, Martin. 1996. "How Well Can Method Substitute for Data? Five Experiments in Poverty Analysis." *World Bank Research Observer* 11 (2) :199--222.
- Ryten, Jacob. 2000a. "El Programa Mecovi: Ideas para el futuro." Mid-term Evaluation Report on the MECOVI Program. Inter-American Development Bank, Washington, D.C.
- Ryten, Jacob. 2000b. "El Programa Mecovi en Peru: Una evaluación." Mid-term Evaluation Report on the MECOVI Program. Inter-American Development Bank, Washington, D.C.
- Scott, Kinnon. 2003. "Mulit-topic Household Surveys." In Francois Bourguignon and Luiz Pereira Da Silva, eds., *The Impact of Economic Policies on Poverty and Income Distribution: Evaluation Techniques and Tools*. New York: Oxford University Press.
- Scott, Kinnon, Diane Steele, and Tilahun Temesgen. Forthcoming. "Living Standards Measurement Study Surveys." In Yansaneh, Ibrahim, ed. *Household Surveys in Developing and Transition Countries: Design, Implementation and Analysis*. New York: United Nations Statistical Division.
- Yansaneh, Ibrahim, ed. Forthcoming. *Household Surveys in Developing and Transition Countries: Design, Implementation and Analysis*. New York: United Nations Statistical Division.
- U.S. Bureau of Labor Statistics (BLS). 1994. "How the Government Measures Unemployment. U.S. Department of Labor, Report 864, Washington, D.C.
- United Nations. 1984. "Handbook of Household Surveys", Studies in Methods, Series F, No. 31, (ST.ESA.TATS.SER.F/31), New York.

World Bank. 2000. Nicaragua: "Ex-Post Impact Evaluation of the Emergency Social Investment Fund (FISE)." Report 20400-NI, Washington, D.C.