GUIDE TO THE ANALYSIS AND USE OF HOUSEHOLD SURVEY AND CENSUS EDUCATION DATA

UEPA / UAPS

The Network on Family and Schooling in Africa (FASAF), created in 1997, is a network of researchers, statisticians and education planners dedicated to using secondary data sources on the demand for schooling, and to strengthening national research capacities and collaborative research between institutions and national statistics and education offices. FASAF operates under the purview of UAPS.

UERD, BP 7118, Ouagadougou 03, Burkina Faso
Telephone: (226) 36.21.15 Fax: (226) 36.21.38 E-mail: pilon@ird.bf; yacoubay@yahoo.com

The Union for African Population Study (UAPS), created in 1984, is a non-profit pan-African scientific organization composed of specialists working on population issues. UAPS publishes scientific reviews, supports researchers, offers technical assistance, and coordinates seven thematic research networks.

UAPS, BP 21007, Dakar Ponty – Dakar, Sénégal
Telephone: (221-8) 25.59.51 Fax: (221-8) 25.59.55 E-mail: uepa@cyg.sn

The United Nations Educational, Scientific and Cultural Organization (UNESCO) is an agency that seeks to contribute to the maintenance of peace and security by fostering collaboration between nations through education, science and culture. The UNESCO Institute for Statistics has as its global mandate the collection, analysis and dissemination of comparable statistics and indicators on education, science and culture and enabling national capacity in member states to address these same areas.

UNESCO Institute for Statistics, C.P. 6128, Succursale Centre-Ville, Montréal, Québec H3C 3J7, Canada
Telephone: (1-514) 343-6880 Fax: (1-514) 343-6882 E-mail: information@uis.unesco.org
Website: http://www.uis.unesco.org

The United Nations Children's Fund (UNICEF) is an agency that aims to help children living in poverty in developing countries, including the care and stimulation they need in the early years of life, and protection from illness, war and natural disasters.

Strategic Information Section, UNICEF, 3 UN Plaza, New York, NY 10017 USA
Telephone: (1-212) 824-6745 Fax: (1-212) 824-6474 E-mail: eloaiza@unicef.org
Website: http://www.unicef.org

The United States Agency for International Development (USAID) Office for Human Capacity Development in the Bureau for Economic Growth, Agriculture, and Trade is responsible for implementing USAID's goal of building human capacity through education and training. The office provides field support, technical leadership and research to help nations and USAID field missions improve education and training and to help develop stable, democratic countries with thriving market economies and healthy, well-educated families.

USAID/G/HCD, DHS EdData Mgr., 1300 Pennsylvania Ave., NW, Washington, D.C. 20523-3901 USA
Telephone: (1-202) 712-4273 Fax: (1-202) 216-3229
Website: http://www.usaid.gov/educ_training

ORC Macro is an international research, management consulting and information technology firm, which provides survey research, policy analysis, evaluation, and training. ORC Macro implements MEASURE *DHS***+** with the Population Council and the East-West Center, and the DHS EdData Activity, a USAID Office for Human Capacity Development activity. DHS EdData provides information for education policy and programme planning through the collection and dissemination of education data.

ORC Macro, 11785 Beltsville Drive, Calverton, MD 21202 USA
Telephone: (1-301) 572-0200 Fax: (1-301) 572-0999 E-mail: measure@macroint.com
Website: http://www.dhseddata.com

ISBN 92-9189-016-2

Ref: UIS/TD/04-03
Cover design: JCNicholls Design
Printed by: ICAO

CONTENTS

Page

Page

TABLES AND FIGURES

Page

LIST OF ACRONYMS

AIR	Apparent intake ratio
DHS	Demographic and health survey
EFA	Education for All
EP	*Enquête prioritaire*
GAR	Gross attendance rate
GER	Gross enrolment ratio
MICS	Multiple indicator cluster survey
LSMS	Living standards and measurement studies
NAR	Net attendance rate
NER	Net enrolment ratio
UIS	UNESCO Institute for Statistics

PREFACE

There is a renewed awareness of the potential of household surveys and population censuses to provide the kind of information on household and individual characteristics and their multiple interactions that is needed to formulate and monitor development.

More and more countries are organising multiple-topic household surveys using standardised approaches to measuring on a regular basis. If properly designed and implemented, these surveys can provide important and unique information. This Guide has been compiled with the aim of facilitating and promoting education policy analysis and utilisation of education indicators based on household survey and census data.

The Guide was first conceived and produced in French by the Research Thematic Network, *Family and Schooling in Africa* (FASAF) of the Union for the African Population Study (UAPS), which comprises a working group of researchers, statisticians and education planners from eight countries (Benin, Burkina Faso, Cameroon, Democratic Republic of the Congo, Côte d'Ivoire, Mali, Niger, and Togo). Other researchers and members of the UNESCO Institute for Statistics also made contributions to the Guide. It was published by the French Center for Population and Development (CEPED) in collaboration with UAPS and UNESCO in 1999.

This English version is a collaborative effort between FASAF, the UNESCO Institute for Statistics, UNICEF, the United States Agency for International Development (USAID), and ORC Macro. It is not simply a translation of the French version but rather an extension of the original document, refining its main tenets and adding new insights.

Within each contributing organization, the following staff collaborated on the English version of the Guide:

- Jean-Francois Kobiane and Marc Pilon (FASAF)
- S.K. Chu, Albert Motivans, Diane Stukel, Bertrand Tchatchoua and Nghia Bui Quang (UNESCO Institute for Statistics)
- Edilberto Loaiza (UNICEF)
- Kimberly Bolyard and Linda Padgett (USAID)
- Kristi Fair and Anne Genereux (ORC Macro)

We hope the readers of this Guide find information and ideas that will enable them to carry out meaningful and useful analysis of education data from demographic censuses and household surveys. The authors invite you to share your experiences and findings in order to further enrich and upgrade this Guide into a more substantive and practical reference tool.

INTRODUCTION

The demand for data on education has never been greater, both at the national and international levels. In many countries, large-scale household surveys are conducted regularly to provide information on population, health, education, household income and expenditure, employment, and other critical areas of study. In many instances, the education information from these data collection efforts is underutilised. However, these data can provide invaluable information to inform policy debates.

The information is underutilised for two main reasons: first, because potential data users are unaware of the education data collected by these censuses and surveys; and, second, because reports of the findings present education data primarily as a background characteristic rather than an object of study *per se*.

For example, reports often discuss data on education in relation to household expenditure or as a factor in decisions about childbearing rather than presenting data on differences in educational attainment by sex, age, urban/rural residence, and other characteristics.

At the same time, it is important to accurately report and interpret survey data that are subject to error (due to sampling design and to non-sampling error), particularly in developing countries where special sampling designs, such as cluster approaches, may affect the likelihood for error.

This Guide is intended for staff in the Ministry of Education (MOE) and others who work with education indicators, including education trainers and planners, education researchers, staff from national statistics offices responsible for conducting large-scale household surveys and census operations, and international agencies involved in education issues.

The Guide may also be useful to staff from Ministries of Education and the organizations that collaborate in survey design and data analysis in the framework of census and household surveys. Such collaboration allows education experts and those in survey methodology to confer on which data are of most interest, how best to collect these data, and how to analyse the data and present results. Hopefully, the collaboration between these organizations will also result in greater attention to education issues in the design of future multiple-topic surveys. In addition, MOE staff collaborating with survey implementing organizations are likely to utilise the data set to conduct ongoing analyses, possibly in conjunction with administrative data.

Specifically, this Guide intends to:

- provide an overview of household-level education data that are available from various sources and to describe their usefulness;
- provide guidance on the analysis and presentation of descriptive data so that they are most useful to Ministries of Education and to policy-makers; and

- contribute to the standardisation of indicators and calculation methods based on household-level education data.

The guide is organised in five chapters. The first chapter examines different sources for education data. Chapter 2 describes and discusses some of the education indicators that can be produced based on these data sources. Chapter 3 describes national population censuses and the major international household survey programmes. Chapter 4 discusses survey methodology and its implications for the analysis of education data, while Chapter 5 addresses various ways of presenting education survey data and provides ideas for graphically representing survey data.

SOURCES OF EDUCATION DATA 1

1.1 EDUCATION STATISTICS BASED ON ADMINISTRATIVE DATA

Governments regularly collect education statistics to monitor and manage the education system. Schools and other educational institutions regularly report data on students, teachers, expenditures, and physical facilities. These data and the indicators derived from them are then used to gauge the capacity and performance of schools in relation to national education goals and plans, and to determine future development policies, plans and management arrangements. Typically, they are inexpensive, as a by-product of the administrative data collection. They usually cover the whole country and are available on an annual basis.

As in other areas (such as health and social welfare), the data collected through administrative channels on education sometimes face shortcomings related to data quality. Some aspects of quality include coverage, accuracy and timeliness. In the case of education, in many countries, there is less than universal reporting by schools, so various data points for the non-reporting schools have be estimated. Also, in many countries, administrative data are only available for schools, teachers and students in the formal government sector or for private schools that receive significant government funding. Private educational institutions and non-formal educational programmes managed by non-governmental organizations and local communities are often not covered by administrative statistics.

Questions also arise as to the reliability of data reported by government and public schools, particularly when resource allocation from the government is tied to enrolment and the number of teachers or classrooms. Schools may report higher enrolment figures in order to obtain greater resources, producing distortions in the estimates of student enrolments. There is also anecdotal evidence of private schools that under-report income and expenditure in order to derive greater benefits or lower their tax liability.

1.2 HOUSEHOLD SURVEY AND CENSUS EDUCATION DATA AS A COMPLEMENT TO ADMINISTRATIVE DATA

Administrative data typically provides limited information on the individual characteristics of pupils (such as age, sex and residence), and little information on the characteristics of their households. Moreover as school-based surveys and censuses focus on children who attend school, there is no information on the individual characteristics and family backgrounds of children who do not attend school.

Household surveys and censuses provide important education data that can be analysed according to household and individual characteristics. The availability of multiple censuses and surveys conducted over time enables changes to be tracked within a given timeframe, assuming consistency in survey questions and methodology.

Data from household surveys and censuses can complement the school-based data by providing information on aspects of children's background that may influence household schooling decisions. Censuses and multi-topic household surveys provide considerable information on household and individual household member characteristics, including data on children's school participation. These data on children's school enrollment or school attendance can then be analysed according to household and child characteristics.

For example, depending on the survey, the percentages of school-age children attending school can be compared by urban-rural residence, household wealth, sex, and other characteristics critical for decision-making. Knowing that 33% of poor rural girls of primary school age currently attend school, compared to 57% of wealthy rural girls, for example, provides insight into the relationship between household wealth and school participation.

Censuses and household surveys also provide data on adult educational attainment and often, on self-reported literacy skills. These surveys and censuses can provide national-level sources of data on adult educational attainment and literacy and allow comparisons by different household characteristics.

1.3 Comparing Education Indicators Based on Administrative and Survey Data

Administrative and household survey data sources measure educational participation in different ways. Administrative data are based on school reporting at the beginning of the school year, but in some cases it can include reporting at the middle or end of the school year. Enrolment ratios are based on the numbers of children enrolled in school and the school-age population estimated from national censuses and/or vital statistics.

Ideally, household surveys collect data on enrolment and/or school attendance among a representative sample of children. Questions concerning children's school participation are typically asked of the head of household. The timing of the survey is not related to any point in the school year and may actually even cross two different school years.

Estimates of educational participation from these two sources may differ for a number of reasons. One major factor is that the question asked in the household surveys querying children's school attendance is different from that answered by school censuses: Attending school is not necessarily the same as being enrolled in school. Children may be recorded in school enrolment records and yet not actually be attending school.

The different rates of participation can also be attributed to the timing of data collection relative to the school year. A school census conducted at the beginning of the school year and a household survey collecting data at the end of the school year will likely find different rates of participation since some children will have enrolled in school without ever actually attending, and other children will have dropped out of school during the school year.

In addition, the accuracy of the population estimate and the completeness of school-level data can affect the calculation of participation ratios from administrative data. Similarly, the completeness of the census enumeration and the sample design for the household survey may also affect the accuracy of estimates produced by censuses and surveys.

In short, many factors may contribute to variations in estimates of enrollment and attendance ratios. Further research is needed to explore the reasons for similarities or differences between these two different measures of participation.

EDUCATION INDICATORS FROM NATIONAL POPULATION CENSUSES AND HOUSEHOLD SURVEYS 2

2.1 AN ANALYTICAL FRAMEWORK

National censuses and multiple-topic household surveys collect data on the characteristics of both households and individuals within those households. This chapter looks at education indicators that are commonly derived from data sets or that can be constructed using existing variables.

Most surveys and censuses collect information on household characteristics, including those related to location (such as urban-rural residence or region); household composition and characteristics (household size, sex of the household head and so on.); and socio-economic characteristics (household wealth, dwelling quality and type).

Measuring the socioeconomic status of a household is a complex issue. Some surveys collect detailed data on household income and expenditures which are then used to construct distributions of households, for example, to identify the 20% of the population with the greatest income and the 20% of the population with the lowest income.

However, in most multi-purpose surveys, it is not possible to collect this information. A proxy used for household socioeconomic status when detailed data are not available is commonly called a household wealth or assets index. An asset index, for example, proxies socio-economic status by household access to amenities or ownership of items. The composition of an asset index varies by survey, but assets commonly included in the index are electricity, radio, television, refrigerator, bicycle, motorcycle, car, telephone, number of persons per sleeping room, water source, sanitation facilities, and type of flooring. Each household asset used in the index is assigned a weight generated through an analysis of principal components, which calculates the importance of each element of the index.

Data on individual household members are collected on variables such as sex, age, relationship to the household head, marital status, schooling attainment by level and grade at that level (and sometimes the highest degree completed). Data are also often collected on employment status and other characteristics.

For children in the household (usually those under age 15), data are often collected on parental survivorship (whether each child's mother and father is alive, and whether each parent lives in the household).

Data on current school participation (whether currently attending, and if so, at what level and grade) are collected for youth. The age ranges for data on education participation vary widely, and may include children of pre-school age, those of primary and secondary school age, youth and even adults. Some surveys also collect literacy data for adults and youth using either self-assessment or testing as a baseline. Many surveys also collect time-use data or labour data (such as unemployed or employed or working unpaid/paid in kind/paid in cash).

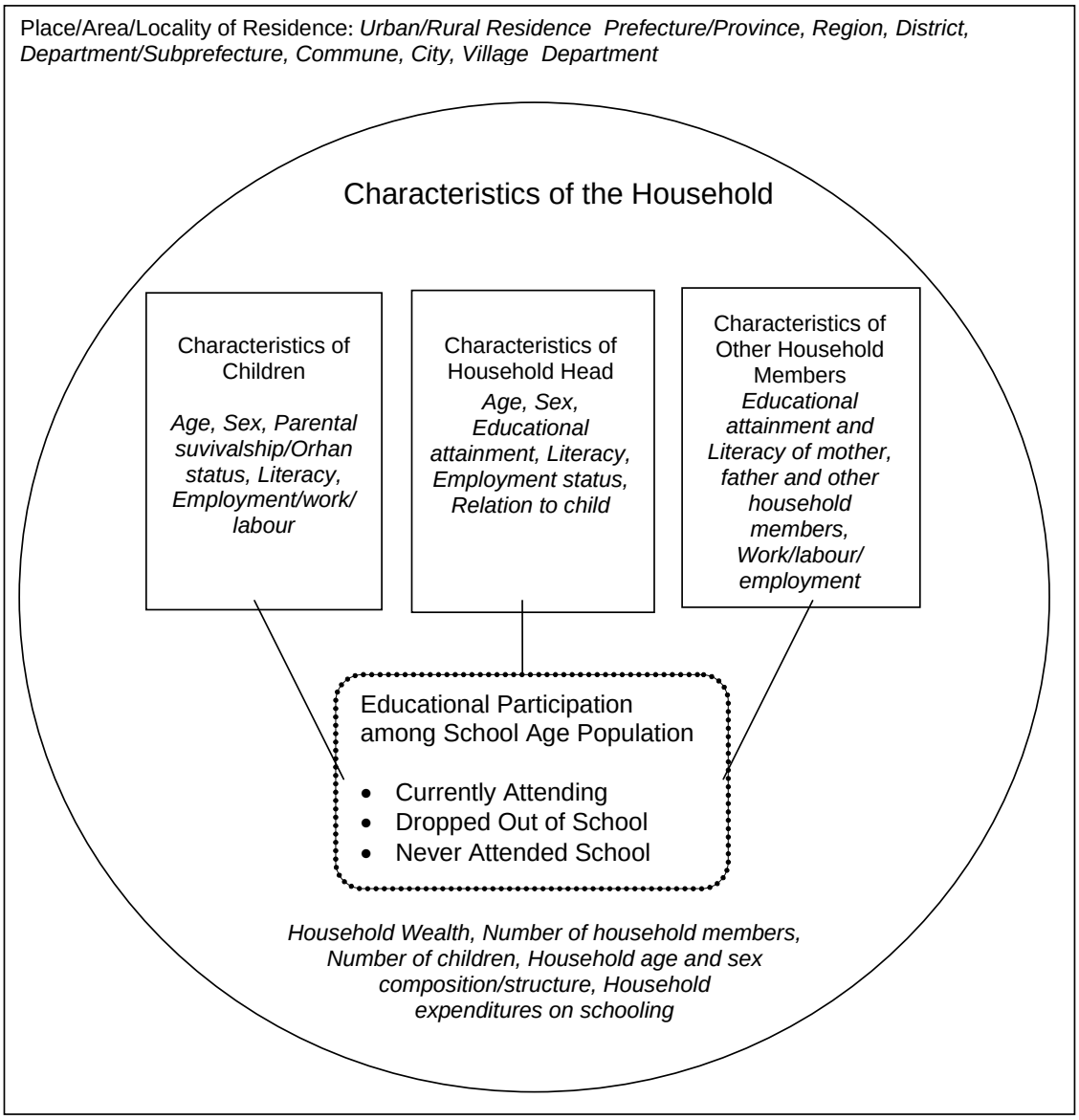
Access to background information on children and their households facilitates data analysis by various characteristics and groups. The analytical framework (*see Figure 1*) illustrates the kinds of data often available from national censuses and household surveys, as well as interrelationships between the data First, the geographic context in which households are situated is determined: urban or rural areas, villages or towns, districts, and so on. The household is located within that square and is characterised by selected variables (such as, wealth, size, and number of school-age children). Within the household, the household members are characterised by age, sex, educational attainment and literacy, and employment status. The rounded rectangle with the dashed border outlines the schooling status of children (attending school, dropped out of school, and never attended school).

This approach provides a framework that can be used to analyse data on children's school participation by urban-rural location, region, household wealth, parents' educational attainment, children's sex and age, and other household and individual characteristics. Another way of studying these differences in educational participation is to attribute causation through statistical analysis. This more advanced analysis is not the topic of this Guide, though a brief overview to statistical analysis is provided in Chapter 4.

Various education indicators can be calculated based on the data collected in a given survey or census. This chapter presents a set of education indicators and explains briefly how they are calculated (*see Appendix A*). The discussion includes indicators that have been established through many years of use in the international education community and those that are the result of more recent indicator development.

Before discussing each indicator in turn, it must be emphasised more broadly that statistics and indicators—which are data organised and presented as information—must be interpreted before being used to inform policy-making or programme planning. Indicators themselves do not explain anything; rather, they point to achievements or change (Bottani, 1990). Indicators must be interpreted and placed in different contexts, including the socioeconomic, cultural, and demographic characteristics of the population; educational programmes and reforms underway; the quality of the curriculum, teachers, school management, school facilities; and so on. In summary, the indicators can be used in conjunction with other available contextual data to analyse changes in education indicators.

Figure 1
A framework for analysing education data from household surveys and censuses



Note: Not all of the indicators listed in the framework will be available in all household surveys. This analytical framework is meant to be indicative of the types of information available and suggestive of possible analyses.

2.2 LITERACY

Literacy generally denotes the ability to read and write and to use written words (in any written language) in everyday life. Literacy is one of the intended outcomes of schooling, as well as a measure of a person's ability to function in society and to continue to learn.

Literacy is measured in different ways across national censuses and household surveys. A common approach is to ask whether each person within a household can read and write, with understanding, a simple statement about his/her everyday life. A person who can do so is considered to be literate; a person who cannot is classified as illiterate. Some surveys classify a person as "semi-illiterate" if he/she can read but cannot write or can read and write only figures or his/her own name; but this category is not used in international comparisons.

Some critics suggest that literacy estimates based on self-reported declarations may not be reliable since some illiterate persons may be reluctant to admit to their illiteracy. In addition, if one household respondent, such as the household head, reports on literacy for all household members, he or she may be mistaken in judging the reading or writing ability of other household members. It is therefore recommended that actual tests of literacy be administered either as part of national censuses or household surveys, or as part of a post-enumeration evaluation survey. Tests may be simple reading tests using newspaper cuttings or a standard text, or more elaborate literacy tests using multiple criteria-based and norm-based methods that gauge ability based on literacy, numeracy and information-use scales.

Literacy indicators can be structured in various ways. Typically, the age range generally covers only adults, partly because children's literacy is less commonly tested in household surveys. Whenever possible, these indicators should be compared by sex and residence in order to illustrate potential literacy gaps among these population groups.

The adult literacy ratio is the percentage of the adult population (generally defined as those age 15 and older) that is literate.

The literacy ratio of the population age 15-24 is the percentage of the youth population that is literate. Assuming that most youth have attended school, the literacy ratio among adults age 15-24 provides a gross measure of literacy among youth who have never attended school, still attend school, or who recently left school. This indicator is sometimes used as a rough proxy for youth learning achievement.

A key policy issue is the illiteracy ratio and the characteristics of illiterates who represent the key target group for literacy programmes. Literacy ratios among population cohorts can also be a useful way to discern the change in literacy ratios over time. For example, ratios can be compared, by sex and by 10-year age cohorts (such as, 15-24 or 25-34.), to assess how literacy ratios change over time.

2.3 EDUCATIONAL ATTAINMENT

Educational attainment is used as an indicator of human capital and is measured by the percentage distribution of the adult population (generally those age 15 and older) by the number of years or highest level of schooling attended, or completed. For example, perhaps 22% of the adult population have never attended school, with 58% having attended primary school, 17% secondary, and 3% post-secondary school. These indicators are generally presented by age and sex to illustrate how patterns of school attainment and completion have changed over time.

The data can be standardised for ease in interpretation across countries and within countries where the education system has changed over time. The standard international approach to level of education is as follows:

- No schooling (persons with less than one year of primary education);
- Incomplete primary education (persons who have not reached grade 5 of primary education);
- Completed primary education (persons who have attended grade 5 and higher or successfully completed primary education);
- Completed lower secondary education;
- Completed upper secondary education;
- Completed post-secondary non-tertiary or tertiary education.

If necessary, a separate category can be added for persons who have attended or completed only pre-primary school. Similarly, the category of post-secondary non-tertiary and tertiary education can be further sub-classified in accordance with both the national education system as well as with levels 4, 5 and 6 of the International Standard Classification of Education (ISCED).

Using the same data on the highest level of schooling attended and highest grade completed, *adult primary and secondary school completion ratios* can also be calculated. The primary school completion ratio is the percentage of the adult population (age 15 and older) that has either successfully completed primary school, or has attended a higher level of schooling. The secondary school completion ratio is the percentage of the adult population (age 20 and older) that has either successfully completed secondary school, or has attended tertiary-level schooling.[1] In contrast with the indicators of educational attainment, these indicators depict the educational composition of the adult population in terms of levels of schooling completed.

In calculating both attendance and completion ratios for adult populations, it is important to bear in mind the official age range of primary school and the prevalence of over-age attendance. If the official age range for primary school is 6-12, for example, but there is

[1] The age range used to calculate the secondary school completion ratio among adults typically is age 20 and older to allow for the fact that the official age range for secondary school ends in the late teenage years.

a significant over-age primary school population, the primary school completion ratio among adults age 15-24 will be affected if many of these young adults currently attend primary school. The same principle applies at the secondary level.

2.4 SCHOOL PARTICIPATION RATIOS

As discussed in Chapter 1, indicators of current school participation produced by household surveys and national censuses are generally based on questions about a child's attendance at school (*see Appendix A*). An indicator of participation based on enrolment can also be calculated, depending on the questions asked in the census or survey.

To calculate net and gross participation ratios, the national census or household survey must provide information on children's ages, whether children currently attend school, and if so, at what level and grade. The *net attendance ratio (NAR)* usually is calculated separately by ISCED level (primary, lower secondary and upper secondary school). The primary NAR is the percentage of the official primary school-age population that attends primary school. If there are 6 primary school grades and the official entry age is age 6, the official primary school age range would be 6-11. The secondary school NAR is the percentage of the official secondary school age population that attends secondary school. These indicators show the extent of participation by level of schooling among children of the "official" school age.

The NAR is an important indicator, as it provides the percentage of children of primary or secondary school age currently attending the age-appropriate level of schooling. That is to say that, although there may be grade repeaters or late starters for the individual grades within the primary or secondary level who are not "on-time for grade", they are in fact on-time when considering the entire level as a whole. However, in terms of cross-national comparisons, there are two approaches to comparing NARs for countries with differently-structured school systems. For example, in Angola, there are only 4 years of primary school, with the official primary school age range being 6-9; in neighboring Namibia, by contrast, there are 7 years of primary school, with the official primary school age range being 7-13. Consequently, the first approach entails applying the International Classification of Education Systems (ISCED), making sure that national categories of educational programmes are mapped to the ISCED system. In this way, cross-national comparisons can be better accommodated.

The cohort approach is another way to compare NARs across countries by using attendance ratios among children in a specific age range. For example, the UNICEF World Summit for Children Indicators measures the percentage of children age 6-12 attending school. This approach compares children in the same age group, although national variations in education systems are not taken into account.

The *gross attendance ratio (GAR)* is calculated for each level of education. For example, the primary school GAR is the total number of students attending primary school—regardless of age—expressed as a percentage of the official primary school age population. The GAR indicates the level of participation in a specific level of

schooling by pupils at any age, and when compared with the NAR, indicates the extent of over- and under-age participation by level of schooling. In countries with high ratios of primary/secondary school attendance, the GAR can exceed 100 if there are many over-age (or under-age) students in school.

The *gender parity index* (GPI) measures gender-related differences in school participation. The GPI is calculated by dividing the gross attendance/enrolment ratio for females by the gross attendance/enrolment ratio for males. Using an example from Mozambique cited in the UNESCO database, in 2000, the female primary school gross enrolment ratio (GER) was 73 and the male GER was 98. Thus the GPI was 0.91. In contrast, in Guinea, where the gender gap in school participation is much wider than in Mozambique, the female primary school GER in 1997 was 41 and the male GER was 68, producing a GPI of about .60. As these examples illustrate, the closer the GPI is to 0, the greater is the gender disparity in favor of males. A GPI of 1.0, on the other hand, indicates gender parity. A GPI greater than 1.0 indicates a gender disparity in favor of females, meaning more females than males attend school. However, the GPI is only one measure comparing male and female participation ratios in schooling; other measures can also be quite useful, particularly when data are available for more than one point in time.[2]

The calculation of the *gross* and *net intake ratios* requires information on children's school participation over a two-year period. The net intake ratio is the number of new entrants to the first year of primary school who are of the official school starting age, divided by the total number of children in the population of the same age. The net intake ratio for Guinea, for example, is calculated as follows: (number of 7-year-old children who did not attend the first year of primary school last year, but did attend the first year of primary school this year) / (number of children who are age 7). Note that the indicator does not reflect the number of 7 year-olds currently attending the first year of primary school now, but those who are in the first year of primary school for the first time.

The *gross intake ratio* is the number of new entrants to the first year of primary school at any age, divided by the total number of children of the official school entry age. The gross intake ratio is calculated as follows: (number of children of any age who did not attend the first year of primary school last year, but did attend the first year of primary school this year) / (number of children who are of official school entry age). The gross intake ratio cannot be lower than the net intake ratio, and can exceed 100 if participation in the first year of primary school is high and if there is significant over-age or under-age participation.

The *proportion of students over-age, under-age and on-time, by grade* can be calculated for different education levels. Based on the official target age for each grade, one can calculate the proportion of children in that grade who are older or younger or at

[2] There are several different approaches to monitoring change in the gender gap. The absolute gender gap, for example, is the male GAR (or the gross enrolment ratio) minus the female GAR (or the gross enrolment ratio). Comparing the absolute gender gap at two points in time shows changes in the gender gap.

the target age for the grade.[3] Over-age among primary school students, in particular, is prevalent in many developing countries. Over-age children occupy places in school that might have been given to children at the target age for the grade. In addition, over-age children may be more likely to drop out of primary school before completing that level of schooling.

The *age-specific attendance ratio* (ASAR) indicates the percentage of youth of a given age that currently attends school, regardless of the level of participation (primary, secondary, or higher). This indicator is an important complement to the NAR and GAR, as the ASAR captures current school participation at any level of schooling rather than only at the level intended for a child at a specific age. The ASAR cannot exceed 100%.

2.5 STUDENT FLOW RATIOS

Student flow ratios describe the progression of students through the school system, including students leaving the system. The *promotion ratio* is the percentage of students in a grade in a given school year promoted to the next grade in the following school year.

The *repetition ratio* is the proportion of students in a grade in a given school year who also attend that same grade in the following school year. Together with the dropout ratio, the repetition ratio helps to describe the flow of students through a school system, and suggests where problems with student flow ratios are most acute. For instance, repetition ratios at the primary school level are often highest at the first and last grades of the cycle. More students may repeat the first grade because of perceived or actual unreadiness to proceed; similarly, many students may repeat the last grade of primary school in an effort to perform well enough to progress to secondary school.

The *dropout ratio* is the proportion of students in a grade in a given school year who no longer attend school the following year. Both rates are calculated from data on children's school participation during two consecutive school years.

Household survey data on dropouts, in particular, provide a useful complement to administrative data. Administrative data estimate dropout ratios by assuming that those students who, from one year to the next, are not promoted to the next grade and who are not repeating the grade, dropped out of school. Typically, records of whether students who left a particular school went to another school elsewhere or left school altogether do not exist. As a result, some student transfers are inevitably counted as dropouts and thus overstate the real dropout rate. Under these circumstances, it is useful to have another source of data on dropouts that is not affected by student transfers between schools.

[3] Students are considered over-age if they are two or more years older, and under-age if they are one or more years younger than the official age for the grade. Students are considered to be on-time if they are of the official age, or are one year older than the official age for their grade: for example, if the official entry age for grade 1 is 7, a student age 7 or 8 is considered on-time in grade 1; a student age 9 or older is considered over-age; and a student age 6 or younger is considered under-age.

Survival ratios estimate the percentage of students who start grade 1 of primary school who will eventually attain a given grade, with or without grade repetition. Often, survival ratios are calculated to grade 5 (a proxy for attaining literacy) and to the final grade of primary school. The survival ratio to grade 5 estimates the percentage of grade 1 students in a specific year that will "survive" or progress through the beginning of grade 5.

Ideally, survival ratios are calculated using the true cohort method, which involves either a longitudinal study of a cohort of students moving through the school system, or a retrospective study of students using existing school records (UNESCO, 1998). National census and household survey data, however, do not allow for the true cohort method approach. Instead, the survival ratio is estimated using the reconstructed cohort method, which requires a series of assumptions (UNESCO, 1999). The survival ratio is calculated using ratios of promotion, dropout, and repetition, with the projection based on several assumptions, including: a) that there are no new entrants to the school system (including dropouts returning to school); b) that at any grade, the same promotion, repetition, and dropout ratios will apply to students, regardless of whether a student is in the grade for the first time or is repeating; c) that the same promotion, repetition, and dropout ratios observed apply for students while they are in that cycle of schooling; and d) that the number of times students are allowed to repeat is defined. Using the reconstructed cohort method, survival ratios are linked most closely to dropout ratios—the lower the dropout ratio, grade by grade, the higher the estimated survival ratio.

The *transition ratio from primary to secondary school* is the proportion of students attending the final grade of primary school in a given year, who attend the first grade of secondary school in the following school year. This indicator does not project the proportion of students from a particular cohort that eventually may make the transition to the first grade of secondary school, but rather those who made the transition in a given school year.

In addition to indicators calculated at the individual level, education indicators can also be examined at household level. One interesting question to ask is what percentage of households have children of school age. This can be further refined according to school level (pre-primary, primary, secondary, and tertiary).

With this information, the percentage of those households with one or more children attending school can be calculated, as well as the percentage of households with one or more school-age children with none of the children attending school. More detailed analyses are possible, such as the percentage of school-age children in school among households with one or more school-age children. Household level data on children's school participation can be used to supplement indicators related to individual participation.

NATIONAL POPULATION CENSUSES AND SELECTED HOUSEHOLD SURVEYS 3

3.1 NATIONAL POPULATION CENSUSES

A population census represents a complete enumeration of the entire population. As a comprehensive source of demographic, social, and economic data on a population, a national census provides valuable information for policies and the planning of socio-economic development from the national to the local level. A national census also provides the framework for drawing household and population samples for household surveys.

National population censuses are large-scale statistical operations requiring considerable resources, organization and preparation, and are carried out about every 10 years over the entire national territory. Censuses can cover the *de facto* population, or all individuals residing in the country at the time of the censuses or the *de jure* population, which also covers nationals living abroad. Because census data provide information on the entire population, data can be disaggregated better than survey data (for example, to region or district level or small sub-populations), which may not be representative at the sub-national level.

National population censuses collect data on the individual and household characteristics of the population. Characteristics include age, sex, household structure, migration, and a range of other socio-economic characteristics such as wealth, language, religion, education, employment, occupation, disability and so on.

The latest *Principles and Recommendations for Population and Housing Censuses* published by the United Nations provides useful guidelines for collecting data on literacy, school attendance, educational attainment, field of study and educational qualifications (United Nations Statistical Office, 1998). These guidelines recommend that a literate person be defined as one who can both read and write, with understanding, a short simple statement on his/her everyday life. Literacy may be in any written language. It should not be assumed that persons who attended school are literate and that those who did not attend are not literate. Data on school attendance should be collected from persons age 5 to 29 years, although the age range may vary by country. Educational attainment is defined as the highest level and grade of education attended or completed, taking into account all types of education and training as far as they are measurable in terms of level and grade equivalence. Population censuses may also include a question on the field of education for persons 15 years of age and over who attended at least one grade in secondary school or in equivalent training. Some countries include a separate question on educational certificates, diplomas, and professional titles and degrees obtained.

3.2 SELECTED HOUSEHOLD SURVEYS

The use of education indicators from international multi-topic survey programmes in less developed countries has grown markedly since the mid-1990s, thanks in part to the efforts of multilateral and bilateral institutions that fund survey programmes, such as UNICEF, the World Bank, USAID and others. Generally, these surveys have collected data on a wide range of development issues and provide information on population, health, education, household income and expenditure, employment, and other critical public policy domains.

This Guide encourages greater use of the education data available from these survey programmes, as well as from national censuses and from other national household surveys. The Guide's emphasis is on surveys that cover multiple topics, including education, rather than on surveys that primarily collect education data. In addition to the larger multi-topic international household surveys mentioned above, country statistical offices and other agencies have carried out other useful national household surveys to collect data. Some education data can be obtained from these efforts, which typically include surveys on the labour force, and income and expenditure. Data from national surveys will probably have already been used by education constituents in country as opposed to data from multiple-topic surveys. These surveys may have been conducted once or multiple times over a period of years.

This Guide focuses on several widely-implemented multiple-topic household surveys in order to familiarise potential data users with the kinds of data collected by each survey, and where possible, to provide specific information on the limitations of each type of survey. While many surveys use core sets of questions across countries, country-specific variations often make it difficult to generalise about survey content and available data. Furthermore, generalisation is made even more difficult because the core instruments used often evolve over time so that the data collected early on in a programme change with time. Under these circumstances, while the Guide considers specific surveys, it does not methodically list all variables available from each type of survey, but instead focuses on general survey characteristics and approaches.

Another important note regarding data sets is the question of access to the data. Generally, survey implementing organizations determine how data sets may be accessed and by which individuals or organizations. Some programmes disseminate data sets via the internet (*see Appendix B*).

Demographic and Health Surveys (DHS)

The Demographic and Health Surveys (DHS) is an international survey programme designed to provide current and reliable information on key indicators of social development, including fertility levels and trends, family planning knowledge and use, infant and child mortality, and maternal and child health. Since 1984, DHS has conducted a total of over 140 surveys in 70 countries, with many countries having conducted multiple surveys at five-year intervals. Sample size varies greatly across surveys, with a trend toward larger sample sizes over time in order to disaggregate the

data to lower levels. For example, the DHS sample in Malawi included about 5,000 households in 1992 compared with 15,000 households in the 2000 DHS. The sample size of 15,000 households allowed data on key indicators to be disaggregated to the district level for priority districts; the smaller sample did not permit analysis at the sub-regional level.

The DHS includes a household questionnaire, which collects data on household and individual characteristics, and a questionnaire targeted on women of child-bearing age, usually those age 15-49.[4] Often, an individual questionnaire for men age 15-59 is also included.

Surveys from Phase I of the DHS programme (those conducted through 1989) included few questions on education. Subsequent phases of the DHS have included more questions on educational attainment among household members, literacy among adults, and school participation among youth. In the latest phase of the DHS, MEASURE *DHS+* (1997-2002), women and men were asked to demonstrate literacy by reading from a sentence provided them. This approach is an improvement on the self-reported literacy measurement previously used.[5] In addition, the latest DHS includes a question on adults' exposure to literacy-fostering programmes. The household questionnaire was also revised to include questions on school participation among youth over a two-year period so that student repetition and dropout ratios could be calculated in addition to school attendance ratios.

Two other data collection activities bear mention: the Service Provision Assessment (SPA) and the DHS EdData Survey. Either in conjunction with the DHS or separately, some countries conduct a SPA, which collects data from health providers and communities on the characteristics of health services and family planning services available in a given country, as well as information on nearby schools. A new USAID activity, DHS EdData, conducts education surveys that are statistically linked to the DHS. DHS EdData Surveys provide data on factors affecting household demand for schooling. The goal is to provide information on household decisions regarding how households invest in education.[6]

[4] Some surveys include all women in this age range; others include only currently-married or ever-married women.

[5] In this instance, "adults" refers to those women and men interviewed individually through the women's and men's questionnaires, rather than referring to all adults in the household.

[6] Specific topics in the DHS EdData core survey include: the reasons for school-age children never having attended school or having dropped out of school; reasons for over-age first-time school enrollment; household expenditures on schooling and other contributions to schooling; parents'/ guardians' perceptions of the benefits of schooling and of school quality; distances and travel times to schools; age of children's first school attendance and dropout; and the frequency of and reasons for student absenteeism.

Living Standards Measurement Studies (LSMS)

The Living Standards Measurement Study (LSMS) was initiated by the World Bank in 1980 to assist national statistics institutes improve their ability to collect good data, and to use those data for policy-making. An LSMS survey collects comprehensive data on most aspects of household welfare: consumption; income from activities in the labour market, household enterprises or agriculture; asset ownership; migration; health; education; nutrition; fertility; savings and credit; and anthropometrics. Typical LSMS survey sample sizes range from 2,000 to 5,000 households.[7] A typical LSMS survey includes three different kinds of questionnaires: the household questionnaire to collect information on all household members; a community questionnaire to gather information on local conditions that are common to all households in the area; and a price questionnaire to permit adjustments in countries where prices vary considerably among regions. Facilities questionnaires are sometimes used to gather detailed information on schools or health facilities. LSMS surveys are often administered over the course of an entire year to be able to adjust for seasonal variations.

The standard education module of an LSMS survey collects data for all household members who are primary school age or older. Most LSMS surveys also solicit information on pre-school age children regarding their participation in various programmes such as government-provided early childhood or school feeding programmes. The standard education module includes questions on self-reported literacy and numeracy; whether children have ever attended school; current school attendance; school completion; current enrollment; current grade/level; repetition; highest degree/diploma received; attendance at private/public/religious school; distance to school; and transportation method used to go to school.[8] For those who attended school in the past 12 months, data are collected on household expenditures on schooling and receipt of scholarships. In a few countries, children were given brief literacy and numeracy tests. The LSMS provides the most comprehensive education information among the surveys discussed in this Guide.

In 1987, the World Bank implemented another programme called the Social Dimensions of Adjustment (SDA) in Sub-Saharan Africa. Under this programme, several different types of surveys were developed: the Integrated Survey, which is identical to the LSMS survey; the Priority Survey, which is a relatively simple household survey aimed at identifying policy target groups. It provides a mechanism to produce key socio-economic variables easily on a regular basis in order to describe and monitor the well-being of different household groups. Finally, the Core Welfare Indicators Questionnaire (CWIQ) uses simple indicators to identify who benefits from social programmes. On education issues, the CWIQ indicators are usually limited to distance to the nearest school; number of children enrolled in school by age and sex; and household satisfaction with the quality of the schooling.

[7] LSMS survey samples are designed to represent the population of the country as a whole as well as that of certain subgroups of the population, and are small to balance sampling and non-sampling errors.

[8] The actual content of the module varies by country.

Multiple Indicators Cluster Surveys (MICS)

The Multiple Indicator Cluster Survey (MICS) methodology was developed by UNICEF to support governments and other partners in measuring progress for children and women. UNICEF has conducted two rounds of the Multiple Indicator Cluster Surveys (MICS): first in 1995-1996 as part of the mid-decade review; and subsequently in 1999-2000 (called MICS2), as part of the end-of-decade review of progress toward the goals of the 1990 World Summit for Children. In addition to assessing progress over the last decade, the data can be used for setting the baseline for the next decade and for monitoring programmes.

The MICS studies also provide data on additional children's rights measures and on other UNICEF priority topics, including the IMCI initiative and malaria, and HIV/AIDS. Under MICS, about 60 surveys were conducted; under MICS2, close to 70 surveys have been conducted to date, with sample sizes varying from about 2,000 to 20,000 households.

The MICS2 includes a household questionnaire, a questionnaire for women age 15-49, and a questionnaire to collect data on children under age 5. MICS2 includes questions on literacy and early childhood education. The education questions in MICS2 include self-reported literacy and educational attainment for household members age 5 and older, and school attendance for children age 5-17.

Statistical Information and Monitoring Programme on Child-Labour (ILO)

Based in the Bureau of Statistics of the ILO, the Statistical Information and Monitoring Programme On Child-Labour (SIMPOC) made its debut in 1998 as the statistical branch of the International Programme on the Elimination of Child Labour (IPEC). Its mandate is to create a statistical knowledge-base on child labour. Since its inception, the programme expanded many-fold, and is currently in different stages of completion of close to 150 surveys in about 50 countries. Although its show-case is household-based surveys, since the universal adoption of Convention No.182 on the worst forms of child labour (WFCL) in 1999, the programme has also been undertaking surveys of limited scope and coverage with greater focus on the WFCL – rapid assessment, base-line, street children, school-based, and establishment-based surveys.

SIMPOC household-based surveys run at a minimum of about 10,000 households as the sample size, rising to as high as 120,000 households in some large countries. They are designed to be large enough to allow for disaggregation of key indicators down to the district level in order to facilitate the development of comprehensive policies and programmes against child labour.

A typical SIMPOC questionnaire is divided into three principal sections – household characteristics, parents' section and children's section. Focusing on boys and girls between the ages of 5 and 17, the aim of these surveys is to draw a quantitative picture of the economic and non-economic activities (such as household chores) of children; demographic, social and economic characteristics of household members; working hours and conditions; health and safety issues including injuries at work and perceptions of parents about children's work. On the education front, in addition to capturing children's school participation, these surveys also gather information on literacy and school attendance of children. SIMPOC surveys also examine the effect of children's work on their schooling and the relationship between parents' socio-economic background and child labour.

Centers for Disease Control and Prevention (CDC) Surveys

The Centers for Disease Control and Prevention (CDC) conducts Reproductive Health Surveys, focusing on females and males of reproductive age. Since 1990, the CDC has conducted about 20 surveys in Eastern Europe and Latin America. Sample sizes vary from 6,000 to 20,000 respondents, depending on a country's data needs. The CDC surveys do not employ a core questionnaire that is used in each country; instead, the questionnaires are individually tailored to the needs and interests of each country.

The household questionnaire does not include education questions. Education data are collected only for individual men and women of reproductive age and/or youth (15-24), depending on the particular survey design. As a consequence, indicators on children's school attendance are generally not available, although some recent surveys have included questions asked of women about the school attendance of their children age 5-14. In most CDC surveys, the only education information collected is men's and women's educational attainment, although in some countries additional questions on women's education are asked.

The Network of Migration and Urbanization Surveys in West Africa/le réseau d'enquêtes migration et urbanisation en Afrique de l'Ouest (REMUAO)

The Migration and Urbanization in West Africa Network (REMUAO) conducted eight migration surveys in 1993 in the following countries: Burkina Faso, Côte d'Ivoire, Guinea, Mali, Mauritania, Niger, Nigeria, and Senegal. The surveys aimed at measuring the flows, characteristics, attitudes and aspirations of migrants; the determinants and consequences of migration; and the degree of societal integration. In Burkina Faso, a module on reproductive behavior was included. Furthermore, modules on agricultural operations and utilisation of revenues were added in Côte d'Ivoire. The surveys addressed migration both within and across countries. Sample sizes ranged from about 7,000 to 13,000 households. The age range of interest among respondents varied by country, with some surveys interviewing migrants age 6 and older and some migrants age 15 and older. The surveys collected information on the level of education attained, literacy, and current schooling status. The surveys also collected data on education as a factor in the decision to migrate. The data are nationally representative and can be disaggregated by urban-rural residence.

Other household surveys

The surveys listed above are not an exhaustive list of household surveys providing education data. A number of other household surveys and survey programmes offer data on education. In Appendix B, these surveys are described briefly and information about how to access data sets is provided. Appendix B also gives contact information for several organizations together with useful links to household survey data sets or databanks.

SURVEY METHODOLOGY AND THE USE OF EDUCATION DATA 4

This chapter discusses how certain methodological limitations of multiple-topic household surveys and national censuses affect the ways in which the data should be analysed. While this chapter discusses these constraints in general, it is important to become familiar with the design of the survey when using data from survey reports and in the secondary analysis of these data. For example, survey results can be used better and the data set manipulated further if we know more on the questions asked about education, the timing of survey fieldwork, the sampling frame, and the survey design and weighting/ estimation scheme. Data users should access not only the data set and the report of results, but also the supporting documentation, which includes the survey questionnaires and interviewer manuals prepared for a particular survey.

4.1 DEFINING THE HOUSEHOLD

By necessity, national censuses and household surveys must first define what constitutes a household. Generally, a household is defined as a person or group of persons that usually lives and eats together. A household is not the same as a family since a family only includes people who are related; rather, a household includes all people who live together, whether they are related or not. For example, three unrelated men who live and cook meals together would not be considered as one family but as one household. A member of the household is any person who usually lives in the household. Likewise, national censuses and surveys specify what constitutes a visitor: a visitor is someone who is not a usual member of the household, but who did sleep in the household the night before the day of the interview.[9]

As discussed in Chapter 2, patterns in education participation are commonly examined in terms of various household characteristics, including the sex of the household head, parents' educational attainment, household wealth, and so on. However, the above approach to defining a household may not capture important information on other family influences on schooling outcomes. For example, the head of the household may be identified as a woman who is the mother of the children in the household. At the same time, even if he lives elsewhere, the children's father may exert considerable influence over whether the children ever attend school, how much money is spent on schooling, and other decisions within the household. In addition, perhaps an uncle (living in another household) provides significant support for the schooling of his nieces and nephews. In such a case, the wealth or household asset measure may underestimate the resources (remittances) available to support the children of the household, and the classification of the household as female-headed will not capture the potential influence of the children's father on decision-making processes.

Another example of the difficulty of capturing and interpreting family relationships is the practice of fostering children: Children may be sent to live with relatives or non-relatives

[9] These specific definitions are adapted from the DHS Interviewer's Manual.

primarily so that they can be educated; other children, however, are sent to provide domestic and/or agricultural labour in those households. It is reasonable to assume that children fostered to households to provide labour are less likely to enroll and continue in school than children living with their own parents or living with other households for the primary purpose of attending school. Large-scale household surveys provide information on the relationship between a child and the household head and whether the child lives with his/her parents, but do not elaborate on the specific circumstances of a child living away from his/her parents.

In short, the typical approach to defining the household does not capture the complexities of kinship and extended family structures operating in many countries. As a result, variables included in household survey and national census data sets may not completely capture the various influences on schooling decisions made within those households. In spite of these limitations, national census and household survey data in the aggregate allow patterns of school participation to be examined according to important household characteristics.

4.2 VARIATION IN MEASURES OF EDUCATIONAL PARTICIPATION

Survey questions about educational attainment and current school attendance/ enrollment may be phrased quite differently. A survey may ask "*What is the highest grade that (name of person) has completed?"* or "*What is the highest grade that (name of person) has attended*?" Attending a particular grade is not the same as having passed or successfully completed a particular grade and these differences should be taken into account in data analysis. Selected surveys, notably many of the LSMS surveys, may also provide data on education qualifications by asking about the highest diploma or certificate earned.

Some surveys collect information on the grade that household members currently attend or the grade attended in the last school year. It should be noted that "currently" attending or "still" attending school typically refers to whether a child generally attends school, and not to attendance on that particular day or even during that particular week. Other surveys—notably the early DHS surveys—inquire as to the highest level/grade completed by household members, followed by a question about whether the person still or currently attends school. In the latter case, in order to create current attendance indicators such as NAR or GAR, an assumption is required about the level/grade currently attended by the household member. For example, if a child has completed grade 4 and currently attends school, it is assumed that the child currently attends grade 5.[10]

In order to calculate the NAR and GAR from this set of questions, certain assumptions must be made, which may not hold: for instance, that a child age 13 who completed grade 6 and still attends school currently attends grade 7. The child may be repeating

[10] In the previous rounds of the DHS, interviewers were instructed that completion of a grade refers to successful completion of that grade, rather than mere attendance of that grade. These instructions were intended to permit the assumption that the child, if he/she "currently" attends school, attends the next grade after the highest grade completed.

grade 6. If the age range for primary school is age 7 to 12, and for secondary school is age 13 to 18, by assuming the child is now in the first year of secondary school, we would overestimate the NAR at the secondary level (and underestimate the GAR at the primary level), particularly in countries where repetition is common in the last year of primary school.

As mentioned in Chapter 2, in recent years, both the DHS and MICS surveys ask about household members' school attendance for the last two years. This enables dropout and repetition ratios to be calculated. These revised questions specify the school year and ask about attendance during that year, and as a result, no assumptions are made as to the level/grade currently attending.

As described above, questions about education may be worded differently across surveys, while collecting essentially the same information. In other cases, surveys collect different kinds of information about attainment or participation. Knowing how questionnaires were designed and what kind of information is collected is essential to the proper use of the results.

4.3 TIMING AND DURATION OF SURVEY FIELDWORK

When considering education data from multiple-topic surveys, it is important to know at what time of year, and over how long a period of time data are collected, particularly data on children's school attendance or enrolment. A national census is generally implemented over a short period of time, and so household respondents are likely to answer questions about schooling during one given school term. Household surveys, on the other hand, may take three to five months or longer to implement, and as a consequence, data on "current" school attendance may be collected during two school terms or years, over school vacations, or some combination thereof.

Under these conditions, questions arise as to how to interpret data on current attendance. Do data on attendance (or non-attendance) refer to the school year which was just completed or to the coming school year? The DHS survey conducted in Madagascar in 1992 is a case in point: according to the Malagasy school calendar, the between-year school vacation runs from July through August. Because the DHS was in the field from May to November, the survey covered two school years and the vacation between years. As a consequence, it is impossible to know whether children in households surveyed before the beginning of the 1992-1993 school year (surveyed from May to the end of August) were going to attend school in the next school year. For children whose households were surveyed from September to the end of November, it is not known whether they attended school during the previous school year (1991-1992). As in the case of Madagascar, survey reports generally present data on school attendance as though they were from a single school year even though that may not be the case. Under the circumstances, the school attendance data presented in this report must be interpreted cautiously.[11] In using the data set, of course, data on school

[11] In recent years, DHS and MICS have made more explicit the instructions on handling questions about education participation when some of the survey fieldwork is conducted between two school years.

attendance might be studied separately for households surveyed during the 1991-1992 school year, over the vacation, and during the 1992-1993 school year.

A more serious problem in the calculation of children's school attendance ratios is the fact that a national census or household survey collects data on children's school attendance during the "current" school year, which may have begun many months before the data were collected. By contrast, data on age usually are collected in completed years during the national census or household survey. As a result, the net attendance ratio, for example, is calculated for children who were of primary school age, say 6-12 at the time the data were collected, rather than for those who were age 6-12 at the time the school year began. In other words, the estimate of NAR includes some children who were actually age 5 (not 6) at the time the school year began. The NAR calculation also excludes some children who were 12 (not 13), at the time the school year began. If school attendance ratios are low among children who are under-age (in this case, age 5 or younger) and high among those age 12, the NAR will probably be underestimated. This becomes more problematic the greater the time between the beginning of the school year and the data collection, and the longer the duration of the survey.

Furthermore, children's attendance ratios may fluctuate considerably depending on the point at which the survey is conducted. For example, attendance ratios very early in the school year may be low in countries where children may not start attending school in the first few weeks of the school year; in countries with significant school dropout during the school year, attendance ratios near the end of the year may be noticeably lower than at the beginning of the school year. School attendance ratios may be lower at certain points during the year if there is seasonal variation owing to the demand for agricultural and herding work.

In summary, two identical surveys conducted in the same year in the same country over two periods of time might produce markedly different school attendance ratios. It cannot be assumed that one set of results is less valid than the other. In comparing within-country data at different points in time or comparing data across countries, the timing of the survey relative to the timing of the school year must be taken into account.

4.4 SAMPLE SIZES AND REPRESENTATIVENESS OF DATA

A national census is designed to collect information from all households in a country. As a consequence, at least in theory, the entire household population is included in data collection. In practice, of course, not every single household provides information in a national census and so the issue of non-response arises. This issue will be discussed more fully later in this chapter.

By contrast, a household survey is designed to provide a "snapshot" of the population using a representative sample of households; thus, by design not all households in the population are present in the survey. What constitutes a representative sample depends on the survey design, which is influenced by two factors: the level of accuracy sought in the estimates for various indicators; and the level of data disaggregation.

Many surveys, for instance, provide estimates of indicators (such as adult educational attainment) at the national level, for urban and rural areas, and for regions of the country. However, a larger overall sample size will be required to calculate sub-regional estimates simply because a large sample size will enable reasonably accurate indicators to be estimated.

Many multiple-topic surveys are not specifically designed to study education, possibly complicating the analysis of education data. For example, in addition to the standard education questions that allow the NAR and GAR to be calculated, the 1999 Guinea DHS included a set of questions on why children do not attend school, how much money is spent on schooling, and other similar questions, for children age 6-15 who lived with their mothers age 15-49.[12] As a consequence, data on some questions were not collected for children who were not living with their mothers or who lived with their mothers who happened to be age 50 or older. Hence, the data on reasons for not attending school do not represent all children age 6-15 in Guinea, but only children age 6-15 living with their mothers age 15-49 at the time of the survey. Data analysis is complicated by the limitations of the sub-sample of children for whom data were collected.[13]

In doing secondary analysis of existing data sets from sample surveys, it is critical to understand the survey design. Typically, surveys are designed in such a way that the entire population is stratified (often geographically) into regions for which estimates are desired. Finer sub-stratification will take place if estimates are desired at even lower levels or sub-regions. Then sufficient sample size is allocated to each of these substrata or sub-regions, and sample units are drawn within the sub-region according to a predetermined sampling scheme involving some randomisation procedure.

While a sample survey may be designed to produce estimates at the province or state level, lower-level estimates can be produced even if the sampled households or individuals at this lower level are not drawn in such a way as to be representative of the entire population. However, sufficient sample size at this lower level is essential to produce accurate estimates of indicators.

In fact, it is possible to produce estimates for levels that even cut across geographic design strata where representativity may again be an issue, such as in the case of age groups, provided there is sufficient sample size within each age group. Such estimates are called “domain estimates”. In fact, an estimate for a primary school NAR is an example of such a domain estimate that cuts across geographic design strata. Since most household surveys do not use age groups as design strata, and hence cannot ensure *a priori* the sample size within the age group, the sample size that is obtained for this group is a random outcome of the survey. Yet, it is still reasonable to produce estimates for this group, since it is likely that the sample size will be sufficiently large,

[12] The 1998 National Family Health Survey conducted by the CDC in El Salvador included a similar education module which was asked only of women age 15-49 who had children age 7-14, and so suffers from the same limitations.

[13] For further discussion of this issue, see the report of survey results: *Schooling in Guinea: Findings from the GDHS-II 1999*.

given that the occurrence of 6-12 year olds in a survey is not uncommon. If it is believed that the representativity is adversely affected, often a "calibration" procedure (also called "post-stratification" or "benchmarking") will be implemented. In essence, a weighting up to external and independent demographic counts of age is made, which has the effect of pulling the representativity back into line. This procedure is embedded in the final weights and is transparent to the analyst. However, the analyst should check the accompanying documentation to see if this has been done.

Finally, for specific information on the sampling designs of the household surveys cited in Chapter 3, data users should consult survey reports and/or contact representatives of survey organizations (*see Appendix B*).

4.5 General Guidance in Data Analysis

The first step in using a data set is to become familiar with its structure and the nature of its variables, the circumstances of data collection, and any limitations on the use of the data set. The documentation for a national census or household survey, such as reports and a codebook, will provide important background information on the survey, such as sample size and data quality indicators.

Generally speaking, data files made available for analysis should be "cleaned". These files will have been checked for structural and range errors and edited for internal consistency. Provisions that compensate for non-response should also be incorporated into the files and fully explained in the accompanying documentation.

Data manipulation and analysis can be demanding and complex. This brief section of the Guide does not intend to provide a comprehensive set of guidelines for the use of data sets; instead, this section reviews several key issues to be considered in analysing survey data.

Try to familiarise yourself with how the data set is structured and what is the appropriate way to form estimates using the given data set

For example, find out whether records within the data files are at the household or individual level and whether household or individual weights are intended to be used in estimation procedures. To clarify, if each record in the data set of a household survey represents a sampled individual, then typically there is a weight associated with that record or individual as well. Census files will typically contain one record for each household or individual in the population; thus, there will be no need for weights since each individual represents himself or herself in the population. In contrast, sample surveys will contain one record for each sampled individual or household (as opposed to the entire population). It is in this context that weights arise since individual records represent or "carry the weight" of others in the population with similar attributes. The weights are calculated in advance of survey taking and are directly related to the randomisation procedure used in drawing the sample. Further details on how weights are formed can be found in any basic sampling textbook (Särndal *et al*, 1992; Cochran, 1977*).*

To form estimates that correctly represent the population coming from sample surveys, one must take the weighted sum of the attribute of interest over all the sample records. To illustrate, consider the example of primary school NAR, where this is defined as the number of children age 6-12 attending primary school divided by the number of children age 6-12. In estimating the numerator, one would sum the weights of all children age 6-12 in the sample survey attending primary school; in estimating the denominator, one would sum the weights of all children age 6-12 in the sample survey. To form the final estimator, one would divide the estimated numerator by the estimated denominator. Weights can automatically be taken into account when forming estimators through the use of standard statistical packages such as SAS, SPSS or BMDP. These packages often have "weight" or equivalent options that enable weighted sums to be calculated.

Try to familiarise yourself with the variables, before using them in calculations

If possible, check the questionnaires to see the source questions for variables in order to understand better how to use the data. Be sure to check how the variable is coded (i.e., male=1, female=2 for a variable on sex) and how it may and may not be used.

Replicate published results before proceeding with additional calculations

If there are reports of results from the data collection activity, try to replicate these results before calculating any new indicators. Sorting out the difficulties with calculations already done will bolster your confidence in producing new results.

Be sure to consider the issue of missing values

Non-response in a survey or census can happen in one of two ways. First the entire record representing an individual or household can be missing since the individual or household refused to answer, was not available, could not be contacted, etc; this is called "total non-response".

The second type of non-response arises when attributes or variables within a record are missing and is termed "item non-response". For item non-response, some variables may have a low incidence of missing values, such as the age of each household member since this is easy to collect; other variables, such as whether a youth attends school during the current school year, may have large numbers of missing values because the question was not asked or known for all household members. National census and survey efforts handle missing values in different ways. Regardless of the chosen method used, compensation needs to be made for non-respondents since they often have attributes that are distinctly different from those who respond. Therefore the resulting indicators can be severely biased if we ignore the non-response and assume that the two groups behave similarly.

A technique called imputation is often used to compensate for missing data values in the case of item non-response. Using this method, missing values are placed by those that are believed to represent those missing values. A range of methods can be employed for undertaking imputation (Groves *et al*, 2001). The resulting complete or "square" file allows indicators to be estimated. National statistical offices will sometimes

release files without compensating for the missing item values; it should be noted that estimates of indicators based on these files will render analyses that are biased.

In the case of total non-response, weight adjustments are often used to compensate rather than imputing all the variables separately for a missing record. For the weight adjustment method, the non-responses are not filled in but rather are omitted from the file. To compensate, the weights of respondent records are boosted up within predetermined "weighting classes" to compensate for those individuals or households that are not present (who in turn represent individuals or households in the population that must be accounted for). In this case, what is typically presented to the analyst is a file having "final weights" that include the weight adjustments for non-response on the respondent records. If the analyst verifies that the weight adjustment is embedded in the final weight, indicators can be established as described above without concern for the missing values.

However, in the cases of both imputation and weight adjustment methods, special variance or standard error formulae should be used to compensate for the variation due to non-response. Often, however, analysts ignore this component and simply use usual textbook formulae for variances. This is reasonable to do provided the overall non-response ratio in the survey is not too high (say, less than 10%) and thus the contribution to the overall variance coming from non-response can be assumed to be low.

Be sure to calculate a measure of accuracy of the basic estimate (coefficient of variation) in order to be able to make a determination regarding the reliability of the estimates of indicators

Depending on the overall size of your sample, some tabulations may yield cells that are based on very small numbers of cases. This may happen, for example, when you are tabulating results by categories of background variables in which relatively few respondents fall (e.g. women with higher education, as in Table 4 presented in Chapter 5). These estimates may not be reliable. For this reason, it is paramount to calculate some measure of accuracy and to disseminate it alongside the basic estimate to be able to gauge the reliability of all estimates produced. A good rule of thumb in this regard is to use the coefficient of variation (CV).

The coefficient of variation (CV) is defined as the square root of the variance divided by the estimate itself and multiplied by 100 – expressed as a percentage.

For example, in the case of the primary school NAR, the CV of the primary NAR is defined as the square root of the variance of the primary school NAR divided by the estimate of the primary school NAR multiplied by 100. One of the advantages of using CVs as measures of reliability is that they are scale-less, permitting comparisons with other estimates that are measured in entirely different units. Often national statistical offices will advocate basic quality guidelines that say that estimates having CVs greater than 35% should not be used to draw statistical inferences and should not be released to the public. Be sure to properly account for complex survey designs in your analysis, particularly when calculating variances.

Often household surveys will have complex sample designs, involving many levels of stratification, clustering, multiple stages or phases of sample selection, or unequal probability sample weighting schemes. The analyst needs to be aware of these constructs and to properly account for them. These complex schemes require special attention when producing estimates, and particularly their variances. Incorrect variances can occur if we use statistical packages that do not properly take these complex designs into account. Sometimes sample surveys will have simple designs (such as using Stratified Simple Random Sampling procedures only); in this case, ignoring the sample design by a straightforward application of standard statistical packages will still give the "right answer", so it is really the case of complex sample surveys that need special consideration.

In general, national population censuses collect data on all households and individuals in the population so sample design and weighting are not at issue. The only exception is when a different questionnaire with more detailed questions is presented to a sampled fraction of the population. But even then, issues of complex survey designs do not tend to surface since very simple designs (such as Stratified Simple Random Sampling) tend to be used.

In the case of complex sample survey designs, forming the estimate itself (for example, primary school net attendance ratio (NAR)) is not an issue since it is easy to take the design into account by simply inserting the survey weight w_i into the estimator as discussed. The critical issue is rather variance estimation and thus standard error estimation and CV estimation. If an analyst naively uses the textbook variance for an estimator based on a complex survey, the variance can be wrong and off by several orders of magnitude. And simply inserting the weight into the variance to account for the design (often called a "naive estimator of the variance") goes only part way towards correcting the problem, but it can still be quite off. Variances under complex survey designs need to be derived separately; again, there is an large body of literature on the subject (Särndal *et al,* 1992; Cochran, 1977).

Furthermore, many statistical packages do not have options for taking into account variances based on complex surveys. However, increasingly, modules are being added to packages to allow for these more complicated scenarios. At a minimum, analysts should be made aware of these problems and pitfalls and should be consulting with statisticians for help.

Note that a further complication in calculating variances arises from the fact that most education indicators of interest, particularly those cited in Chapter 2, take the form of ratios. These are “non-linear” estimators of indicators and are relatively complex, since the variance is due to both an estimated numerator and denominator portion; thus, variances for these kinds of indicators need to be calculated carefully. Some statistical packages (SPSS, for example) allow ratio estimates to be calculated directly. When this is not the case, the advice of a survey statistician should be sought (Wolter, 1985).

PRESENTATION OF HOUSEHOLD SURVEY DATA 5

This chapter focuses on issues that might be of great interest to Ministries of Education (MOEs) and survey organisers when considering how to present data from surveys in the most useful way for policy-makers. The focus here is on descriptive data presented in cross tabulations.

The data presented in these tables and figures show correlations between variables. For instance, more children living in urban areas may attend primary school than children living in rural areas. This means that there is a correlation, or a relationship, between urban-rural residence and the likelihood of attending school. The data do not provide information on causation or quantify the effect of living in an urban or rural area on the likelihood of a child attending primary school. Cross tabulations do not take into account other variables that may also affect the likelihood of attending primary school—such as household wealth, region of residence, parents' education, and other factors. In order to quantify the effect of a particular factor on an outcome of interest, other statistical methods are required.

However, correlations between variables, such as wealth and school participation, are important and interpretable. For instance, if poor children are less likely than wealthier children to attend school, policies affecting the costs and returns of schooling to households need to be reviewed.

The tables and figures below make no attempt to cover the wide range of indicators discussed earlier in this Guide; instead, they provide general guidance on how to present data in user-friendly formats. Not all household surveys provide data on the variables used in the following examples.

5.1 SELECTING AND PRESENTING STANDARD EDUCATION INDICATORS

Attendance

All of the data sources discussed in this Guide include data on school attendance. As illustrated by Tables 1 and 2, the audience and the intent of the table or figure should determine the background characteristics to be presented. Table 1 presents the percentage of primary school age children attending primary school in Azerbaijan, based on the 2000 MICS. The data are disaggregated by age, region, area, household status and household wealth. Note that the household status of interest in Azerbaijan is whether children are residents or refugees; this may not be relevant in other countries.

Table 1
Percentage of children of primary school age attending primary school, Azerbaijan, 2000 MICS

		Male		Female		Total	
		Percent Attending	Number	Percent Attending	Number	Percent Attending	Number
Age	7	71.5	284	72.2	241	71.8	525
	8	94.8	328	93.5	277	94.2	605
	9	96.0	328	96.1	305	96.1	633
	10	88.7	337	89.0	317	88.8	654
Region	Baku area	92.5	322	90.6	287	91.6	609
	Nakhchivan	95.0	40	94.0	50	94.4	90
	Center, North	87.8	483	88.8	445	88.3	928
	West, Southwest	84.8	250	85.5	207	85.1	457
	South	85.7	182	85.4	151	85.6	333
Area	Urban	89.4	634	90.5	598	89.9	1232
	Rural	87.2	643	86.2	542	86.8	1185
Household Status	Resident	88.4	1145	88.2	1031	88.3	2176
	IDP or Refugee	87.9	132	90.8	109	89.2	241
Household Wealth	Poor	85.2	318	86.2	239	85.6	557
	Middle	88.5	793	88.6	735	88.5	1528
	Rich	93.4	166	91.0	166	92.2	332
Total		**88.3**	**1277**	**88.4**	**1140**	**88.4**	**2417**

Produced using data from the 1999 Guinea Demographic and Health Survey, Table 2 presents similar information to the previous table. In contrast, Table 2 presents data on both net and gross attendance ratios (NAR and GAR) at the primary and secondary school levels, as well as the gender parity index. Data are presented according to girls' and boys' background characteristics—residence, household wealth and mother's education, enabling indicators to be compared across groups.

This table provides a good example of country-specific analysis. As school participation in Guinea is relatively low, it is useful to view participation ratios among children of all ages. The table shows that while only 46% of boys age 7-12 attend primary school, the gross attendance ratio is 74, suggesting that for every 46 boys of primary school age attending primary school, 28 boys over age 13 attend primary school. Both the net and gross attendance ratios are shown to emphasise the large numbers of over-age children attending school. In addition, in Guinea, unlike in Azerbaijan, there is a substantial gender gap in primary school participation. Presenting the gender parity index in the table emphasises the gender gap.

Table 2
School attendance rate, Guinea, 1999 DHS

Net attendance ratio (NAR) and gross attendance ratios (GAR) for the *de facto* household population age 6-24, by sex and selected background characteristics

	PRIMARY							
	Net attendance ratio			Gross attendance ratio				
	Male	Female	Total	Male	Female	Total	Gender parity index	Sample size
Residence								
Urban	77.0	63.7	70.1	129.7	95.0	111.6	0.7	1,869
Rural	33.6	20.5	27.1	52.2	27.5	40.0	0.5	4,629
Region								
Lower Guinea	47.7	36.4	42.3	73.0	51.9	62.9	0.7	1,325
Middle Guinea	27.5	19.4	23.3	43.4	26.0	34.5	0.6	1,776
Upper Guinea	28.0	18.6	23.4	49.1	26.2	38.1	0.5	1,069
Forest Guinea	54.2	33.2	43.5	89.5	43.7	66.1	0.5	1,397
Conakry	86.2	71.7	78.6	139.7	111.0	124.7	0.8	931
Household asset index								
Poorest quintile	23.7	11.8	17.6	38.4	15.7	26.8	0.4	1,289
Second quintile	26.7	11.3	19.3	39.3	14.5	27.3	0.4	1,305
Middle quintile	37.4	23.9	31.0	58.0	31.0	451.0	0.5	1,285
Fourth quintile	59.8	44.4	52.2	99.3	62.0	80.8	0.6	1,345
Richest quintile	83.3	72.1	77.3	138.7	109.6	123.0	0.8	1,277
Mother's education								
No education	27.7	27.7	35.1	60.6	38.0	49.5	0.6	4,018
Some education	72.1	72.1	76.2	118.3	98.3	108.2	0.8	579
Mother not in household	33.7	33.7	38.1	91.0	52.6	70.9	0.6	1,802
Total	33.4	33.4	39.5	73.6	47.7	60.6	0.6	6,499

	SECONDARY							
Residence								
Urban	32.1	17.4	24.9	52.2	24.3	38.5	0.5	1,700
Rural	8.7	2.3	5.7	11.4	2.9	7.4	0.2	2,847
Region								
Lower Guinea	23.5	5.0	14.3	29.0	6.3	17.8	0.2	895
Middle Guinea	10.8	5.9	8.5	14.9	6.6	11.1	0.4	1,057
Upper Guinea	6.5	4.4	5.6	11.0	5.7	8.7	0.5	728
Forest Guinea	16.1	4.6	10.8	14.9	7.2	16.7	0.3	1,012
Conakry	32.3	19.3	25.3	57.7	28.0	41.6	0.5	856
Household asset index								
Poorest quintile	4.4	0.3	2.4	7.0	0.3	3.6	0.0	778
Second quintile	5.2	0.6	3.1	6.8	0.6	3.9	0.1	730
Middle quintile	10.0	2.6	6.7	13.3	2.9	8.6	0.2	854
Fourth quintile	19.9	9.4	15.0	31.0	12.9	22.5	0.4	994
Richest quintile	36.8	19.9	28.4	57.9	28.0	42.9	0.5	1,191
Mother's education								
No education	5.4	3.0	4.3	5.8	3.3	4.7	0.6	1,740
Some education	19.6	18.0	18.8	24.4	21.0	22.8	0.9	193
Mother not in household	23.1	9.5	16.4	36.7	13.6	25.2	0.4	2,978
Total	**17.2**	**8.1**	**12.9**	**26.3**	**11.1**	**19.0**	**0.4**	**4,547**

Often, the completion of fourth grade is considered to be a proxy for literacy. Therefore, the persistence or "survival" of grade 1 students to grade 5 becomes critically important (*see Appendix A*). Table 3 provides indicators of educational persistence for Chad. Because each cell has a different denominator, they are not presented in this particular table. The table shows that a large number of children starting grade 1 do not reach grade 5.

Table 3
Percentage of children entering first grade of primary school who eventually reach grade 5,
Chad, 2000 MICS

		Percent in grade 1 reaching grade 2	Percent in grade 2 reaching grade 3	Percent in grade 3 reaching grade 4	Percent in grade 4 reaching grade 5	Percent who reach grade 5 of those who enter grade 1
Household asset index	Poorest	84.4	88.1	89.8	90.2	60.2
	Second	93.3	98.6	93.1	94.9	81.3
	Middle	82.9	88.1	90.5	91.9	60.7
	Fourth	85.0	86.5	87.5	88.6	57.0
	Richest	79.7	86.1	86.4	90.8	53.8
Sex	Male	85.3	89.4	88.3	91.3	61.5
	Female	84.9	89.6	90.4	91.0	62.6
Place of residence	Urban	86.2	89.4	88.8	89.1	60.9
	Rural	84.8	89.5	89.2	92.9	62.9
Region	N'Djaména	83.2	88.8	86.4	89.4	57.0
	Other villages	88.9	89.7	91.2	88.7	64.6
	Rural	84.8	89.5	89.2	92.9	62.9
Total		**85.1**	**89.5**	**89.0**	**91.2**	**61.8**

Some surveys include data on early childhood education programmes in recognition of their importance. Table 4 presents data on early childhood education for Côte d'Ivoire from the 2000 MICS. The table presents the distribution by sex, region, age and mother's education. Note that the more education the child's mother has completed, the more likely the child is to attend an early childhood programme.

Table 4
Percentage of children age 36-59 months who are attending some form of organised early childhood education programme, Côte d'Ivoire, 2000 MICS

		Percent attending programme	Number of children
Sex	Male	6.8	1,422
	Female	5.7	1,412
Region	Central	3.4	152
	North Central	7.6	251
	North East	4.8	94
	Centre East	3.5	74
	South (without Abidjan)	4.4	478
	South West	6.5	258
	West Central	1.8	397
	West	7.0	334
	North West	6.8	118
	North	1.7	198
	Abidjan	13.7	480
Area	Urban	12.6	1,228
	Rural	1.4	1,606
Age	36-47 months	4.2	1,586
	48-59 months	8.8	1,248
Mother's education	None	3.0	1,985
	Primary	7.1	565
	Secondary	23.1	241
	Tertiary	64.7	26
	Non-formal Training	25.0	15
	Missing/Don't Know	16.8	2
Total		**6.2**	**2,834**

The numerator includes children for whom the mother/caretaker indicated that the child was attending some form of organised early childhood education programme. The denominator is children age 36-59 months.

Literacy

Questions on literacy are also frequently asked in household surveys and censuses. These data can also be presented in different ways, depending on the objectives of the researcher and the available data. Taken from the 2000 MICS in Côte d'Ivoire, Table 5 presents data on literacy among men and women age 15 or older by region, area, and age. Reporting of literacy data should include the definition of literacy used and information on how the figures are calculated.

Table 5
Percentage of the population age 15 years and older that is literate, Côte d'Ivoire, 2000 MICS

		Sex					
		Male		Female		Total	
		% Literate	Number	% Literate	Number	% Literate	Number
Region	Central	44.8	790	21.8	962	32.1	1,752
	North Central	56.2	1,254	36.5	1,504	45.5	2,758
	North East	42.5	390	17.0	645	26.6	1,035
	Centre East	57.8	312	32.9	347	44.7	660
	South (w/o Abidjan)	59.2	2,153	36.2	2,528	46.8	4,681
	South West	46.7	1,072	32.5	1,130	39.4	2,202
	West Central	53.9	1,599	25.6	1,833	38.8	3,432
	West	66.4	810	34.6	1,124	48.0	1,934
	North West	26.0	427	9.7	475	17.4	902
	North	39.4	754	21.0	891	29.4	1,645
	Abidjan	83.5	3,262	64.0	3,912	72.9	7,174
Area	Urban	74.1	7,098	53.2	8,395	62.8	15,493
	Rural	42.8	5,717	20.1	6,956	30.3	12,672
Age	15-24	69.6	5,164	51.5	6,042	59.9	11,206
	25-34	67.6	2,613	41.2	3,714	52.1	6,327
	35-44	60.7	1,987	32.3	2,759	44.2	4,747
	45-54	52.0	1,540	17.9	1,464	35.4	3,004
	55-64	26.6	879	5.2	837	16.1	1,715
	65+	17.1	638	3.6	536	11.0	1,173
Total		**60.1**	**12,821**	**38.2**	**15,351**	**48.2**	**28,173**

5.2 PRESENTING CLEAR AND COMPLETE INFORMATION

Tables and figures can be powerful tools for communicating survey results. In presenting data, tables or figures must be clear, internally consistent, and complete. Tables should be complete *per se* and contain all relevant information on the data presented and the population concerned, including, as applicable, the age ranges, the background characteristics and so on. When applicable, tables should also include sample sizes.

Taken from the report on the 1996 LSMS in Ghana, Table 6 defines the indicator presented - school attendance, and presents data by expenditure quintile (Statistical Service, Republic of Ghana and World Bank, 1998). Presentation could be improved in two ways: by adding "expenditure quintile" to the top left column to clarify that "first" through "fifth" and "all" refers to expenditure quintile; and specifying that the table covers youth age 6-25 in the title or below it.

Table 6
School attendance ratio by location, sex and expenditure quintile (%), Ghana, 1996 LSMS

	Accra		Other urban		Rural		Country		
	Male	Female	Male	Female	Male	Female	Male	Female	All
First	37.5	42.9	59.8	43.4	54.9	32.8	55.5	35.5	45.6
Second	65.6	58.1	73.5	50.0	69.2	46.0	70.0	47.7	59.1
Third	68.0	64.3	72.9	56.3	63.5	40.1	66.0	45.5	55.6
Fourth	90.0	62.2	72.6	53.0	59.3	38.2	66.9	45.1	55.9
Fifth	72.2	64.6	39.4	47.1	44.4	20.3	50.4	41.8	46.1
Total	**73.0**	**60.8**	**67.0**	**49.9**	**60.5**	**38.2**	**63.0**	**42.9**	**53.0**

Note: School attendance ratio is the percentage of children currently enrolled in school or having attended school (at any level) during the last twelve months.

The table in the LSMS Ghana report that precedes Table 6, above, presents attendance ratios by location, sex, and age group, and includes ages 6-25; thus, the reader assumes that Table 6 includes the same age groups. However, any doubts regarding the population covered could be removed by specifying the age group in each table.

Whereas Table 6 presents data on children's school attendance ratios in terms of net and gross attendance ratios, not all tables are similarly configured. For example, like many DHS reports of the time, Table 7 is from the report on the 1992 Namibia Demographic and Health Survey. It presents data on children's school participation according to age groups of interest to demographers: age 6-10, 11-15, 6-15, 16-20, and 21-24.[14] Notice that the table presents data on whether youth attended school at the time of the survey without reference to the level of schooling attended. For

[14] The table (2.4 in the original DHS report) reports data on "enrolment", although technically speaking, the table should report data on attendance, since the data are based on the question: "Is (NAME) still in school?"

instance, the table indicates that about 80% of children age 6-15 attended school, but does not specify whether they attended primary or secondary school or both.

Table 7
Percentage of the *de facto* household population attending school, by age group, sex and location, Namibia, 1992 DHS

	Male			Female			Total		
Age group	Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
6-10	73.4	74.4	74.2	76.8	79.3	78.7	75.0	76.9	76.5
11-15	88.0	87.6	87.7	90.8	89.4	89.8	89.4	88.5	88.7
6-15	80.4	80.5	80.5	83.5	84.0	83.9	81.9	82.3	82.2
16-20	61.8	69.0	67.1	46.5	59.2	55.1	53.2	64.2	60.9
21-24	14.8	32.0	25.6	15.7	23.6	20.3	15.3	27.4	22.7

From the perspective of the policy-maker, Table 7 does not present particularly useful data on children's school attendance. The reader of this report will not find data on the net attendance ratio at the primary or secondary school level (for which the appropriate age ranges are 7-13 and 14-18). Further examination of the 1992 DHS Namibia data indicates that over one-third of the primary school learners were over-age for that level of schooling (age 14 or older); therefore, it would not be safe to assume that children age 16-20 who attended school attended at the secondary level. In summary, the primary and secondary school NAR cannot be estimated with information from the results presented in the report on the 1992 Namibia DHS.

One particular disadvantage of not presenting data on school attendance according to established education indicator definitions is that report users may use data on the available age range in place of data on the appropriate age range. For instance, in UNICEF's "The State of the World's Children" 1999 report, which uses the results presented in many DHS survey reports, including that from the 1992 Namibia DHS, the primary school NAR in Namibia is cited as 74% for males and 79% for females—which correspond to the ratios of school attendance reported for children age 6-10 (*see Table 3*). In the UNICEF report, these estimates are footnoted as differing from the standard definition of NAR. Based on data from the 1992 Namibia DHS, the actual primary school NAR estimates (for children age 7-13) are 88% for males (not 74%) and 91% for females (rather than 79%). Had the DHS reports presented NAR data in addition to attendance ratios and for the appropriate age ranges, the estimates of NAR presented in the UNICEF report would have been more accurate.

5.3 PRESENTING DATA GRAPHICALLY

One way of emphasising a particular finding is to present information visually in a graph or figure. Figures are particularly useful for disseminating the data to the informed public, but are just as useful for researchers and decision makers. Figures distill and synthesise the most pertinent and interesting findings. They emphasise the relationship between indicators. Therefore, it is important for the choices to be made with care and for the figures to be legible and easy to interpret.

Pie charts are among the simplest and most easily understood figures. This type of figure requires categories to equal 100%. Figure 2 presents literacy data for men and women in Guinea based on the results of the 1999 DHS. Literacy data were collected for different age ranges for men and women.

Figure 2
Literacy rate among men age 15-59 and women age 15-49, Guinea, 1999 DHS

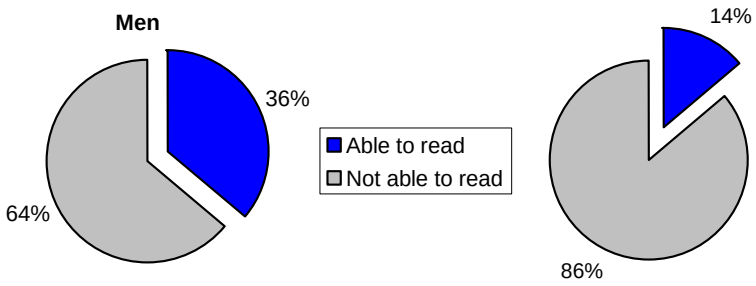
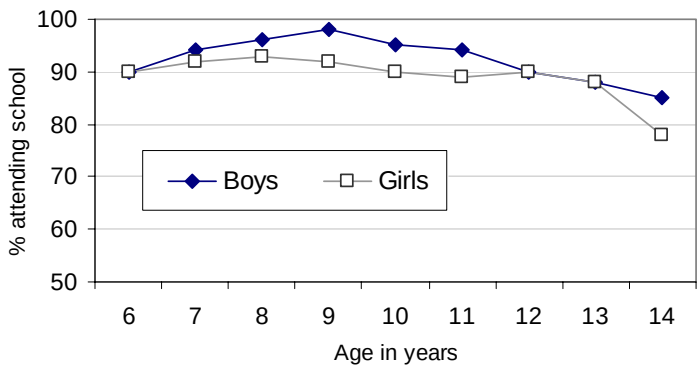


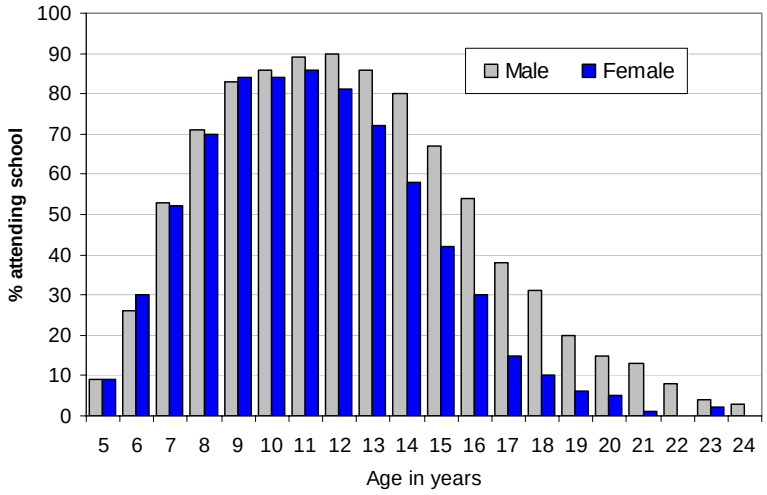
Figure 3 presents information from the MICS on the attendance ratios of boys and girls in rural Algeria. Among children in rural areas, there is a noticeable gender difference in attendance ratios between the ages of 9 and 12. Attendance ratios decline for both boys and girls from age 12 to 14.

Figure 3
Attendance ratio among boys and girls in rural areas, Algeria, 2000 MICS



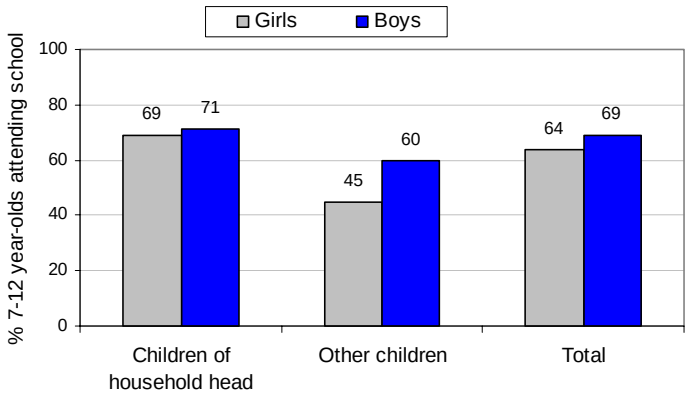
Another way to present data graphically is to create a bar graph that shows the percentage of children of each age attending school. Figure 4 presents the Age-Specific Attendance Rates (ASAR) for youth in Cambodia. There is gender parity from age 7 to 9, but from age 10 onward, far more male youths attend school than female youths, with the gap widening among older youth.

Figure 4
Age-specific attendance rates by sex, Cambodia, 2000 DHS



Based on the 1996 census in Burkina Faso, Figure 5 presents enrollment ratios for children age 7-12 by sex and according to their relationship to the head of the household. Male and female children of the household head are almost equally likely to be enrolled in school; in contrast, children who are not the children of the household head are considerably less likely to be enrolled in school, with girls far less likely than boys to be enrolled in school. There is evidence that many of the children in Ouagadougou who are not children of the household head work in households in domestic service; as may be expected, children in domestic service may be unlikely to attend school because of work or other reasons.

Figure 5
Percentage of children age 7-12 in Ouagadougou who are enrolled in school, by sex and relationship to household head, Burkina Faso, 1996 Census



Based on data from the 1996 Zambia DHS, Figure 6 presents data on the proportion of students, by grade, who are over-age, under-age, and on-time, by sex. The bars illustrate graphically the increasing proportion of over-age students by primary school grade. This type of bar chart can also illustrate adults' educational attainment, using categories such as no schooling, primary, secondary, and higher.

Figure 6
Proportion of primary school students under-age, over-age, and on-time by grade and sex, Zambia, 1996 DHS

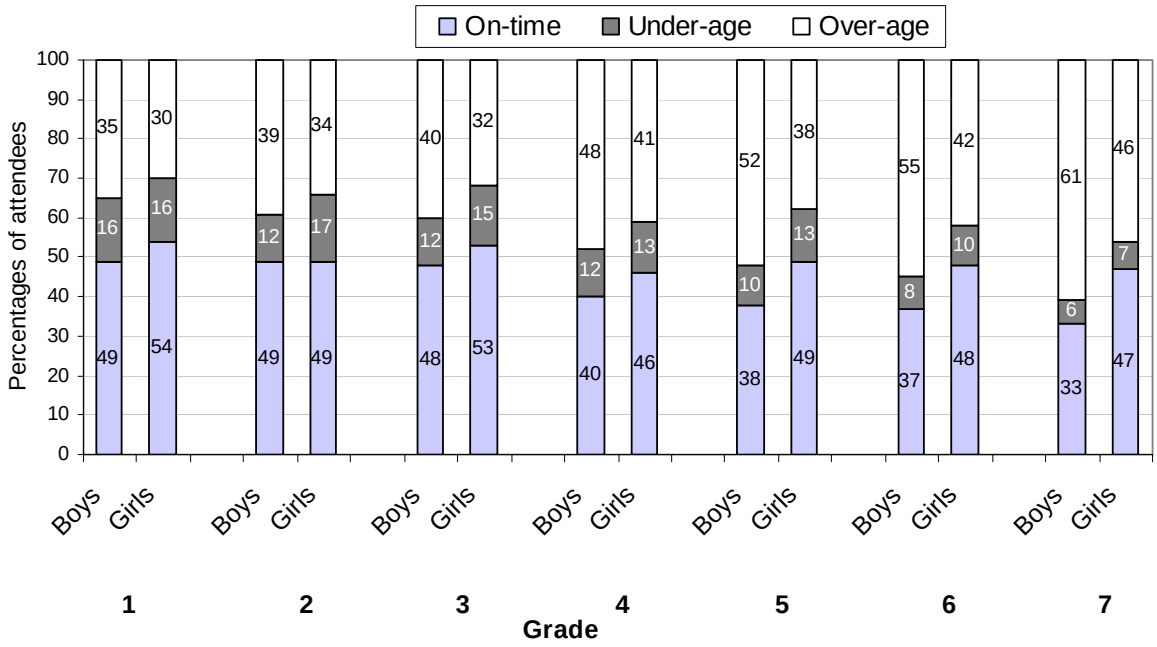


Figure 7 illustrates yet another chart format for data presentation. This chart uses data from the 1996 census in Burkina Faso to illustrate the distribution of students age 6 to 29 across schooling levels. The target age range for primary school is age 7 to 12. The chart shows that most children in that age range attending school are in primary. Generally speaking, youth of secondary school age (13 to 19) attend secondary school. However, many youths up to the age of 15 attend primary school, suggesting that a considerable number of students are over-age for the level of schooling attended.

Figure 7
Distribution of students by age and education, Burkina Faso, 1996 Census

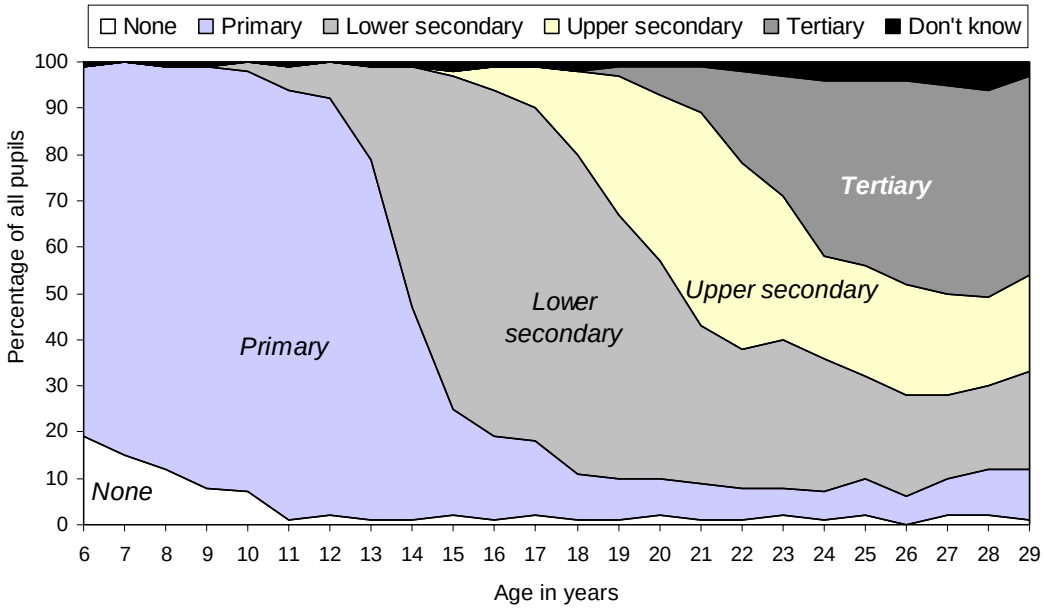
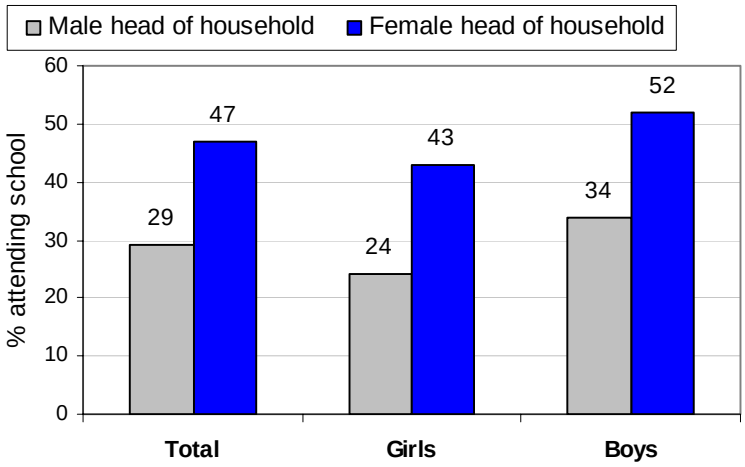


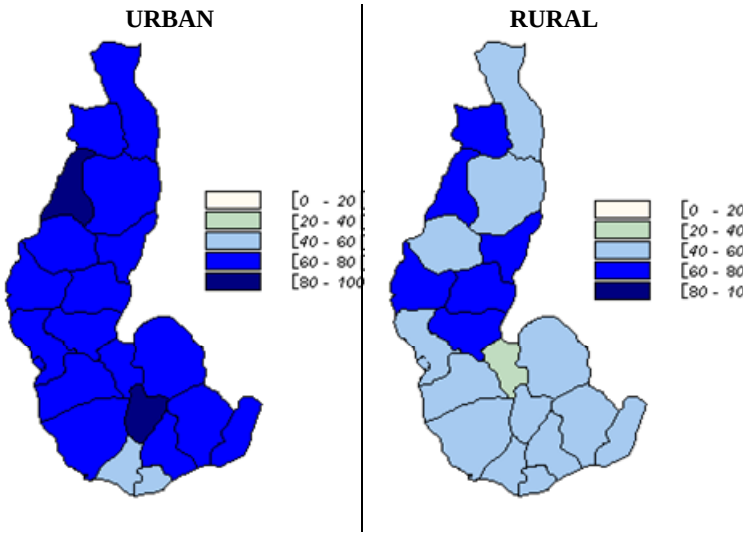
Figure 8 presents findings on gender in Burkina Faso. Households headed by women are much more likely to send both boys and girls to school than households headed by men. However, generally more boys attend school than girls regardless of the sex of the household head.

Figure 8
Percentage of boys and girls who attend school, according to the sex of the head of the household, Burkina Faso, 1994-95 EP



Maps are also a helpful way of presenting regional variation in educational participation. Figure 9 highlights differences in educational participation in Madagascar. In urban areas, educational participation of all children 6-14 in the household is extremely high except in the southernmost region. Generally, there is lower educational participation of all children 6-14 in the household in rural areas, with a few exceptions in the central and northwestern districts.

Figure 9
Percentage of households with children 6-14 years old attending school by location, Madagascar, 1993 Census



The above tables and figures illustrate considerations to bear in mind when presenting data. As mentioned previously, these tables and figures are by no means exhaustive. Tables and figures presenting data on other education indicators can be patterned after those presented in this Guide.

REFERENCES

Bottani, N. (1990). "The background of the CERI/OECD project on international educational indicators," *International Journal of Educational Research,* 14, 335-342.

Central Statistical Office [Zambia], Ministry of Health [Zambia] and Macro International. (1993). *Zambia Demographic and Health Survey, 1992*. Calverton, Maryland: Central Statistical Office and Macro International Inc.

Central Statistical Office [Zambia], Ministry of Health [Zambia] and Macro International. (1997). *Zambia Demographic and Health Survey, 1996.* Calverton, Maryland: Central Statistical Office and Macro International Inc.

Cochran, WG. (1977). *Sampling Techniques*. John Wiley & Sons.

Grosh, Margaret and Paul Glewwe. (2000). *Designing Household Survey Questionnaires for Developing Countries: Lessons from Fifteen Years of the Living Standards Measurement Study.* Washington, DC: World Bank and Oxford University Press.

Grosh, Margaret and Paul Glewwe. (1995). "A Guide to Living Standards Measurement Study Surveys and Their Data Sets," *LSMS Working Paper 120.* Washington, DC: World Bank.

Grosh, Margaret and Juan Muñoz. (1996). "A Manual for Planning and Implementing the Living Standards Measurement Study Survey," *LSMS Working Paper 126*. Washington, DC: World Bank.

Groves R.M., D.A. Dillman, J.L. Eltinge and R.J.A. Little (2001). *Survey Nonresponse*. Jossey-Bass.

Ministère de l'Enseignement Pré-Universitaire et de l'Éducation Civique, Service Statistique et Planification [Guinea] and ORC Macro. (2001*). Schooling in Guinea: Findings from the GDHS-II 1999.* Calverton, Maryland USA: Ministère de l'Enseignement Pré-Universitaire et de l'Éducation Civique, Service Statistique et Planification and ORC Macro.

Ministry of Health and Social Services[Namibia] and Macro International Inc. (1993*). Namibia Demographic and Health Survey 1992.* Windhoek, Namibia and Calverton, Maryland, USA: Ministry of Health and Social Services and Macro International Inc.

Pilon, Marc. (Ed.). (1999). Guide d'exploitation et d'analyse des données de recensement et d'enquêtes en matière de scolarisation (Les Documents et Manuels du CEPED, n°9). Paris, France: Centre français sur la population et le développement(CEPED), Union pour l'étude de la population africaine(UEPA), et

Organisation des Nations unies pour l'éducation, la science et la culture(UNESCO).

Särndal C-E, Swensson B. and Wretman J. (1992). *Model Assisted Survey Sampling*. Springer-Verlag.

Statistical Service, Republic of Ghana and World Bank. (1988*). Ghana Living Standards Survey Preliminary Report.* Lkd. LSMS. <http://www.worldbank.org/html/prdph/ lsms/country/gh/docs/GHA04.pdf.> (2001, June 26).

UNICEF (1999). *The State of the World's Children*. New York, NY: UNICEF.

UNESCO (1988). *World Education Indicators, Working Document.* Paris, France: UNESCO.

UNESCO (1998a). Division of Statistics*. Gender-Sensitive Education Statistics and Indicators: A Practical Guide.* Lkd. UNESCO. <http://unescostat.unesco.org/en/ pub/doc/gend-stat.pdf> (2000, June 19).

UNESCO (1998b). *Use of Cohort Analysis Models for Assessing Educational Internal Efficiency,* UNESCO. <http://www.uis.unesco.org/statsen/statistics/indicators/ i_pages/indspec/eff-box.html >.

UNESCO (2000). *Education for All: The Year 2000 Assessment Technical Guidelines published by UNESCO for the International Consultative Forum on Education for All* (*see website www.education.unesco.org/efa*).

United Nations. Statistical Office (1998*). Principles and Recommendations for Population and Housing Censuses*. New York. Series M No. 67, Rev. 1.

Wolter, K. (1985). *Introduction to Variance Estimation*. Springer-Verlag.

CALCULATION OF EDUCATION INDICATORS

APPENDIX A

As discussed in Chapter 4, missing values can be handled in different ways. Household surveys like the DHS do not exclude cases with missing values. In this way, tables with consistent sample sizes across indicators can be generated. Other data collection efforts choose to exclude cases with missing values on key variables. Decisions on how to handle missing values need to be made on a case-by-case basis.

Literacy

The adult literacy ratio is calculated as follows:

$$\frac{\text{Number of adults age 15 and older in the sample or population who are literate}^{15}}{\text{Number of adults age 15 and older in the sample or population}}$$

As mentioned in the text, the adult literacy ratio can also be calculated for adults age 25 and older, and for various age groups (such as 15-24, 25-34, 35-44…).

Furthermore, the definition of literacy employed in the national census or household survey is critical. In household surveys such as the DHS, the percentage of literate persons includes respondents who can read a simple statement easily or with difficulty.

Adult educational attainment

Adult educational attainment is a measure of the highest level of schooling attended, and is calculated by level of schooling, as follows:

Percentage with no schooling:

$$\frac{\text{Number of adults age 15 and older who have attended less than one year of school}}{\text{Number of adults age 15 and older in the sample or population}}$$

Percentage with primary schooling:

$$\frac{\text{Number of adults age 15 and older whose highest level attended is primary school}}{\text{Number of adults age 15 and older in the sample or population}}$$

Percentage with secondary schooling:

$$\frac{\text{Number of adults age 15 and older whose highest level attended is secondary school}}{\text{Number of adults age 15 and older in the sample or population}}$$

Percentage with post-secondary schooling:

$$\frac{\text{Number of adults age 15 and older whose highest level attended is post-secondary school}}{\text{Number of adults age 15 and older in the sample or population}}$$

Notice that these categories are mutually exclusive and together should total 100%.

[15] Literacy may be self-reported or tested, according to the survey.

Adult school completion ratios

Adult primary school completion ratios are calculated as follows:

$$\frac{\text{Number of adults age 15 and older who completed primary school (or have attended a higher grade)}}{\text{Number of adults age 15 and older in the sample or population}}$$

Adult secondary school completion ratios are calculated as follows:

$$\frac{\text{Number of adults age 20 and older who completed secondary school (or have attended a higher grade)}}{\text{Number of adults age 20 and older in the sample or population}}$$

Notice that in these calculations, the age ranges differ for primary (age 15 and older) and secondary (age 20 and older) school completion ratios among adults. The age range begins at 20 for the indicator at the secondary level because the official age range for secondary school extends through the late teenage years in most countries. Owing to the way these indicators are calculated, adults who have completed secondary school or higher are counted as having completed both primary and secondary school.

Current school attendance ratios

The *net attendance ratio* (NAR) at primary school level is the percentage of the official primary school age population attending primary school. This indicator is calculated as follows:

$$\frac{\text{Number of children of primary school age who attend primary school}}{\text{Number of children of primary school age in the sample or population}}$$

At the secondary level, the indicator is calculated as follows:

$$\frac{\text{Number of children of secondary school age who attend secondary school}}{\text{Number of children of secondary school age in the sample or population}}$$

Note in particular that the numerator for the secondary school NAR includes only those children of secondary school age who attend secondary school; those children who attend primary school are included in the denominator but not in the numerator. Also note that the official primary and secondary school age ranges differ across countries.

The gross attendance ratio (GAR) at the primary level is the total number of students attending primary school—regardless of age—expressed as a percentage of the official primary school age population. The GAR indicates the level of participation in primary schooling by people of any age, and when compared with the NAR, indicates the extent of over- and under-age participation in primary schooling. In countries with high ratios of primary school attendance, the GAR can exceed 100 if there are significant numbers of over-age (or under-age) students in primary school. The GAR at the primary level is calculated as follows:

$$\frac{\text{Number of children of } \textit{any} \text{ age who attend primary school}}{\text{Number of children of primary school age in the sample or population}}$$

At the secondary level, the indicator is calculated as follows:

$$\frac{\text{Number of children of } \textit{any} \text{ age who attend secondary school}}{\text{Number of children of secondary school age in the sample or population}}$$

The gender parity index (GPI) at the primary level indicates whether male and female children are equally likely to attend primary school, or whether there is a gender gap in participation. The GPI at the primary level is calculated as follows:

$$\frac{\text{Female primary school GAR}}{\text{Male primary school GAR}}$$

At the secondary level, the GPI is calculated as follows:

$$\frac{\text{Female secondary school GAR}}{\text{Male secondary school GAR}}$$

The net intake ratio is the number of new entrants to the first year of primary school who are of the official primary school entrance age, expressed as a percentage of the population that is of the official school entrance age. Indicating the extent of timely entry to primary school, the net intake ratio is a more sensitive measure of changes in school enrollment ratios than net enrollment or attendance ratios.

The net intake ratio requires data on school attendance over a two-year period, and is estimated as follows, assuming that the official starting age for the first grade of primary school is age 7:

$$\frac{\text{Number of 7-year-old children who did not attend the first grade (or any higher grade) of primary school last year, but did attend the first grade of primary school this year}}{\text{Number of children who are age 7}}$$

The gross intake ratio (or apparent intake ratio) is the number of new entrants to the first year of primary school—regardless of age—expressed as a percentage of the population that is of the official school entrance age. The gross intake ratio indicates the extent of access to the first grade of primary school, relative to the population of primary school entrance age. If participation in primary schooling is high and there are many over-age and/or under-age new entrants in the school system, the gross intake ratio can exceed 100 percent.

The gross intake ratio also requires data on school attendance over a two-year period. Assuming that the official starting age for the first grade of primary school is age 7, it is calculated as follows:

$$\frac{\text{Number of children of any age who did not attend the first grade (or any higher grade) of primary school last year, but did attend the first grade of primary school this year}}{\text{Number of children who are age 7}}$$

The proportion of students over-age, under-age, and on-time, by grade is calculated as detailed below.

Proportion of grade 1 students who are on-time for grade:

$$\frac{\text{Number of grade 1 students who are at the official age for the grade (age x) or age (x + 1)}}{\text{Number of grade 1 students in the sample or population}}$$

Proportion of grade 1 students who are under-age for grade:

$$\frac{\text{Number of grade 1 students who are younger than the official age for the grade (age x)}}{\text{Number of grade 1 students in the sample or population}}$$

Proportion of grade 1 students who are over-age for grade:

$$\frac{\text{Number of grade 1 students who are older than the official age for the grade + 1 (x + 1)}}{\text{Number of grade 1 students in the sample or population}}$$

The age-specific attendance ratio (ASAR) is calculated as follows, for all ages for which data are available. If, for example, data on current school attendance, at any level of schooling, is available for children age 6-24, the ASAR can be calculated for each age in that range.

$$\frac{\text{Number of children age 6 currently attending any level of schooling}}{\text{Number of children age 6 in the sample or population}}$$

$$\frac{\text{Number of children age 7 currently attending any level of schooling}}{\text{Number of children age 7 in the sample or population}}$$

And so on, through age 24.

Student flow ratios

Student flow ratios describe the movement of students in and out as well as through the education system. Using data on school attendance and level and grade attended during the current and last years that are collected by recent demographic surveys, the following student flow ratios can be derived:

- Promotion ratio
- Repetition ratio
- Dropout ratio
- Transition ratio
- Survival ratio

These indicators can be calculated from national census or household survey data, providing that the following basic questions are included in the survey:

Q1. Is (name) currently attending school ?
Q2. If 'YES', what level and grade is (name) currently attending ?
Q3. Did (name) attend school last year ?
Q4. If 'YES', what level and grade did (name) attend last year ?

These basic student flow indicators can in turn be used in calculating survival ratios as well as other indicators by applying the cohort student flow model.

For household members whose answers to both Q1 and Q3 are 'YES', the comparison of the answers to Q2 and Q4 allows the *promotion ratio* to be calculated: those who were in grade (n-1) last year, and who are in grade (n) this year were promoted. This

number is then divided by the number of students who were attending grade (n-1) the last year to obtain the promotion ratio for grade (n-1) for last year.

The repetition ratio, by grade, is the proportion of students who attend a given grade in a given school year who attend the same grade in the following year. The repetition ratio is calculated as follows:

$$\frac{\text{Number of students attending a given grade in a given school year who attended the same grade in the last school year)}}{\text{Number of students attending the given grade in the last school year}}$$

In other words, the repetition ratio for grade (n) is calculated by dividing the number of students who attended grade (n) during both the current and last school years (*see Q. 2 and Q.4*), by the total number of students who attended grade (n) in the last school year.

The transition ratio from primary to secondary school is calculated as follows:

$$\frac{\text{Number of students in last grade of primary school in a given school year who attend the first grade of secondary school in the following school year}}{\text{Number of students in last grade of primary school in a given school year}}$$

In other words, the transition ratio from primary to secondary school is calculated by dividing the number of students in the last grade of primary school in the last school year who attend the first grade of secondary in the current school year, by the number of students who attended the last grade of primary school in the last school year.

Some students may skip grades from one year to the next (for example, from grade 1 to grade 3), or descended from a higher to a lower grade (for example, from grade 3 to Ggade 2). If such occurrences are common, grade-skipping and demotion ratios can be calculated.

The dropout ratio by grade is calculated in a similar way as is the repetition ratio:

$$\frac{\text{Number of students who attended a given grade in a given school year who do not attend school in the following school year}}{\text{Number of students attending the given grade in the last school year}}$$

In other words, the dropout ratio for grade (n) is calculated by dividing the number of students who attended grade (n) during the last school year (*see Q. 4*) and who did not attend school at all during the current school year (Q.1 is NO), by the total number of students who attended grade (n) in the last school year.

SURVEY AND CENSUS CONTACT INFORMATION

APPENDIX B

Access to data sets varies by survey and may be under the purview of the specific country. The survey programme website is the first place to start to learn about what data may be available from countries of interest and how to go about accessing the data required.

Demographic and Health Surveys

Measure DHS +
Macro International Inc.
11785 Beltsville Drive, Suite 300
Calverton, MD 20705
USA

Phone: 301 572-0456
Fax: 301 572-0999
E-mail: measure@orcmacro.com
http://www.measuredhs.com

Access to data sets: Data sets are publicly available through Macro International. Data sets can be requested via the website. Researchers need to apply for a download account. There is no charge for downloading the data sets and accompanying documentation. The site also enables customized tables to be built from hundreds of DHS surveys and indicators with the STATcompiler.

Access to DHS education data: DHS EdData education data can be accessed from the DHS EdData Education Profiles; USAID Global Bureau of Human Capacity Development's Global Education Database; DHS's STATcompiler; and DHS's web data set distribution. (See the USAID's DHS EdData Activity Website for additional information and links: http://www.dhseddata.com.)

Living Standards Measurement Study

LSMS Data Manager
DECRG
The World Bank
1818 H Street, NW
MSN MC3-306
Washington, DC 20433
USA

Fax: (202) 522-1153
E-mail: lsms@worldbank.org
Website: http://www.worldbank.org/lsms/

Access to data: The Development Economics Research Group (DECRG) of the World Bank, formerly the Policy Research Department, maintains the website to provide researchers worldwide with the data sets and methodological lessons from these surveys. It is possible to access results from 38 surveys across 25 countries conducted from 1985 to 2000. Documentation and basic information for all of the LSMS surveys are available through the LSMS Office or through the LSMS website. Many of the data sets can also be downloaded free of charge from the website. http://www.worldbank.org/lsms/

Information on data available in Africa can be found on the Africa Poverty Monitoring Website: http://www4.worldbank.org/afr/poverty/default.cfm.

Multiple Indicator Cluster Surveys

UNICEF, MICS Coordinator
3 UN Plaza
New York, NY 10017

Phone: 212 824-6745
E-mail: childinfo@unicef.org
Website: http://www.childinfo.org

Access to data sets: Access to MICS survey data varies by country as the data are verified at the national level. Contact the MICS Coordinator or the statistics office in country to obtain data sets.

REMUAO Surveys

CERPOD (Centre d'Etudes et de Recherche sur la Population pour le Développement, Bamako, Mali) provided international coordination through REMUAO and published the results of the surveys. To gain access to the data sets or learn more about the REMUAO Surveys, contact CERPOD or the statistical office for the country of interest.

Sadio TRAORE
CERPOD- Centre d'Etudes et de Recherche sur la Population pour le Développement,
BP 1530, Bamako, Mali
Tel: (223) 22 30 43 / 22 80 86
Fax: (223) 22 78 31
E-mail: btraore@cerpod.insah.ml

Africa Household Survey Databank

The Africa Household Survey Databank contains data sets for 110 surveys conducted as early as 1985 and as recently as 1998 for 37 countries. These surveys include both international or donor-driven surveys and national surveys. The types and number of surveys available are: Priority Surveys (Social Dimension of Adjustment - SDA): 28;

Household Budget Surveys / Income Expenditure Surveys: 15; Integrated Surveys (SDA/LSMS): 21; Core Welfare Indicators Questionnaire Survey (CWIQ): 1; Demographic and Health Survey (DHS): 40; Demographic and Health Survey (non-DHS): 1; Others: 4. For further information on survey types, coverage, sampling units, status of data and access policies, see the website.

http://www4.worldbank.org/afr/poverty/databank/about.cfm

Household Surveys Initiative (ISLC/MECOVI)

The Program for the Improvement of Surveys and the Measurement of Living Conditions in Latin America and the Caribbean (ISLC/MECOVI) covers household survey data sets for seven countries: Argentina, Bolivia, El Salvador, Guatemala, Nicaragua, Paraguay and Peru. The objective is to develop an organized, documented and standardized data bank composed of household surveys executed within the ISLC/MECOVI Programme and household surveys executed independently from this programme. Web access to data sets is currently u*nder construction.*

http://www.iadb.org/sds/pov/site_19_e.htm

Poverty Monitoring Database

This database provides quick access to comprehensive poverty information, including household surveys, poverty assessment summaries, participatory poverty assessments, and social indicators.

It provides information on key features of surveys, data availability and general information on recent income/consumption surveys, including many by national agencies.

http://wbln0018.worldbank.org/dg/povertys.nsf/Surveys+By+Country?openview&Count=1999

Luxembourg Income Study/Luxembourg Employment Study

Started in 1983, the Luxembourg Income Study (LIS) is an international comparative study on income distribution aimed at setting up a database containing social and economic micro-data collected in household surveys from different countries and to promote comparative research on the economic status of populations in different countries. The LIS database has a total of over 60 income and expenditure survey data sets for 26 countries, covering the period from the 1980s to 1992.

http://www.lisproject.org/techdoc.htm

Started in 1993, the Luxembourg Employment Study (LES) is designed to facilitate the study of different labour market related issues, including analysis of labour market behaviour on an individual level, or in the framework of the household; educational and

occupational patterns; and retirement decisions. The LES database includes labour force surveys for 16 OECD countries from the late 1980s.

http://www.lisproject.org/lestechdoc.htm

HEIDE database

The Household Expenditure and Income Data for Transitional Economies (HEIDE Database) includes household expenditure and income data for transitional economies including: Bulgaria, Hungary, Poland, Slovak Republic, Armenia, Estonia, Latvia, Russia and Kyrgyz Republic.

http://econ.worldbank.org/view.php?id=3444

Rand Family Life Surveys

The Rand Family Life Surveys were household and community surveys conducted by RAND in collaboration with local institutions in Bangladesh, Guatemala, Indonesia and Malaysia.

http://www.rand.org/

National Censuses

For further information on data available from National Censuses and National Household Surveys conducted by local agencies, contact the statistical office in country.

Organizations with links to education data sources

The William Davidson Institute
http://www.wdi.bus.umich.edu/

The Development Gateway
http://www.ids.ac.uk/eldis/health/health.htm

University of California site
http://biko.sscnet.ucla.edu/dev_data/

SURVEYS BY TYPE AND COUNTRY

APPENDIX C

DHS surveys by region

Country	Year	Fieldwork		Type	Status	Implementing Organization	Female			Households Sample
		Start	End				Resp.	Age	Sample	
Sub-Saharan Africa										
Benin	2001	Aug-01	Oct-01	DHS	Completed	INSAE	All Women	15-49	6,219	5,796
Benin	1996	Jun-96	Aug-96	DHS	Completed	Inst. Nat. de la Statistique	All Women	15-49	5,491	4,499
Botswana	1988	Aug-88	Dec-88	DHS	Completed	Ministry of Health	All Women	15-49	4,368	4,473
Burkina Faso	1999	Dec-98	Mar-99	DHS	Completed	Inst. Nat. de la Stat. et de la Démo.	All Women	15-49	6,445	4,812
Burkina Faso	1992	Nov-92	Jan-93	DHS	Completed	Inst. Nat. de la Stat. et de la Démo	All Women	15-49	6,354	5,143
Burundi	1987	Apr-87	Jul-87	DHS	Completed	Dép. de la Pop., Min. de l'Intérieur	All Women	15-49	3,970	3,868
Cameroon	1998	Feb-98	Jun-98	DHS	Completed	Bur. Cen. Recensements et Etudes de Prop.	All Women	15-49	5,501	4,697
Cameroon	1991	Apr-91	Sep-91	DHS	Completed	Min. du Plan et de l'Amén. du Terr.	All Women	15-49	3,871	3,538
CAR	1994	Sep-94	Mar-95	DHS	Completed	Dir. des Stat. Dém. et Sociales	All Women	15-49	5,884	5,551
Chad	1997	Dec-96	Jul-97	DHS	Completed	Bureau Central du Recensement	All Women	15-49	7,454	6,840
Comoros	1996	Mar-96	May-96	DHS	Completed	Centre Nat. de Doc. et de Rech. Sci.	All Women	15-49	3,050	2,252
Côte d'Ivoire	1998	Sep-98	Mar-99	DHS	Completed	Inst. Nat. de la Statistique	All Women	15-49	3,040	2,122
Côte d'Ivoire	1994	Jun-94	Nov-94	DHS	Completed	Inst. Nat. de la Statistique	All Women	15-49	8,099	5,935
Eritrea	2002	Mar-02	Jul-02	DHS	Completed	National Statistics and Evaluation Office	All Women	15-49	8,754	9,389
Eritrea	1995	Sep-95	Jan-96	DHS	Completed	Nat. Statistics Office	All Women	15-49	5,054	5,469
Ethiopia	2000	Feb-00	Apr-01	DHS	Completed	Central Statistical Authority	All Women	15-49	15,367	14,072
Gabon	2000	Oct-00	Dec-00	DHS	Completed	Direction Générale de la Stat. Et des Etudes Economiques	All Women	15-49	6,183	6,203
Ghana	2003	Aug-03	Oct-03	DHS	Completed	Ghana Statistical Service	All Women	15-49	4,500	6,500
Ghana	1998	Nov-98	Feb-99	DHS	Completed	Ghana Statistical Service	All Women	15-49	4,843	6,003
Ghana	1993	Sep-93	Jan-94	DHS	Completed	Ghana Statistical Service	All Women	15-49	4,562	5,822
Ghana	1988	Feb-88	May-88	DHS	Completed	Ghana Statistical Service	All Women	15-49	4,488	4,406
Guinea	1999	May-99	Jun-99	DHS	Completed	Direction Nat. de la Statistique et de l'Information	All Women	15-49	6,753	5,090
Guinea	1992	Feb-92	Mar-92	Special	Completed	Direction Nat. de la Statistique	All Women	15-49	6,065	6,899

Country	Year	Fieldwork		Type	Status	Implementing Organization	Female			Households Sample
		Start	End				Resp.	Age	Sample	
Sub-Saharan Africa										
Kenya	2003	Apr-03	Aug-03	DHS	Completed	Central Bureau of Statistics (CBS)	All Women	15-49	8,195	8,561
Kenya	1999	Apr-99	Aug-99	SPA	Completed	Nat. Council for Pop. and Dev./Min. of Health	Not specified		Not specified	Not specified
Kenya	1998	Feb-98	Jul-98	DHS	Completed	Nat. Council for Pop. and Dev.	All Women	15-49	7,881	8,380
Kenya	1993	Feb-93	Aug-93	DHS	Completed	Nat. Council for Pop. and Dev.	All Women	15-49	7,540	7,950
Kenya	1989	Dec-88	May-89	DHS	Completed	Nat. Council for Pop. and Dev.	All Women	15-49	7,150	8,173
Lesotho	2004	Nov-03	Jul-04	DHS	Ongoing	Ministry of Health and Social Welfare and the Bureau of Statistics	All Women	15-49	8,200	9,400
Liberia	1986	Feb-86	Jul-86	DHS	Completed	Min. of Planning & Economic Affairs	All Women	15-49	5,239	5,023
Madagascar	2004	Nov-03	Mar-04	DHS	Ongoing	Direction de la Dém. et des Stat. Sociales (DDSS)	All Women	15-49	9,000	9,000
Madagascar	1997	Sep-97	Dec-97	DHS	Completed	Dir. de la Dém. et des Stat. Sociales/INSTAT	All Women	15-49	7,060	7,171
Madagascar	1992	May-92	Nov-92	DHS	Completed	Centre Nat. de Recherches sur l'Env.	All Women	15-49	6,260	5,944
Malawi	2002	May-02	Jul-02	EdData	Completed	Nat. Statistical Office	Not specified		Not specified	3,290
Malawi	2000	Jul-00	Nov-00	DHS	Completed	Nat. Statistical Office	All Women	15-49	13,220	14,213
Malawi	1996	Jun-96	Oct-96	KAP	Completed	Nat. Statistical Office	All Women	15-49	2,683	2,798
Malawi	1992	Sep-92	Nov-92	DHS	Completed	Nat. Statistical Office	All Women	15-49	4,850	5,323
Mali	2001	Jan-01	May-01	DHS	Completed	CPS/MSSPA et DNSI	All Women	15-49	12,817	12,285
Mali	1996	Nov-95	Apr-96	DHS	Completed	CPS/MSSPA et DNSI	All Women	15-49	9,704	8,716
Mali	1987	Mar-87	Aug-87	DHS	Completed	Instit de Sahel: USED/CERPOD	All Women	15-49	3,200	3,048
Mauritania	2003	Jun-03	Aug-03	Special	Ongoing	Office Nat. de la Statistique	All Women	15-49	6,000	Not specified
Mauritania	2001	Oct-00	Apr-01	Other	Completed	Office Nat. de la Statistique	All Women	15-49	7,728	6,149
Mozambique	2003	Jul-03	Sep-04	DHS	Ongoing	INE (National Statistical Institute)	All Women	15-49	12,193	12,087
Mozambique	1997	Mar-97	Jun-97	DHS	Completed	Inst. Nacional de Estadistica	All Women	15-49	8,779	9,282
Namibia	2000	Sep-00	Dec-00	DHS	Completed	Ministry of Health and Social Services	All Women	15-49	6,755	6,392
Namibia	1992	Jul-92	Nov-92	DHS	Completed	Min. of Health and Social Services	All Women	15-49	5,421	4,101
Niger	2000	Apr-00	Aug-00	Special	Completed	Macro International Inc.	All Women	15-49	5,664	4,321
Niger	1998	Mar-98	Jul-98	DHS	Completed	Care International	All Women	15-49	7,577	5,928
Niger	1992	Mar-92	Jun-92	DHS	Completed	Dir. de la Stat. et des Comptes Nationaux	All Women	15-49	6,503	5,242

Country	Year	Fieldwork		Type	Status	Implementing Organization	Female			Households Sample
		Start	End				Resp.	Age	Sample	
Sub-Saharan Africa										
Nigeria	2004	Feb-04	Jun-04	EdData	Ongoing	National Population Commission	Not specified		Not specified	4,600
Nigeria	2003	Mar-03	Aug-03	DHS	Completed	The National Population Commission	All Women	15-49	7,620	7,225
Nigeria	1999	Mar-99	May-99	DHS	Completed	Nat. Pop. Comm.	All Women	10-49	7,647	7,647
Nigeria	1990	Apr-90	Oct-90	DHS	Completed	Federal Office of Statistics	All Women	15-49	8,781	8,999
Ondo State	1986	Sep-86	Jan-87	DHS	Completed	Min. of Health	All Women	15-49	4,213	3,437
Rwanda	2004	Mar-04	Dec-04	DHS	Ongoing					
Rwanda	2001	Sep-01	Nov-01	SPA	Completed	ONAPO - Office National de la Population	Not specified		Not specified	Not specified
Rwanda	2000	Jun-00	Aug-00	DHS	Completed	Office National de la Popluation	All Women	15-49	10,421	9,696
Rwanda	1992	Jun-92	Oct-92	DHS	Completed	Office National de la Popluation	All Women	15-49	6,551	6,252
Senegal	2004	Nov-04	Apr-05	DHS	Ongoing	Centre de Recherche et de Développement Humain (CRDH)	All Women	15-49	12,000	7,950
Senegal	1999	Oct-99	Dec-99	DHS	Completed	SERDHA	All Women	15-49	17,189	9,085
Senegal	1997	Jan-97	Apr-97	DHS	Completed	Min. de l'Economie et des Finances	All Women	15-49	8,593	4,772
Senegal	1993	Nov-92	Aug-93	DHS	Completed	Dir. de la Prévision et de la Stat.	All Women	15-49	6,310	3,528
Senegal	1986	Apr-86	Jul-86	DHS	Completed	Min. de l'Economie et des Finances	All Women	15-49	4,415	3,736
South Africa	2004	Nov-03	Dec-04	DHS	Ongoing	Dept. of Health	All Women	15-49	12,000	10,000
South Africa	1998	Feb-98	Sep-98	DHS	Ongoing	Dept. of Health/Med. Research Council	All Women	15-49	11,735	12,247
Sudan	1990	Nov-89	May-90	DHS	Completed	Dept. of Stat., Min. of Econ. & Nat. Plan.	Ever Married Women	15-49	5,860	6,891
Tanzania	2004	Sep-04	Jul-05	DHS	Ongoing					
Tanzania	2003	Sep-03	Jan-04	AIS	Ongoing	National Bureau of Statistics	All Women	15-49	6,000	7,000
Tanzania	1999	Sep-99	Nov-99	Interim	Completed	Nat. Bureau of Statistics	All Women	15-49	4,029	3,615
Tanzania	1996	Jul-96	Nov-96	DHS	Completed	Bureau of Statistics, Planning Comm.	All Women	15-49	8,120	7,969
Tanzania	1995	Jun-95	Oct-95	In Depth	Completed	Bureau of Statistics, Planning Comm.	All Women	15-49	2,130	1,488
Tanzania	1994	Jul-94	Sep-94	KAP	Completed	Bureau of Statistics, Planning Comm.	All Women	15-49	4,225	4,023
Tanzania	1992	Oct-91	Mar-92	DHS	Completed	Bureau of Statistics, Planning Comm.	All Women	15-49	9,238	8,327
Togo	1998	Feb-98	May-98	DHS	Completed	Direction de la Statistique	All Women	15-49	8,569	7,517
Togo	1988	Jun-88	Nov-88	DHS	Completed	Unité de Recherche Dém., U. du Benin	All Women	15-49	3,360	3,432
Uganda	2004	Feb-04	May-04	AIS	Ongoing	Ministry of Health	All Women	15-59	9,800	10,425
Uganda	2001	Apr-01	Jul-01	EdData	Completed	UBOS - Uganda Bureau of Statistics	Not specified		Not specified	Not specified

Country	Year	Fieldwork		Type	Status	Implementing Organization	Female			Households Sample
		Start	End				Resp.	Age	Sample	
Sub-Saharan Africa										
Uganda	2000	Sep-00	Feb-01	DHS	Completed	Uganda Bureau of Stat. (formerly Dept. of Stat.)	All Women	15-49	7,246	7,885
Uganda	1995	Mar-95	Aug-95	DHS	Completed	Dept. of Stat., Min. of Fin. & Econ. Plan.	All Women	15-49	7,070	7,550
Uganda	1995	Oct-95	Jan-96	In Depth	Completed	Inst. Stat. & Applied Econ. Makerere U.	All Women	20-44	1,750	3,610
Uganda	1988	Sep-88	Feb-89	DHS	Completed	Ministry of Health	All Women	15-49	4,730	5,101
Zambia	2002	Aug-02	Oct-02	EdData	Completed	CSO - Central Statistical Office	Not specified		Not specified	4,245
Zambia	2001	Nov-01	May-02	DHS	Completed	Central Statistical Office	All Women	15-49	7,658	7,126
Zambia	1996	Jul-96	Jan-97	DHS	Completed	Central Statistical Office	All Women	15-49	8,021	7,286
Zambia	1992	Jan-92	May-92	DHS	Completed	University of Zambia	All Women	15-49	7,060	6,209
Zimbabwe	1999	Sep-99	Dec-99	DHS	Completed	Central Statistical Office	All Women	15-49	5,907	6,369
Zimbabwe	1994	Jul-94	Nov-94	DHS	Completed	Central Statistical Office	All Women	15-49	6,128	5,984
Zimbabwe	1988	Sep-88	Jan-89	DHS	Completed	Central Statistical Office	All Women	15-49	4,201	4,107

Country	Year	Fieldwork		Type	Status	Implementing Organization	Female			Households Sample
		Start	End				Resp.	Age	Sample	
Near East/N. Africa										
Egypt	2004	Feb-04	Jul-05	SPA	Ongoing	El-Zanaty, Inc.	Not specified		Not specified	Not specified
Egypt	2003			Interim	Completed	The National Populaton Council/Ministry of Health and Population	Ever Married Women	15-49	9,159	10,089
Egypt	2002	Sep-02	Sep-02	SPA	Ongoing	El-Zanaty & Assoc; Ministry of Health	Not specified		Not specified	Not specified
Egypt	2000	Mar-00	May-00	DHS	Completed	Nat. Population Council	Ever Married Women	15-49	15,573	16,957
Egypt	1998	Nov-98	Dec-98	Interim	Completed	El-Zanaty & Associates	Ever Married Women	15-49	6,406	6,759
Egypt	1997	Oct-96	Mar-97	In Depth	Completed	Nat. Population Council	Ever Married Women	15-49	2,444	
Egypt	1997	Nov-97	Dec-97	Interim	Completed	El-Zanaty & Associates	Ever Married Women	15-49	5,554	
Egypt	1995	Nov-95	Jan-96	DHS	Completed	Nat. Population Council	Ever Married Women	15-49	14,779	15,567
Egypt	1992	Nov-92	Dec-92	DHS	Completed	Nat. Population Council	Ever Married Women	15-49	9,864	10,760
Egypt	1988	Nov-88	Jan-89	DHS	Completed	Nat. Population Council	Ever Married Women	15-49	8,911	9,805
Jordan	2002	Jul-02	Sep-02	DHS	Completed	Dept. of Statistics (DOS)	Ever Married Women	15-49	6,006	7,825
Jordan	1997	Jun-97	Oct-97	DHS	Completed	Dept. of Statistics	Ever Married Women	15-49	5,548	7,335
Jordan	1990	Sep-90	Dec-90	DHS	Completed	Dept. of Statistics/Min. of Planning	Ever Married Women	15-49	6,461	8,333
Mauritania	2000	Oct-00	Dec-00	DHS	Ongoing	Office Nat. de la Statistique	All Women	15-49	6,500	
Morocco	2003	Oct-03	Jan-04	DHS	Ongoing	SEIS-Ministry of Health	All Women	15-49	12,000	12,000
Morocco	1995	Apr-95	May-95	Panel	Completed	Min. de la Santé Publique	All Women	15-49	4,753	2,751
Morocco	1992	Jan-92	Apr-92	DHS	Completed	Min. de la Santé Publique	All Women	15-49	9,256	6,577
Morocco	1987	May-87	Jul-87	DHS	Completed	Min. de la Santé Publique	Ever Married Women	15-49	5,982	6,960
Tunisia	1988	Jun-88	Oct-88	DHS	Completed	Office Nat. de la Fam. et de la Pop.	Ever Married Women	15-49	4,184	5,645
Yemen	1997	Oct-97	Dec-97	DHS	Completed	Central Statistical Organization	Ever Married Women	15-49	10,414	10,701
Yemen	1991	Nov-91	Jan-92	DHS	Completed	Central Statistical Organization	Ever Married Women	15-49	5,687	12,836

Country	Year	Fieldwork		Type	Status	Implementing Organization	Female			Households Sample
		Start	End				Resp.	Age	Sample	
Europe/Eurasia										
Armenia	2000	Oct-00	Dec-00	DHS	Completed	National Statistical Service/MOH	All Women	15-49	6,430	5,980
Kazakhstan	1999	Jul-99	Sep-99	DHS	Completed	Academy of Preventive Medicine	All Women	15-49	4,800	5,844
Kazakhstan	1995	May-95	Aug-95	DHS	Completed	Nat. Institute of Nutrition	All Women	15-49	3,771	4,178
Kyrgyz Republic	1997	Aug-97	Nov-97	DHS	Completed	Inst. of Obst. & Ped., MOH	All Women	15-49	3,848	3,672
Turkey	1998	Aug-98	Nov-98	DHS	Completed	Hacettepe University Inst. of Pop. Studies	Ever Married Women	15-49	8,576	8,059
Turkey	1993	Aug-93	Oct-93	DHS	Completed	Hacettepe University Inst. of Pop. Studies/Min. of Health	Ever Married Women		6,519	8,619
Turkmenistan	2000	Jul-00	Oct-00	DHS	Completed	MCH/MOH and MIT	All Women	15-49	7,919	6,303
Uzbekistan	2002	Sep-02	Dec-02	Special	Completed	Uzbekistan and National Department of Statistics of Uzbekistan	All Women	15-49	5,463	4,168
Uzbekistan	1996	Jun-96	Oct-96	DHS	Completed	Inst. of Obst. & Gynec./MOH	All Women	15-49	4,415	3,703

Country	Year	Fieldwork		Type	Status	Implementing Organization	Female			Households Sample
		Start	End				Resp.	Age	Sample	
Asia										
Bangladesh	2001	Jan-01	May-01	Special	Completed	Mitra & Associates/ACPR/NIPORT	Ever Married Women	13-49	103,750	99,165
Bangladesh	2000	Oct-99	Mar-00	DHS	Completed	Mitra & Associates/NIPORT	Ever Married Women	10-49	10,544	9,854
Bangladesh	1999	Jul-99	Dec-99	SPA	Completed	Mitra & Associates/NIPORT	Not specified		Not specified	Not specified
Bangladesh	1997	Nov-96	Mar-97	DHS	Completed	Mitra & Associates/NIPORT	Ever Married Women	10-49	9,127	8,682
Bangladesh	1994	Nov-93	Mar-94	DHS	Completed	Mitra & Associates/NIPORT	Ever Married Women	10-49	9,640	9,174
Cambodia	2000	Feb-00	Jun-00	DHS	Completed	Nat. Inst. Of Statistics/Min of Health	All Women	15-49	15,351	12,236
Cambodia	1998	May-98	Jul-98	Special	Completed	SAWA Cam./Nat. Inst. of Public Health	All Women	15-49	7,630	
India	1999			Benchmark	Ongoing	Various	All Women		Not specified	Not specified
India	1999	Nov-98	Jul-00	DHS	Completed	International Inst. for Pop. Sciences	Ever Married Women	15-49	90,303	92,486
India	1993	Apr-92	Sep-93	DHS	Completed	International Inst. for Pop. Sciences	Ever Married Women	13-49	89,777	88,562
Indonesia	2002	Oct-02	Mar-03	DHS	Completed	Central Bureau of Stat./NFPCB/MOH	Ever Married Women	15-49	29,483	33,088
Indonesia	1997	Sep-97	Dec-97	DHS	Completed	Central Bureau of Stat./NFPCB/MOH	Ever Married Women	15-49	28,810	34,255
Indonesia	1994	Jul-94	Nov-94	DHS	Completed	Central Bureau of Stat./NFPCB/MOH	Ever Married Women	15-49	28,168	33,738
Indonesia	1991	May-91	Jul-91	DHS	Completed	Central Bureau of Stat./NFPCB/MOH	Ever Married Women	15-49	22,909	26,858
Indonesia	1987	Sep-87	Dec-87	DHS	Completed	Central Bureau of Statistics/NFPCB	Ever Married Women	15-49	11,884	14,142
Myanmar	1996	Apr-96	Jan-97	Special	Completed	Settlmt. and Land Rec. Dep., Min. of Agr.				20,270
Nepal	2001	Jan-01	Jun-01	DHS	Completed	Min. of Health/New ERA	Ever Married Women	15-49	8,726	8,602
Nepal	1996	Jan-96	Jun-96	DHS	Completed	Min. of Health/New ERA	Ever Married Women	15-49	8,429	8,082
Nepal	1987	Feb-87	Apr-87	In Depth	Completed	New ERA	Currently Married Women	15-49	1,630	4,709
Pakistan	1991	Dec-90	May-91	DHS	Completed	Nat. Institute of Population Studies	Ever Married Women	15-49	6,611	7,193
Philippines	2003	Jun-03	Sep-03	DHS	Ongoing	National Statistics Office	All Women	15-49	13,633	12,586
Philippines	1998	Feb-98	Apr-98	DHS	Completed	Nat. Statistics Office, Dept. of Health	All Women	15-49	13,983	12,407
Philippines	1993	Mar-93	May-93	DHS	Completed	Nat. Statistics Office	All Women	15-49	15,029	12,995
Philippines	1993	Oct-93	Dec-93	In Depth	Completed	Nat. Statistics Office	All Women	15-49	8,431	12,995
Sri Lanka	1987	Jan-87	Mar-87	DHS	Completed	Dept. of Cen. & Stat./Min. of Plan Impl.	Ever Married Women	15-49	5,865	7,669
Thailand	1987	Mar-87	Jun-87	DHS	Completed	Inst. of Pop. Studies, Chulalongkorn U.	Ever Married Women	15-49	6,775	9,045
Vietnam	2002	Oct-02	Dec-02	DHS	Completed	General Statistical Office	Ever Married Women	15-49	5,665	7,048
Vietnam	1997	Jul-97	Oct-97	DHS	Completed	Nat. Committee for Pop. and Fam. Plan.	All Women	15-49	5,664	7,001

Country	Year	Fieldwork		Type	Status	Implementing Organization	Female			Households Sample
		Start	End				Resp.	Age	Sample	
Latin America and the Caribbean										
Bolivia	2003	Jul-03	Sep-04	DHS	Completed	National Statistical Institue, INE	All Women	15-49	19,207	19,207
Bolivia	1998	Mar-98	Sep-98	DHS	Completed	Inst. Nacional de Estadistica	All Women	15-49	11,187	12,109
Bolivia	1994	Nov-93	May-94	DHS	Completed	Inst. Nacional de Estadistica	All Women	15-49	8,603	9,114
Bolivia	1989	Feb-89	Jul-89	DHS	Completed	Inst. Nacional de Estadistica	All Women	15-49	7,923	8,439
Brazil	1996	Mar-96	Jun-96	DHS	Completed	Soc. Civil Bem-Estar Fam.	All Women	15-49	12,612	13,283
Brazil	1991	Sep-91	Dec-91	DHS	Completed	Soc. Civil Bem-Estar Fam. no Brasil	All Women	15-49	6,222	6,064
Brazil	1986	May-86	Aug-86	DHS	Completed	Soc. Civil Bem-Estar Fam.	All Women	15-44	5,892	13,283
Colombia	2004	Jul-04	Jun-05	DHS	Ongoing					
Colombia	2000	Mar-00	Jul-00	DHS	Completed	PROFAMILIA	All Women	15-49	11,585	10,907
Colombia	1995	Mar-95	Jun-95	DHS	Completed	PROFAMILIA	All Women	15-49	11,140	10,112
Colombia	1990	May-90	Aug-90	DHS	Completed	PROFAMILIA	All Women	15-49	8,644	7,412
Colombia	1986	Oct-86	Dec-86	DHS	Completed	Corp. Cen. Reg. de Prb./Min. de Salud	All Women	15-49	5,329	4,273
Dominican Republic	2002	Jun-02	Oct-02	DHS	Completed	CESDEM, Centro de Estudios Sociodemograficos	All Women	15-49	23,384	27,135
Dominican Republic	1999	Aug-99	Dec-99	Pre-test	Completed	Cen. Estud. Soc. y Dem. (CESDEM)	All Women	15-49	1,286	1,381
Dominican Republic	1996	Sep-96	Dec-96	DHS	Completed	Cen. Estud. Soc. y Dem./PROFAMILIA	All Women	15-49	8,422	8,831
Dominican Republic	1991	Jul-91	Nov-91	DHS	Completed	Assoc. Domin. Pro-Bienestar de la Fam.	All Women	15-49	7,320	7,144
Dominican Republic	1986	Sep-86	Dec-86	DHS	Completed	Consejo Nac. de Población y Familia	All Women	15-49	7,649	7,152
Dominican Republic	1986	Sep-86	Dec-86	Experimental	Completed	Consejo Nac. de Población y Familia	All Women	15-49	3,885	7,152
Ecuador	1987	Jan-87	Mar-87	DHS	Completed	Cen. de Estud. de Pob. y Pater. Res.	All Women	15-49	4,713	4,578
El Salvador	1985	May-85	Jun-85	DHS	Completed	Asociación Demográfica Salvadoreña	All Women	15-49	5,207	4,922
Guatemala	1999	Nov-98	Apr-99	Interim	Completed	Inst. Nacional de Estadistica	All Women	15-49	6,021	5,587
Guatemala	1997	Mar-97	Jun-97	In Depth	Completed	Inst. Nacional de Estadistica				2,603
Guatemala	1997	Feb-97	Jun-97	SPA	Completed	Inst. Nacional de Estadistica				
Guatemala	1995	Jun-95	Dec-95	DHS	Completed	Inst. Nacional de Estadistica	All Women	15-49	12,403	11,754
Guatemala	1987	Oct-87	Dec-87	DHS	Completed	Inst. de Nutrición de Cent. y Panamá	All Women	15-44	5,160	5,459
Haiti	2000	Mar-00	Jul-00	DHS	Completed	Inst. Haitien de l'Enfance	All Women	15-49	10,159	9,595
Haiti	1994	Jul-94	Jan-95	DHS	Completed	Inst. Haïtian de l'Enfance	All Women	15-49	5,356	4,818

Country	Year	Fieldwork		Type	Status	Implementing Organization	Female			Households Sample
		Start	End				Resp.	Age	Sample	
Latin America & Caribbean										
Mexico	2000	Jan-00	Mar-00	SPA	Completed	Nat. Institute of Public Health	Not specified		Not specified	Not specified
Mexico	1987	Feb-87	May-87	DHS	Completed	Dir. Gen. de Plan. Fam./ Sec. de Salud	All Women	15-49	9,310	7,786
Nicaragua	2001	Sep-01	dec-01	DHS	Completed	INEC (Instituto Nacional De Estadisticas Y Censos)	All Women	15-49	13,060	11,328
Nicaragua	1997	Nov-97	Jan-98	DHS	Completed	Inst. Nacional de Estadisticas y Censos	All Women	15-49	13,634	11,528
Paraguay	1990	May-90	Aug-90	DHS	Completed	Cen. Paraguayo de Estudios de Pob.	All Women	15-49	5,827	5,683
Peru	2003	May-03	Dec-06	DHS	Ongoing	INEI	All Women	15-49	6,000	6,000
Peru	2000	Jul-00	Nov-00	DHS	Completed	Inst. Nacional de Estadistica e Informática	All Women	15-49	27,843	28,900
Peru	1996	Aug-96	Nov-96	DHS	Completed	Inst. Nacional de Estadistica e Informática	All Women	15-49	28,951	28,122
Peru	1992	Oct-91	Mar-92	DHS	Completed	Inst. Nacional de Estadistica e Informática	All Women	15-49	15,882	13,479
Peru	1986	Sep-86	Dec-86	DHS	Completed	Inst. Nacional de Estadistica e Informática	All Women	15-49	4,999	4,497
Peru	1986	Sep-86	Dec-86	Experimental	Completed	Inst. Nacional de Estadistica e Informática	All Women	15-49	2,534	
Trinidad & Tobago	1987	May-87	Aug-87	DHS	Completed	Family Plan. Assoc.	All Women	15-49	3,806	4,122

World Bank Living Standards Measurement Surveys (LSMS)/Integrated Surveys

			FIELD WORK				
Region	**Country**	**Year**	**Start Date**	**End Date**	**Type**	**Implementing Organization**	**Sample Size**
ECA	Albania	2002	Apr-02	Jul-02	LSMS		3,599
ECA	Albania (excludes Tirana)	1996	Sep-96	Oct-96	LSMS	Ministry of Labor and Social Protection	1,500
MNA	Algeria	1995			LSMS		5,900
ECA	Armenia	1996	Nov-96	Dec-96	LSMS		5,000
ECA	Azerbaijan	1995	Nov-95	Dec-95	LSMS	State Statistical Committee	2,016
LAC	Bolivia	1999			LSMS	National Institute for Statistics	3,247
LAC	Bolivia	2000			LSMS	National Institute for Statistics	
ECA	Bosnia Herzegovina	2001	Sep-01	Nov-01	LSMS		5,402
LAC	Brazil (Northeast & Southeast only)	1996	Mar-96	Mar-97	LSMS	Institute of Geography and Statistics	4,940
ECA	Bulgaria	1995	May-95	Jul-95	LSMS	Gallup International	2,468
ECA	Bulgaria	1997	Mar-97	Sep-97	LSMS	Gallup International	2,314
ECA	Bulgaria	2001	Apr-01	May-01	LSMS	Gallup International	2,634
EAS	Cambodia	1997	May-97	Jun-97	LSMS	National Institute of Statistics	6,010
EAS	China (Hebei & Liaoning only)	1995	Jul-95	Jul-95	LSMS	World Bank	1,000
AFR	Côte d'Ivoire	1985	Feb-85	Jan-86	LSMS	Statistics Department	1,588
AFR	Côte d'Ivoire	1986	Feb-86	Jan-87	LSMS	Statistics Department	1,600
AFR	Côte d'Ivoire	1987	Mar-87	Feb-88	LSMS	Statistics Department	1,600
AFR	Côte d'Ivoire	1988	May-88	Apr-89	LSMS	Statistics Department	1,600
LAC	Ecuador	1994	Jul-94	Nov-94	LSMS	Ecuadorian Professional Training Service	4,536
LAC	Ecuador	1995	Aug-95	Nov-95	LSMS	Ecuadorian Professional Training Service	5,760
LAC	Ecuador	1998	Feb-98	May-98	LSMS	National Institute for Statistics and Census	5,760
LAC	Ecuador	1998-99	Nov-98	Sep-99	LSMS	National Institute for Statistics and Census	5,824
AFR	Gambia	1992	Nov-92	Mar-93	IS	Central Statistics Department	1,400
AFR	Ghana	1987/88	Sep-87	Aug-88	IS	Ghana Statistical Service	3,200
AFR	Ghana	1988/89	Sep-88	Sep-89	IS	Ghana Statistical Service	3,200
AFR	Ghana	1991/92	Sep-91	Sep-92	IS	Ghana Statistical Service	3,200
AFR	Ghana	1998/99	Apr-98	Feb-99	IS	Ghana Statistical Service	3,200
LAC	Guatemala	2000	Jul-00	Dec-00	LSMS	National Institute for Statistics and Census	8,500
AFR	Guinea	1994	Feb-94	Jan-95	IS	National Department of Statistics and Information	4,705
LAC	Guyana	1992/93	Jan-93	Jul-93	LSMS	Bureau of Statistics	1,800

SAS	India (Uttar Prades hand Bihar)	1997-98	Dec-97	Mar-98	LSMS		2,250
LAC	Jamaica	1988	Aug-88	Sep-88	LSMS	Planning Institute of Jamaica/Statistical Institute	1,909
LAC	Jamaica	1989-1	Jun-89	Aug-89	LSMS	Planning Institute of Jamaica/Statistical Institute	2,005
LAC	Jamaica	1989-2	Oct-89	Mar-90	LSMS	Planning Institute of Jamaica/Statistical Institute	3,937
LAC	Jamaica	1990	Oct-90	Mar-91	LSMS	Planning Institute of Jamaica/Statistical Institute	1,828
LAC	Jamaica	1991	Nov-91	Feb-92	LSMS	Planning Institute of Jamaica/Statistical Institute	1,786
LAC	Jamaica	1992	Aug-92	Mar-93	LSMS	Planning Institute of Jamaica/Statistical Institute	4,485
LAC	Jamaica	1993	Nov-92	Mar-93	LSMS	Planning Institute of Jamaica/Statistical Institute	1,963
LAC	Jamaica	1994	Nov-94	Jan-95	LSMS	Planning Institute of Jamaica/Statistical Institute	1,940
LAC	Jamaica	1995	May-95	Aug-95	LSMS	Planning Institute of Jamaica/Statistical Institute	1,976
LAC	Jamaica	1996	May-96	Aug-96	LSMS	Planning Institute of Jamaica/Statistical Institute	1,825
LAC	Jamaica	1997	May-97	Jul-97	LSMS	Planning Institute of Jamaica/Statistical Institute	2,020
LAC	Jamaica	1998	May-98	Aug-98	LSMS	Planning Institute of Jamaica/Statistical Institute	7,375
LAC	Jamaica	1999	May-99	Aug-99	LSMS	Planning Institute of Jamaica/Statistical Institute	6,554
LAC	Jamaica	2000	May-00	Aug-00	LSMS	Planning Institute of Jamaica/Statistical Institute	6,309
ECA	Kazakhstan	1996	Jul-96	Jul-96	LSMS	GOSKOMSTAT	2,000
ECA	Kosovo	2000	Sep-00	Dec-00	LSMS	Statistical Office of Kosovo	2,880
ECA	Krygyz Republic	1993	Oct-93	Dec-93	LSMS	GOSKOMSTAT	1,937
ECA	Krygyz Republic	Spring 1996			LSMS	GOSKOMSTAT	2,398
ECA	Krygyz Republic	Fall 1996	Nov-96	Dec-96	LSMS	GOSKOMSTAT	1,951
ECA	Krygyz Republic	1997	Sep-97	Oct-97	LSMS	GOSKOMSTAT	1,428
ECA	Krygyz Republic	1998	Oct-98	Dec-98	LSMS	GOSKOMSTAT	2,962
AFR	Madagascar	1993	Jul-93	Jul-94	IS	National Institute for Statistics	4,504
AFR	Malawi	1990	Jul-90	Jul-91	IS	National Statistical Office	6,000
AFR	Mauritania	1987	Nov-87	Oct-88	IS	Statistics Department and National Accounts	1,600
AFR	Mauritania	1989	Oct-89	Sep-90	IS	Statistics Department and National Accounts	1,600
AFR	Mauritania	1995	Oct-95	Jun-96	IS	Statistics Department and National Accounts	3,540
MNA	Morocco	1991	Oct-90	Oct-91	LSMS	Statistics Directorate	3,323
MNA	Morocco	1998			LSMS	Statistics Directorate	?
SAS	Nepal	1996	Jun-95	May-96	LSMS	Central Bureau of Statistics	3,373
LAC	Nicaragua	1993	Jun-93	Aug-93	LSMS	National Institute for Statistics and Census	4,454
LAC	Nicaragua	1998	Apr-98	Aug-98	LSMS	National Institute for Statistics and Census	4,209
LAC	Nicaragua	2001	Apr-01	Jul-01	LSMS	National Institute for Statistics and Census	4,950
AFR	Niger	1989	Feb-89	Mar-90	IS	Statistics Department and National Accounts	1,872
AFR	Niger	1992	Nov-92	Nov-93	IS	Statistics Department and National Accounts	2,070

AFR	Niger	1995	Nov-95	Dec-95	IS	Statistics Department and National Accounts	4,383
SAS	Pakistan	1991			LSMS	Federal Bureau of Statistics	4,800
LAC	Panama	1997	Jun-97	Oct-97	LSMS	Ministry of Planning and Political Economy	4,945
EAS	Papua New Guinea	1996	Jan-96	Dec-96	LSMS	University of Waikato	1,396
LAC	Paraguay	1997/98	Aug-97	Jul-98	LSMS	General Directorate for Statistics, Surveys & Census	4,353
LAC	Paraguay	1999	Aug-99	Dec-99	LSMS	General Directorate for Statistics, Surveys & Census	
LAC	Paraguay	2000/01	Sep-00	Aug-01	LSMS	General Directorate for Statistics, Surveys & Census	
LAC	Peru	1985	Jul-85	Jul-86	LSMS	National Institute for Statistics & Information	5,120
LAC	Peru (Lima only)	1990	Jun-90	Jul-90	LSMS	Cuanto	1,500
LAC	Peru	1991	Oct-91	Nov-91	LSMS	Cuanto	2,200
LAC	Peru	1994	Jun-94	Aug-94	LSMS	Cuanto	3,500
ECA	Romania	1994/95	Apr-94	Dec-94	LSMS	National Commission for Statistics	31,200
AFR	South Africa	1993	Jul-93	Apr-94	LSMS	University of Cape Town	8,850
ECA	Tajikistan	1999	May-99	Jun-99	LSMS	State Statistical Agency & Center for Strategic Studies	2,000
AFR	Tanzania (National)	1993	Sep-93	Jan-94	LSMS	Planning Commission	5,200
AFR	Tanzania-Kagera	1991			LSMS		800
MNA	Tunisia	1995/96			LSMS		3,800
AFR	Uganda	1992	Mar-92	Mar-93	IS	Ministry of Planning and Economic Development	9,929
EAS	Viet Nam	1992/93	Sep-92	Oct-93	LSMS	General Statistics Office	4,800
EAS	Viet Nam	1997/98	Dec-97	Dec-98	LSMS	General Statistics Office	5,994

Multiple Indicator Cluster Surveys (MICS)

MICS Surveys with education questions.

Eastern and Southern Africa (12)

Angola
Botswana
Burundi
Comoros
Kenya
Lesotho
Madagascar
Malawi
Rwanda
Somalia
Swaziland
Zambia

West and Central Africa (16)

Burkina Faso
Cameroon
Central African Republic
Chad
Côte d'Ivoire
Equatorial Guinea
Gambia
Guinea-Bissau
Liberia
Niger
Nigeria
Sao Tome & Principe
Senegal
Sierra Leone
Togo
Congo, Democratic Republic of

Middle East and North Africa (9)

Algeria
Bahrain
Iraq
Lebanon
Morocco
Sudan
Syria
Tunisia
West Bank & Gaza Strip

South Asia (7)

Afghanistan
Bangladesh
Bhutan
India
Maldives
Nepal
Pakistan

East Asia and the Pacific (8)

Democratic Republic of Korea
Indonesia
Laos
Mongolia
Myanmar
Philippines
Thailand
Viet Nam

Latin America and the Caribbean (7)

Bolivia
Cuba
Dominican Republic
Guyana
Suriname
Trinidad & Tobago
Venezuela

Central and Eastern Europe/Comm.of Indep.States (11)

Albania
Armenia
Azerbaijan
Bosnia and Herzegovina
Georgia
Macedonia
Moldova, Republic Of
Tajikistan
Ukraine
Uzbekistan
Yugoslavia

SAMPLE TABLES

APPENDIX D

Table 1. Distribution (%) of households, from the total population and the population *n*[1] to 14 years according to characteristics of the household head and of household members

Characteristics	*Total*						Male heads of household						Female heads of household					
	House-holds %[2] *(N)*[3]	Total population % *(N)*	Population *n*-14 years				House-holds % *(N)*	Total population % *(N)*	Population *n*-14 years				House-holds % *(N)*	Total population % *(N)*	Population *n*-14 years			
			B+G % *(N)*	B % *(N)*	G % *(N)*	Male ratio			B+G % *(N)*	B % *(N)*	G % *(N)*	Male ratio			B+G % *(N)*	B % *(N)*	G % *(N)*	Male ratio
Head of household																		
Age group																		
Marital status																		
Literacy																		
Level of education			*This table focuses on the characteristics of households and household heads, including the sex of school-age children.*															
Religion																		
Ethnicity																		
Current occupation																		
Job position																		
Household																		
Size of household																		
Type of household																		
Type of housing																		
Water supply																		
Mode of lighting																		
Occupational status																		

[1] The value *n* is defined according to the legal age of school entry. It must be specified when the guide is used by each country and for each table.
[2] The % is given in columns, with 100 for each variable.
[3] Total number for each variable.
B+G: Boys + Girls; B: Boys ; G : Girls.
Source: country, type of operation, year, month.

Totals can be disaggregated into rural, urban and capital.

Table 2. Distribution (%) of households according to number of school-age children (age *n* to 14 years) and characteristics of the household and household head

Characteristics	Number of school-age children								Mean number of school-age children by household
	0	1	2	3	4	5 and +	Total %	Total number	
Head of household									
Sex	(%)	(%)	(%)	(%)	(%)	(%)	100	(*N*)	
Age group							100	(*N*)	
Marital status							100	(*N*)	
Literacy							100	(*N*)	
Level of education							100	(*N*)	
Current occupation							100	(*N*)	
Job position		*This table focuses on the number of school-age children in the household, as well as on characteristics of the household and household head.*							
Household									
Size									
Type									
Type of housing							100	(*N*)	
Water supply							100	(*N*)	
Mode of lighting							100	(*N*)	
Occupational status							100	(*N*)	
The % is given in lines, with 100 for each characteristic. *Source: country, type of operation, year, month.*									

Totals can be disaggreated into rural, urban and capital.

Table 3a. Proportion (%) of children *n* to 14 years in school by residence and according to household and household head characteristics

Characteristics[1]	Total				Rural				Total urban				Capital			
	B+G %[2] *(N)*[3]	B % *(N°)*	G % *(N)*	Male ratio	B+G % *(N)*	B % *(N)*	G % *(N)*	Male ratio	B+G % *(N)*	B % *(N)*	G % *(N)*	Male ratio	B+G % *(N)*	B % *(N)*	G % *(N)*	Male ratio
Children																
Family status																
Age																
Head of household																
Sex																
Age group																
Marital status																
Literacy																
Level of education																
Current occupation																
Job position																
Household																
Size																
Type																
0-4 year-olds in hh					*This table measures and observes differences in school attendance among school-age children, according to their characteristics, those of the household head and other household members, the sex of the child, and area of residence.*											
No.of 15-24 in school																
Number 15+, primary																
Number 25+, secondary																
% 15+, primary																
% 25+, secondary																
Type of housing																
Water supply																
Occupational status																
Total	% *(N)*	% *(N)*	% *(N)*		% *(N)*	% *(N)*	% *(N)*		% *(N)*	% *(N)*	% *(N)*		% *(N)*	% *(N)*	% *(N)*	

[1] The variables mentioned here are not exhaustive. [2] Proportions are calculated for each modality of variable. [3] Reference number for each proportion calculated.

Source : country, type of operation, year, month.

This table can be reproduced making a distinction by sex of head of household.

Table 3b. Proportion (%) of children *n* to 14 years who have been in school by residence and according to characteristics of the household and household head

Characteristics[1]	Total				Rural				Total urban				Capital			
	B+G %[2] *(N)*[3]	B % *(N)*	G % *(N)*	Male ratio	B+G % *(N)*	B % *(N)*	G % *(N)*	Male ratio	B+G % *(N)*	B % *(N)*	G % *(N)*	Male ratio	B+G % *(N)*	B % *(N)*	G % *(N)*	Male ratio
Children																
Family status																
Age																
Head of household																
Sex																
Age group																
Marital status																
Literacy																
Level of education																
Current occupation																
Job position																
Household																
Size					*This table presents data on children of school age, according to the characteristics of households, household heads, and school-age children.*											
Type																
Presence of 0-4 year-olds																
Number of 15-24 in school																
Number 15+, primary																
Number 25+, secondary																
% 15+, primary																
% 25+, secondary																
Type of housing																
Water supply																
Occupational status																
Total	% *(N)*	% *(N)*	% *(N)*		% *(N)*	% *(N)*	% *(N)*		% *(N)*	% *(N)*	% *(N)*		% *(N)*	% *(N)*	% *(N)*	

[1] The variables mentioned here are not exhaustive. [2] The proportions are calculated for each modality of variable. [3] Reference number for each proportion calculated.
Source: country, type of operation, year, month.

This table can be reproduced making a distinction by sex of head of household.

Table 3c. Proportion (%) of children *n* to 14 years who have never been to school by residence and according to characteristics of the household and household head

Characteristics[1]	Total				Rural				Total urban				Capital			
	B+G % *(N)*	B % *(N)*	G % *(N)*	Male ratio	B+G % *(N)*	B % *(N)*	G % *(N)*	Male ratio	B+G % *(N)*	B % *(N)*	G % *(N)*	Male ratio	B+G % *(N)*	B % *(N)*	G % *(N)*	Male ratio
Children																
Family status																
Age																
Head of household																
Sex																
Age group																
Marital status																
Literacy																
Level of education																
Current occupation																
Job position																
Household																
Size																
Type																
Presence of 0-4 year-olds				*This table presents data on children of school age, according to the characteristics of households, household heads, and school-age children.*												
Number of 15-24 in school																
Number 15+, primary																
Number 25+, secondary																
% 15+, primary																
% 25+, secondary																
Type of housing																
Water supply																
Occupational status																
Total	% *(N)*	% *(N)*	% *(N)*		% *(N)*	% *(N)*	% *(N)*		% *(N)*	% *(N)*	% *(N)*		% *(N)*	% *(N)*	% *(N)*	

[1] The variables mentioned here are not exhaustive. [2] The proportions are calculated for each modality of variable. [3] Reference number for each proportion calculated.
Source: country, type of operation, year, month.

This table can be reproduced making a distinction by sex of head of household.

Table 4. Distribution (%) by age of children *n* to 14 years according to their current school participation and, for those in school, according to their year in school

AGE	In school (by year of study)										In school (total)	Dropped out	Never attended school	Total %	Total number
	1	3	3	4	5	6	7	8	9	10					
n	%	%	%	%	%	%	%	%	%	%	%	%	%	100	*(N)*
.														100	*(N)*
.														100	*(N)*
.														100	*(N)*
.		*This table studies the distribution by age of children who (a) currently attend school; (b) have been in school but are no longer in school; (c) have never attended school.*												100	*(N)*
.														100	*(N)*
.														100	*(N)*
.														100	*(N)*
.														100	*(N)*
14														100	*(N)*
Total *n*-14 years														100	*(N)*
% younger											Gross ratio of primary school attendance :				
% older											Net ratio of primary school attendance :				
GRAYS*											Gross ratio of secondary school attendance 1st cycle :				
NRAYS**											Net ratio of secondary school attendance 1st cycle :				

* Gross ratio of attendance by year of study.
** Net ratio of attendance by year of study.
Note: this table must be adapted for the legal school age and educational system of the country.
Source: country, type of operation, year, month.

Totals can be disaggregated into rural, urban and capital, or according to sex.

Table 5. Proportion (%) by age of children who have left primary school, according to sex and relationship to the household head

	Total				Boys				Girls				Male ratio
	Total	Children of HH	OR	NR	Total	Children of HH	OR	NR	Total	Children of HH	OR	NR	for the total
N	% *(N)*	% *(N)*	% *(N)*	% *(N)*									
.	% *(N)*												
.	% *(N)*												
.	% *(N)*			*This table presents information on primary school dropout among children according to their relationship to the household head.*									
	% *(N)*												
.	% *(N)*												
.	% *(N)*												
.	% *(N)*												
.	% *(N)*												
.	% *(N)*												
14	% *(N)*												
n-14 years	% *(N)*												

Children of HH: children of head of household.
OR: other relative.
NR: no kin relationship.
Note: this table must be adapted for the legal school age of the country.
Source: country, type of operation, year, month.

Totals can be disaggregated into rural, urban, or capital, or according to sex of head of household.

Table 6. Proportion (%) by last class completed or attended for children who dropped out of primary school, according to sex and relationship to the household head

Class	Total				Boys				Girls				Male ratio for the total
	Total	Chd. HH	OR	NR	Total.	Chd. HH	OR	NR	Total	Chd. HH	OR	NR	
1	% *(N)*	% *(N)*	% *(N)*	% *(N)*									
2	% *(N)*												
3	% *(N)*				*This table allows for the study of whether there are "threshold" classes at which children are most likely to drop out of primary school, according to sex and relationship to the household head.*								
4	% *(N)*												
5	% *(N)*												
6	% *(N)*												
Primary	% *(N)*												

Children of HH: children of head of household.
OR: other relative.
NR: no kin relationship.
Note: this table must be adapted for the educational system of the country.
Source: country, type of operation, year, month.

Totals can be disaggregated into rural, urban or capital, or according to sex of head of household.

Table 7. Distribution of households (by total number and in %) according to the number of school-age children in the household (age *n* to 14 years; from 1 to maximum *Na*) and the number of children in school (from 0 to maximum *Né*)

Number of children	Number of school-age children				
in school	1	2	3	*Na*	Total
0	*N* (%)* (%)**	*N* (%) (%)	*N* (%) (%)	*N* (%) (%)	*N* 100 (%)
1	*N* (%) (%)	N (%) (%)	*N* (%) (%)	*N* (%) (%)	*N* 100 (%)
2	/	*N* (%) (%)	*N* (%) (%)	*N* (%) (%)	*NE* 100 (%)
3	/	/	*N* (%) (%)	*N* (%) (%)	*N* 100 (%)
Né	/	/	/	*N* (%) (%)	*N* 100 (%)
Total number of households	*N*	*N*	*N*	*N*	*N*
(total % in column)	(100)	(100)	(100)	(100)	(100)
% in line (total)	%	%	%	%	100
% of households having at least one child in school	%	%	%	%	%
Mean number of children in school	*m*	*m*	*m*	*m*	*m*

N: number.
%*: % in line.
%**: % in column.
m: mean number.
Source: country, type of operation, year, month.

This table calculates the following indicators overall and according to the number of school-age children:
a) the % of households having at least one child in school (calculated for the total, this result constitutes the counterpart at the household level of the education ratio);
b) the % of households having no child in school (appears on the line "0 child in school");
c) the % of households having all their children in school (appears on the diagonal between number of children who are school age and in school);
d) mean number of children in school.

Table 8. Distribution (%) of households according to number of children in school by household (age *n* to 14 years) and characteristics of the household and household head

Characteristics	Number of children in school by household								% of households
	0[1]	1	2	3	4	5 and over	Total *(number)*	Mean number[2]	having all children in school[3]
Head of household									
Sex							100 *(N)*		
Age group							100 *(N)*		
Marital status							100 *(N)*		
Literacy							100 *(N)*		
Level of education									
Religion									
Ethnicity									
Current occupation									
Job position									
Household									
Size							100 *(N)*		
Type							100 *(N)*		
Type of housing							100 *(N)*		
Water supply							100 *(N)*		
Occupational status							100 *(N)*		
Total for households							100 *(N)*		

This table presents data on the number of children in school, according to household characteristics.

[1] indicator b), the complement to 100 % equivalent to indicator a), cf. table 7. [2] indicator c), cf. Table 7.
[3] indicator d), cf. .table 7. *Source: country, type of operation, year, month.*

Totals can be disaggregated into rural, urban or capital, oraccording to sex of head of household.

Table 9. Numbers of literate and illiterate household members by sex and age group

Age group (in years)	Literate			Illiterate			Not specified			Literacy ratio			
	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female	Male ratio
10-14													
15-19													
20-24													
25-29													
30-34				*This table examines the general literacy level among household members age 10 years and older, by sex and age group.*									
35-39													
40-44													
45-49													
50-54													
55-59													
60-64													
65-69													
70-74													
75 and over													
Age unknown													
Total 15 and over													

Source: country, type of operation, year, month.

Totals can be disaggregated into rural, urban or capital.
This table can be reproduced for heads of household alone.

Table 10. Distribution (%) of population 15 years and older according to the highest level of school attended, by sex and age group

Level of study	Male						Female					
	15-19	20-24	25-29	.....	65+	Total	15-19	20-24	25-29	.....	65+	Total
Never attended												
school												
Primary												
Total												
1												
2												
3												
4												
5												
6												
Secondary												
Total												
1			*This table presents information on the highest level of schooling attended by household members age 15 and older, according to sex and age group*									
2												
3												
4												
Higher												
Total												
1												
2												
3												
4												
Other												
Not specified												
Total	100	100	100	100	100	100	100	100	100	100	100	100
(Number)	*(N)*	*(N)*	*(N)*	*(N)*	*(N)*	*(N)*	*(N)*	*(N)*	*(N)*	*(N)*	*(N)*	*(N)*
Source: country, type of operation, year, month.												

Totals can be disaggregated into rural, urban or capital. This table can be reproduced for heads of household alone.

Table 11. Distribution (%) of school age children (age *n* to 14 years) according to birthplace and residence

Birthplace/ residence	Total				Boys				Girls			
	Total	Children of HH	OR	NR	Total	Children of HH	OR	NR	Total	Children of HH	OR	NR
Same village												
Same city												
R/R intra depart.												
R/other rural			*By giving the distribution of school-age children according to the variable combining their birthplace and their area of residence at the time of the survey, this table gives an approximation of the geographical mobility of these children, and allows for the examination of differences according to sex and relationship to the household head.*									
R/secondary city												
Rural/Capital												
Sec. city/Sec. city												
Sec. city/Capital												
Capital/Capital												
Sec. city/Rural												
Capital/Rural												
Total *(Number)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*

R: rural.
Intra depart.: within the same department.
Children of HH: children of head of household.
OR: other relative.
NR: no kin relationship.
Source: country, type of operation, year, month.

Table 12. Distribution (%) of children *n* to14 years according to their sex and family status, in relation to the survival and residential status of their parents

Status	Total				Boys				Girls			
and residence of parents	Total	Children of HH	OR	NR	Total	Children of HH	OR	NR	Total	Children of HH	OR	NR
Lives w/both parents												
Lives w/mother, father elsewhere												
Lives w/mother, father deceased												
Lives w/father, mother elsewhere												
Lives w/father, mother deceased												
Father and mother elsewhere												
Father elsewhere, mother deceased												
Father deceased, mother elsewhere												
Father and mother deceased												
Unknown												
Total *(Number)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*

Children of HH: children of head of household. OR: other relative. NR: no kin relationship.
Source: country, type of operation, year, month.

Totals can be disaggregated into rural, urban or capital.

Table 13. Proportion (%) of children *n* to 14 years in school according to their sex and relationship to the household head

	Total				Boys				Girls			
Characteristics	Total	Child. of HH	OR	NR	Total	Child. of HH	OR	NR	Total	Child. of HH	OR	NR
Head of household												
Sex	% *(N)*	% *(N)*	% *(N)*	% *(N)*								
Age group												
Marital status												
Literacy												
Level of education		*This table shows differences in schooling according to the characteristics of the head of household, the size and composition of the household and other household characteristics.*										
Current occupation												
Household												
Size												
Type												
Type of housing												
Water supply												
Mode of lighting												
Occupational status												
Total	% *(N)*	% *(N)*	% *(N)*	% *(N)*	% *(N)*	% *(N)*	% *(N)*	% *(N)*	% *(N)*	% *(N)*	% *(N)*	% *(N)*

Children of HH: children of head of household.
OR: other relative.
NR: no kin relationship.
Source: country, type of operation, year, month.

Totals can be disaggregated into rural, urban or capital.

Table 14. Proportion (%) of children in school (age *n* to 14 years) according to birthplace and residence

Birthplace/ residence	Total				Boys				Girls			
	Total	Child. of HH	OR	NR	Total	Child. of HH	OR	NR	Total	Child. of HH	OR	NR
Same village												
Same city												
R/R intra depart.												
Rural/other rural		*This table shows whether there are differences in school attendance according to children's geographic mobility.*										
R/Secondary city												
Rural/Capital												
Sec. city/Sec. city												
Secondary city/Capital												
Capital/Capital												
Secondary city/Rural												
Capital/Rural												
Total *(Number)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*	100 *(N)*

R: rural.
Intra depart: within the same department.
Children of HH: children of head of household.
OR: other relative.
NR: no kin relationship.
Source: country, type of operation, year, month.

Table 15. Proportion (%) of children in school according to their sex and family status, in relation to the survival and residential status of parents

Survival and residence of parents	Total				Boys				Girls			
	Total	Child. of HH	OR	NR	Total	Child. of HH	OR	NR	Total	Child. of HH	OR	NR
Lives with both parents	% *(N)*	% *(N)*	% *(N)*	% *(N)*	% *(N)*	% *(N)*	% *(N)*	% *(N)*	% *(N)*	% *(N)*	% *(N)*	% *(N)*
Lives with mother, father elsewhere												
Lives with mother, father deceased												
Lives with father, mother elsewhere												
Lives with father, mother deceased												
Father and mother elsewhere												
Father elsewhere, mother deceased												
Father deceased, mother elsewhere												
Father and mother deceased												
Situation not clear												

Children of HH: children of head of household.
OR: other relative.
NR: no kin relationship.
Source: country, type of operation, year, month.

Totals can be disaggregated into rural, urban or capital.

Table 16. Probability of promotion, repetition and drop-out for school population n to 14 years by age and sex

Age	Promotion			Repetition			Dropout			Number		
	B+G	B	G	B+G	B	G	B+G	B	G	B+G	B	G
5												
6			*This table presents information on the probability of promotion, repetition and dropping out, for children in school, by age of the child.*									
7…												
.												
.												
14												
n-14 years												
Source: country, type of operation, year, month.												

Totals can be disaggregated into rural, urban or capital.

Table 17. Probability of promotion, repetition and drop-out for school population *n* to 14 years by year of study and sex

Year of	Promotion			Repetition			Drop out			Number		
study	B+G	B	G	B+G	B	G	B+G	B	G	B+G	B	G
1												
2												
3			*This table presents information on the probability of promotion, repetition and dropping out, for children in school, by age of the child.*									
4												
5												
6												
7												
8												
9												
10												
Note: this table must be adapted for the school system of the country. *Source: country, type of operation, year, month.*												

Totals can be disaggregated into rural, urban or capital.

Table 18a. Proportion (%) of children *n* to 14 years who are both working and in school, by year of age, according to sex and relationship to household head

Age	Total			Boys			Girls		
	Total	Child. Of HH	Others	Total	Child. of HH	Others	Total	Child. of HH	Others
n	% *(N)*								
.	% *(N)*								
.	% *(N)*								
.	% *(N)*		*This table evaluates the portion of children who are working among those who attend school, by age and according to the sex and family status of the child.*						
.	% *(N)*								
.	% *(N)*								
.	% *(N)*								
.	% *(N)*								
.	% *(N)*								
.	% *(N)*								
14	% *(N)*								
n-14 years	% *(N)*								

Children of HH: children of head of household.

Source: country, type of operation, year, month.

Totals can be disaggregated into rural, urban or capital.

Table 18b. Proportion (%) of children *n* to 14 years who are working among children who have attended school, by year of age, according to sex and relationship to household head

	Total			Boys			Girls		
Age	Total	Child. of HH	Others	Total	Child. of HH	Others	Total	Child. of HH	Others
n	% *(N)*								
.	% *(N)*								
.	% *(N)*								
.	% *(N)*		*This table presents information on the percentage of children who are working among those who have been to school.*						
.	% *(N)*								
.	% *(N)*								
.	% *(N)*								
.	% *(N)*								
.	% *(N)*								
.	% *(N)*								
14	% *(N)*								
n-14 years	% *(N)*								

Children of HH: children of head of household.

Source: country, type of operation, year, month.

Totals can be disaggregated into rural, urban or capital.

Table 18c. Proportion (%) of children *n* to 14 years who are working among children who have never been to school, by year of age, according to sex and relationship to household head

Age	Total			Boys			Girls		
	Total	Child. of HH	Others	Total	Child. of HH	Others	Total	Child. of HH	Others
n	% *(N)*								
.	% *(N)*								
.	% *(N)*								
.	% *(N)*		*This table presents information on the percentage of children who are working among those who have been to school.*						
.	% *(N)*								
.	% *(N)*								
.	% *(N)*								
.	% *(N)*								
.	% *(N)*								
.	% *(N)*								
14	% *(N)*								
n-14 years	% *(N)*								

Children of HH: children of head of household.

Source: country, type of operation, year, month.

Totals can be disaggregated into rural, urban or capital.

Table 19a. Distribution (%) of children *n* to 14 years who are in school according to their type of stated work, for the total population and by sex

Type of work				Total	*(Number)*
Working		Not working	Other		
Total	Housework				
Total					
				100	*(N)*
				100	*(N)*
Boys					
				100	*(N)*
				100	*(N)*
Girls					
				100	*(N)*
				100	*(N)*
Source: country, type of operation, year, month.					

This table presents information on type of work done by children who currently attend school.

Table 19b. Distribution (%) of children *n* to 14 years who have been to school according to type of work, for the total population and by sex

Type of work				Total	*(Number)*
Working		Not working	Other		
Total	Housework				
Total					
				100	*(N)*
				100	*(N)*
Boys					
				100	*(N)*
				100	*(N)*
Girls					
				100	*(N)*
				100	*(N)*
Source: country, type of operation, year, month.					

This table presents information on type of work done by children who currently attend school.

Totals can be disaggregated into rural, urban or capital.

Table 19c. Distribution (%) of children *n* to 14 years who have never been to school according to their type of stated work, for the total population and by sex

Type of work				Total	*(Number)*
Working		Not working	Other		
Total	Housework				
Total					
				100	*(N)*
				100	*(N)*
Boys					
				100	*(N)*
				100	*(N)*
Girls					
				100	*(N)*
				100	*(N)*
Source: country, type of operation, year, month.					

This table presents information on type of work done by children who currently attend school.

Totals can be disaggregated into rural, urban or capital.

Table 20a. Distribution (%) of children *n* to 14 years who are in school and working, according to their job position, for each year of age, for the total population and by sex

Age	JOB POSITION					
	Independent	Family helper	Appren-tice	Other	Total	*(Number)*
Ensemble						
N					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
14					100	*(N)*
Total					100	*(N)*
Boys						
N					100	*(N)*
.					100	*(N)*
.						
.						
.						
.						
.						
.						
.					100	*(N)*
14					100	*(N)*
Total					100	*(N)*
Girls						
N					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.	*This table presents information on type of work done by children who currently attend school, according to level of school attended.*					
.						
.						
.					100	*(N)*
.					100	*(N)*
14					100	*(N)*
Total					100	*(N)*
Note: one can distinguish between children who have been to school or not. *Source: country, type of operation, year, month*.						

Totals can be disaggregated into rural, urban or capital.

Table 20b. Distribution (%) of children *n* to 14 years who have been to school and are working, according to their job position for each year of age, for the total population and by sex

Age	Job position					
	Independent	Family helper	Apprentice	Other	Total	*(Number)*
Total						
N					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
14					100	*(N)*
Total					100	*(N)*
Boys						
n					100	*(N)*
.					100	*(N)*
.						
.	*This table presents information on type of work done by children who currently attend school, according to the level of schooling attained.*					
.						
.						
.						
.						
14					100	*(N)*
Total					100	*(N)*
Girls						
n					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
14					100	*(N)*
Total					100	*(N)*
Source: country, type of operation, year, month.						

Totals can be disaggregated into rural, urban or capital.

Table 20c. Distribution (%) of children *n* to 14 years who have never been to school and are working by job position, year of age and sex

AGE	Job position					
Total	Independent	Family helper	Appren-tice	Other	Total	*(Number)*
N					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
14					100	*(N)*
Total					100	*(N)*
Boys						
n					100	*(N)*
.					100	*(N)*
.	*This table presents information on type of work done by children who currently attend school, according to the level of schooling attained.*					
.						
.						
.						
.						
.						
.						
14					100	*(N)*
Total					100	*(N)*
Girls						
n					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
.					100	*(N)*
14					100	*(N)*
Total					100	*(N)*
Source: country, type of operation, year, month.						

Totals can be disaggregated into rural, urban or capital.

ANALYSING DATA: A PRACTICAL EXAMPLE

APPENDIX E

We use data from the 1998 DHS survey for Ghana in order to illustrate an example of a typical data analysis. Household weights are used in the analysis to reflect the complex nature of the design. The non-response adjustment is built into the final household weight and is transparent to the user so is automatically accounted for. In any case, the household response rate is 99.1%. Therefore, it is safe to ignore variability due to non-response when forming variance or standard error estimates. For this particular example, we wish to test the hypothesis that there is no difference between male and female primary school net attendance ratios (NAR). We start by calculating the primary school NAR estimates separately for males and females.

$$NAR_{males} = \frac{Weighted\ sum\ of\ all\ males\ in\ the\ sample\ age\ 6-11\ who\ attend\ primary\ school}{Weighted\ sum\ of\ all\ males\ in\ the\ sample\ age\ 6-11}$$

$$NAR_{males} = \frac{1479.7}{1955.4} = 0.75674$$

$$NAR_{females} = \frac{Weighted\ sum\ of\ all\ females\ in\ the\ sample\ age\ 6-11\ who\ attend\ primary\ school}{Weighted\ sum\ of\ all\ females\ in\ the\ sample\ age\ 6-11}$$

$$NAR_{females} = \frac{1393.5}{1853.87} = 0.75168$$

In order to test the hypothesis that males and females have the same net attendance ratio, we need to compute the difference between the two ratios as well as the standard error of the difference which we will call SE (NAR_{males} - $NAR_{females}$). We use the *descript* procedure of the SUDAAN package to do so. In computing the standard error, SUDAAN, unlike many standard statistical packages, takes into account the complexity of the sample design, namely clustering, stratification and differential probabilities of selection. The standard error of the difference given by SUDAAN is SE (NAR_{males} - $NAR_{females}$) = 1.34. Therefore the t-test statistic is given by:

$$t = \frac{NAR_{males} - NAR_{females}}{SE\ (NAR_{males} - NAR_{females})}$$

$$t = \frac{0.75674 - 0.75168}{1.34} = 0.38$$

Next is to compare this calculated t-statistic with the value of the t-statistic obtained from the t-distribution with 210 degrees of freedom. The degree of freedom for any sample survey is the number of primary sampling units (PSUs) minus the number of strata. For the 1998 Ghana DHS survey, there are 400 PSUs and 190 strata. At the 5% level of significance, the value of the t-statistic from the t-distribution with 210 degrees of freedom is 1.645. The calculated t-statistic is smaller than the value given by the t-distribution. Therefore, we do not reject the hypothesis that males and females have the same NAR. That means there is no statistically significant difference in the NAR between males and females.

We can arrive at the same conclusion by considering what is called the p-value. At the 5% level of significance, a p-value that is less than 0.05 leads to the rejection of the hypothesis of no difference, while we cannot reject this hypothesis when the p-value equals or is greater than 0.05. SUDAAN automatically produces the calculated t-test statistic and the corresponding two-sided p-value (p=0.7054). We therefore conclude that there is no statistically significant difference in the NAR between males and females.

By comparing the values of the calculated NAR for males and females without the use of statistical hypothesis testing, one could believe that there is no gender difference in the NAR. However, there may be cases where the two figures are close but a statistically significant difference exists between males and females. In other cases, the two ratios may be far apart but are not statistically different. To make valid conclusions, one should always use appropriate statistical hypothesis testing.