World Bank

Water and Sanitation Household Survey at Project Level - Ethiopia

Final Report

April 2005

COWI

World Bank

Water and Sanitation Household Survey at Project Level - Ethiopia

Final Report

April 2005

Report no.	60213A-1
Issue no.	6
Date of issue	21 April 2005
Prepared	CJP, CGH, MEB, TK
Checked	CAG, VR
Approved	MEB

Table of Contents

Table of Appendices

Abbreviations

BSU	Basic sampling units
DHS	Demographic and Health Survey
JMP	Joint Monitoring Programme
LSMS	Livings Standards Measurement Studies
MICS	Multiple Indicator Cluster Survey
PPS	Probability proportionate to size
PSU	Primary sampling units
SPSS	Statistical Package for the Social Sciences (statistical software system)
VIP	Ventilated improved pit (latrine)
WHO	World Health Organisation
WHS	World Health Survey
BSU	Basic sampling units

1 Introduction

The Water and Sanitation Sector of the World Bank is facing increasing demands for access and performance data. This is driven in part by a corporate focus on the Millennium Development Goals and the World Bank's contribution to their achievement, and in part by a general need to better monitor the impact of sector interventions. The aim of this report is to propose a range of dedicated project level WSS indicators at household level and a survey methodology on how to obtain these indicators.

The goal of the household survey presented in this report is to obtain a better understanding of households' access to water, sanitation, and hygiene across a project area, and the safety and quality of the services provided from a household perspective. The methodology has been developed and tested on urban households in Ethiopia and rural households in Peru. The report includes the results from the case study in Ethiopia and the result from Peru is presented in a second report.

The survey instrument has been developed taking the work of the Joint Monitoring Programme (JMP) as our starting point. Hence the questions on water and sanitation in the JMP-questionnaire form part of survey instrument presented in this report. However, whereas the JMP approach aims at few comparable indicators which can be measured as part of a general national household survey, the aim of the methodology described here is to obtain more detailed information on water, sanitation and hygiene in a geographically limited project area.

For future monitoring of surveys at project level, it is proposed to take the questionnaire presented in this report as a starting point and to adjust the questions to a local setting using qualitative household interviews. A set of 28 indicators is proposed from which the most locally relevant may be selected as the basis on which the monitoring should be carried out.

For the sample selection, a multiple stage cluster sampling methodology is proposed. When this procedure is followed carefully, a statistically representative sample is obtained which is crucial for the applicability of the results. The approach enables data collection with a balance between the precision of the estimates and practical implementation costs.

The practical steps of survey administration include preparatory activities, testing, and data collection. Here we have placed emphasis on the *training of the enumerators* to ensure that they understand the task they are undertaking, and on *quality control* to ensure that the interviews are conducted in a professional manner which yields more accurate results.

The case study for Ethiopia is reported in the last chapter of the report. A subset of 73 towns in the Oromia region was selected as the target of the survey. These towns were selected for projects to improve the water and sanitation services and are generally characterised by a low level of services. A two-stage cluster sampling procedure was applied to select the sample. From a total of 73 towns, 24 towns were selected randomly with a probability proportionate to size (PPS). In the second stage of the sample selection, enumeration maps were used in order to obtain a sample of households that is spread across each town. A total of 1,013 households were interviewed.

The results of the survey are reported by presenting the 28 selected indicators, their mean values, and the confidence interval for the mean value.

The survey indicates that 46% (±7%) of the households have access to improved water services and 38% (±5%) of the households have access to an improved sanitation facility, according to the definitions adopted in this report.

More than four out of every five households have soap, while few households (5%, ±3%) are observed to have adopted appropriate hand washing practices, (as gauged by the absence of one or more of the necessary preconditions for hand washing: soap, water and sink, bucket or equivalent.) At the same time 30% (±9%) of the respondents claim to wash their hands at critical times.

This means that there is a discrepancy in the results depending on the method used to obtain the data; this may be due to biases in both indicators; the enumerators may not be probing sufficiently to se the hand washing practice of household members, and the respondents may respond with “expected” or “desirable” behaviour rather than actual behaviour when asked to state when they wash their hands. It is proposed to use both observation and questions, although most emphasis should be put on the results from the questions when interpreting the results.

Ten percent of the sample was back-checked after the end of the main data collection phase by having a separate group of enumerators visiting the interviewed households. The back-checking results include information about the two central indicators; access to improved drinking water and access to improved sanitation; and household size.

The back-checking reveals a discrepancy between the results from the main data collection phase and the back-checking. This underlines that the back-checking should be used actively during the data collection in order to ensure early correction of quality flaws.

The remaining of this report is organised as follows. Chapter 2 presents the content of the survey instrument and the selected indicators. Chapter 3 describes the multiple stage cluster sampling method proposed and Chapter 4 presents the implementation procedure. The experience gained in Ethiopia, the first of two pilot studies, is presented in Chapter 5.

2 Design of survey instrument

This chapter presents the main survey instrument, the household questionnaire on water supply, sanitation and hygiene practices, and a set of indicators that is proposed to describe the level of services. A questionnaire is designed with the purpose of collecting the data needed to calculate the proposed set of indicators. The criteria for the selection of the indicators has been to identify a small number of indicators that are comparable between projects and over time.

When designing a monitoring and evaluation approach for a specific project a subset of the 28 indicators can be selected and the questionnaire should be adjusted accordingly. It is important to have one approach that is generally accepted and used to facilitate comparability over time and across projects. In addition, adapting an existing survey instrument to local conditions is considerably less expensive than designing a survey instrument for each project.

Taking the work by the Joint Monitoring Programme (JMP) and other questionnaires as a starting point, a questionnaire is developed, and from the information gathered in the questionnaire a limited set of indicators is selected. The selection is designed such that an overview of the water supply, sanitation and hygiene practices can be obtained. This maintains the focus on important indicators and keeps the cost of the survey down.

The proposed survey instrument and methodology have been tested in Ethiopia whereas the test in Peru remains to be conducted. Chapter 6 presents the results of the indicators found in Ethiopia, the results for Peru are presented in a separate report.

2.1 Background

The WHO/UNICEF Joint Monitoring Programme for Water Supply and Sanitation (JMP) is the official United Nations mechanism for reporting on progress in achieving the water and sanitation targets of the Millennium Development Goals.

Previously only data provided by governments and service providers were used for evaluating water supply and sanitation coverage. Since the "Global Water Supply and Sanitation Assessment 2000", the JMP uses a multiple-survey approach for estimating water supply and sanitation coverage. Household data from multiple household surveys such as Demographic and Health Surveys

(DHS), Multiple Indicator Cluster Surveys (MICS), World Health Surveys (WHS), Living Standards Measurement Studies (LSMS) and national censuses. While this approach is considered to provide the best national estimates possible, i a past drawback has been the lack of comparability between survey results stemming from the use of different questions by each survey. Also, most surveys focus on water supply and sanitation alone and do not include hygiene practices, although it is generally agreed that hygiene is a crucial determinant of the health effect of water and sanitation interventions.

To cope with this lack of uniformity, the JMP has recently made an effort to develop a harmonised core set of questions that have already been adopted by several international survey programmes including the DHS, MICS, and WHS.

Whereas the JMP approach aims at a few comparable indicators on the national level, the aim of the methodology described here is to obtain more detailed information on water, sanitation and hygiene in a geographically limited project area.

The questionnaire proposed here for household surveys on water supply, sanitation and hygiene at project level takes its starting point in this core set of water, sanitation and hygiene questions as reflected in the JMP's "Guide for Water Supply, Sanitation and Hygiene Related Survey Questions" and "Questions for Household Survey Module - Water Supply, Sanitation and Hand Washing", June 2004. This report describes how the instrument was then developed, tested and documented in Ethiopia.

In response to the specific requirements of the instruments, lessons learnt from more detailed and specific questionnaires have been taken into consideration, in particular those related to hygiene behaviour. The following documents were the major sources of inspiration:

- The London School of Hygiene and Tropical Medicine's WASH survey method developed under the Water, Sanitation and Hygiene campaign of the Water Supply and Sanitation Collaborative Council (WSSCC), in particular Kristof Bostoen's water, sanitation and hygiene study in LAO PDR, funded by the World Bank.

- The Environmental Health Project's (USAID) report on Assessing Hygiene Improvement - Guidelines for Household and Community Levels.

- Boot, M.T. & Sandy Cairncross: Actions Speak - The study of hygiene behaviour in water and sanitation projects.

2.2 Core survey instrument - indicators, definitions and questions

The proposed questionnaire that has been tested in Ethiopia is attached as Appendix 1.

This section presents the subjects included in the questionnaire and the indicators that can be derived subsequently. A restricted set of indicators has been selected and the rationale and definitions of this set is presented.

2.2.1 Water supply

In accordance with the JMP questionnaire, reasonable access to water supply is assessed through questions on the following topics:

- main source of drinking water;
- main source and alternative source for other purposes;
- time to collect water;
- quantity of water used;
- individuals who collect water.

The term *reasonable access* has been adapted from the JMP in recognition of the need to further qualify *access* because access is a relative term that only exists as more or less/to a higher or lower degree, depending on the socio-economic and cultural context. The availability of an improved technology is not identical with accessibility, which depends on other factors such as the quality of the service, the location of the service, the reliability of the service etc. It has therefore been necessary to include generally acceptable standards or reasonable standards such as minimum quantity per person per day, minimum time spent on fetching water etc. *Reasonable access* measures the degree to which these standards are fulfilled and gives a more complete picture of access to water supply.

The main source type of drinking water provides an approximation for reasonable access understood as the availability of safe water of at least 20 litres per person per day from a source that can be reached within 15 minutes from the respondents' dwelling (approximately equivalent to one kilometre).

It is assumed by the JMP and others, for the purposes of making the surveys manageable, that certain types of technologies give an adequate supply of water, and these are accordingly defined as the improved water supply technologies. Improved water sources are protected water sources, non-improved water sources are unprotected water sources. A protected water source is constructed in a manner that prevents water from being contaminated, particularly from surface runoff (from rain, snow melt, or irrigation water). The following technologies are considered as improved:

- piped water into dwelling;
- piped water to yard/plot;
- public tap/standpipe;
- tubewell/borehole;
- protected dug well;
- protected spring;
- rainwater collection in closed containers.

Non-improved water supply facilities are:

- unprotected dug well;
- unprotected spring;
- rainwater collection in open containers;
- small-scale vendor;
- tanker-truck;
- surface water.

According to the JMP definitions, bottled water is considered non-improved because of concerns about the supplied quantity, not because of concerns about the quality. Here bottled water is treated differently. If the main source of drinking water is bottled water, the alternative source is used to define if the water source is improved. For instance in this report, if the alternative source is piped water into dwelling, the water supply is considered to be improved.

Water treatment is also included in the survey to discern whether households are treating the water before drinking.

As households frequently use alternative sources for other purposes than drinking, additional information is collected on the main and alternative water source used for bathing, washing, cleaning, etc. Questions on water transport time are included as an important determinant of access to water supply. Finally, a question on who fetches the water is included in order to obtain information about the burden from this work on women and children.

Compared to the JMP questionnaire, the proposed questionnaire has been supplemented with questions on the following topics to obtain a more complete picture:

- Handling of water during storage and for consumption: This is included as safe water from the source could be contaminated during storage and at the time of consumption if not handled properly;

- Seasonality is included because continuous access to improved water only applies if the household has access all year round.

- Reliability of water supply: Questions are included on the occurrence and duration of breakdowns. If the duration of breakdowns is excessive, i.e. more than one month in total during the last half year access is limited.

- Daily availability: Questions are included on the operation time of water supply facilities.

- Attitudes towards water services quality: This is to assess "reasonable access" in terms of the degree of the households' satisfaction with their water services such as water quality, distance, and stability of the services, etc.

- Expenditures on water, water treatment and water storage

Experience gained from the trial of the WASH survey for water, sanitation and hygiene practices referred to in section 2.1 suggests that enumerators have difficulties in understanding and explaining the different water supply and sanitation technologies. Therefore visual aids (picture cards with drawings of the different technologies) are used for training purposes and to support the enumerators during field work. The visual aids are presented in Appendix 4[1].

Questions are combined with direct observations, as observations often achieve more accurate results, especially in the case of delicate issues such as cleanliness, hygiene, etc. "Water handling" is also assessed through observation, i.e. the condition and design of the storage container and water drawing devices.

Below is the set of proposed indicators. The number of indicators has been limited to cover the indicators necessary to obtain a representative picture of the situation.

[1] The visuals aids were adjusted from those used in WHO's Demographic and Health Surveys. A new set of visual aids is under development in the JMP Technical Advisory Group.

Table 2.1 Water supply indicators

Indicator	Definition	Source/ Question
Access to improved water supply		
I1 - Continuous access to improved water sources as main drinking water source	Percentage of households having access to an improved drinking water supply all year round. Improved water supply options are: Piped water into dwelling; Piped water to yard/plot; Public tap/standpipe; Tubewell/borehole; Protected dug well; Protected spring; Rainwater collection in closed container.	Q2, Q32
I2 - Access to improved water supply as the main drinking water source at least part of the year	As previous but including households with access to improved water all year round or only part of the year	Q2, Q32
I3 - Time to collect water is 30 minutes or less.	Number of households with 30 minutes or less per round trip to the water source; the duration excludes other errands undertaken but includes time for queuing..	Q6
I4 - Daily availability	Percentage of households where option 1 Daily/24 hours or 2 Daily/at certain hours applies.	Q17
I5 - Reliable supply of main drinking water source during the last six months.	Percentage of households that did not experience interruptions or experienced interruptions of less than a month during the last six months. The interruptions considered are technical interruptions and hence not interruptions due to seasonal variation.	Q19, Q20
I6 - Adequate quantity of water	Percentage of households (non piped and piped users) using 20 litres or more per capita per day.	Q9 or Q11, Q61-Q63
I7 - Reasonable access	Percentage of households with daily access to an improved water source all year round, proving at least 20 l per capita per day, within 30 minutes from the residence, and interruptions of less than one month during the last six months.	Q2, Q6, Q9, Q11, Q32, Q17, Q20, Q61-Q63
Water storage and quality		
I8 - Appropriate and hygienic drinking water storage facilities.	Percentage of households with storage facilities that were observed to meet the three following conditions: 1 Cup/dipper clean, off the ground and out of reach for children; 2 Container is covered OR 3 Container has narrow neck; 4 Container looks clean.	Q28
I9 - Satisfaction with the quality of their main drinking water services.	Percentage of households who rate all seven options "good" or "acceptable". The options are: 1 Clarity (no sediments); 2 Colour; 3 Smell; 4 Taste; 5 Healthiness; 6 Stability of service; 7 Convenience (time, distance, waiting).	Q31
I10 Water treatment within the last week.	Percentage of households that have applied one or several treatment options during the last week. The options are: 1Boiling; 2 Bleach/chlorine; 3 Sieve through cloth; 4 Water filter; 5 Let stand and settle; 6 Other	Q21, Q22, Q23
Water cost & contributions		
I11 - Average initial costs of water services/facility	Average in local currency	Q14, Q2
I12 - Average monthly costs of water services	Average in local currency	Q10 and Q12
I 13 - Average cost of storage facility	Average in local currency	Q29
I14 - Average costs of water treatment	Average in local currency	Q24

Note: The numbers referred to in the definition is the option numbers in the relevant question, cf. Appendix 1

2.2.2 Sanitation

In line with the JMP core questionnaire, "access to an adequate sanitation facility" is assessed through two questions:

- available sanitation technology;
- shared/non-shared facilities.

Knowing the type of sanitation facility used by a household provides information on the degree to which the facility meets basic sanitation requirements, i.e. if it hygienically separates human excreta from human contact. Meeting basic sanitation requirements is possible through the following facilities:

- flush to piped sewer system;
- flush to septic tank;
- flush/pour flush to pit;
- composting toilet;
- VIP/pit latrine with slab.

Facilities that do not meet basic sanitation requirements include:

- flush/pour flush to elsewhere;
- pit latrine without slab/open pit;
- bucket toilet;
- hanging toilet;
- open air/bush/field etc.

Whether a sanitation facility is shared with other households or is a public facility can impact its hygienic quality and in turn have an impact on its use. Shared facilities that are poorly maintained may be difficult for household members to use because of filth, odour and flies. Shared facilities - either private or public- are not considered adequate.

The JMP questions focus only on the type of available facilities. In the proposed questionnaire, additional indicators are included that measure the actual use of the facilities, cf. Appendix 1. Actual use is assessed by the presence of factors that are likely to encourage or discourage their use. The following are included:

- time to reach the facility;
- free access, signs of regular use;
- hygienic condition of facility.

If it takes a long time to reach the facility - if the facility is some distance away from the place of residence - this may discourage its use, e.g. by children or during the night. Whether a facility is actually used or not can be assessed by looking at certain conditions such as free passage (path is cleared, entrance not blocked, etc) and other signs of use. The hygienic condition of a facility is an indication of its use - unhygienic conditions are more likely to discourage its use. At the same time, the condition of the facility is an indication of hygiene practices.

Picture cards are used to assist the enumerators to identify/validate the type of sanitation available to a household. Also, questions are combined with observations. The time it takes to reach the facility, the accessibility (indication of actual use) of the facility and its hygiene condition are assessed through observation.

The costs of building the sanitation facility and the costs of maintenance and repair are also assessed, because costs are important aspects.

The proposed set of indicators regarding sanitation is given in the table below.

Table 2.2 Sanitation indicators

Indicator	Definition	Source/ Question
Access to improved sanitation facility		
I15 - Access to improved sanitation facilities	Percentage of households that have and use an improved and private facility. Improved options are: 1 Flush to piped sewer system; 2 Flush to septic tank; 3 Flush/pour flush to pit; 5 Composting toilet; 6 VIP/pit latrine with slab.	Q51, Q53
I16 - Access and actual use of improved sanitation facility	Percentage of households with improved services and where none of the below observation options apply. The observations are: Dense vegetation in front of the toilet; Waste or debris on the path; Major crevice or potholes on path; Entrance to toilet obstructed.	Q55
I17 - Sanitation facility within less than 5 min. from household	Percentage of households with sanitation facility which can be reached within five minutes' walk from the household.	Q54
Quality of sanitation		
I18 - Improved sanitation facility in good hygienic condition	Percentage of households with improved sanitation where none of the observation options apply. The options are: 1 Visible faecal residues in and around the drop whole or the basin; 2 Visible faecal residues on the floor, wall or door; 3 Visible used anal cleansing material; 4 Surface flow of sewage; 5 The toilet smells bad/stinks.	Q56
I19 - Satisfaction with the quality of the sanitation facility	Percentage of households that rate all six options "good" or "acceptable". The six options are: 1 Effective operation; 2 Convenience; 3 Cleanliness; 4 Smell; 5 Distance to toilet; 6 Ease of repair/maintenance.	Q60
Cost of sanitation facility		
I20 - Average cost for building facility	Average in local currency	Q59, Q51
I21 - Average annual costs for maintenance and repair	Average in local currency	Q58, Q51

Note: The numbers referred to in the definition is the option numbers in the relevant question, cf. Appendix 1

2.2.3 Grey water discharge

Grey water consists of domestic wastewater exclusive of toilet waste. The rationale for including grey water is that from the point of view of "water hygiene" wastewater disposal is an integral part of the "water handling cycle" comprising drawing of water, storage, handling for consumption, reuse, and wastewater disposal.

Water used for cleaning clothes and nappies can be heavily contaminated with the same disease-causing organisms that hygiene and sanitation are intended to control. Stagnant water resulting from insufficiently drained grey water could also become breeding grounds for mosquitoes. Grey water therefore needs to be disposed of properly.

The following observations are included to assess grey water disposal:

- mode of discharge of the household's wastewater;
- hygiene condition of the point of discharge.

Table 2.3 Grey water disposal indicator

Indicator	Definition	Source/ Question
Access improved grey water disposal		
I22 - Access to improved grey water disposal methods.	Percentage of households that use an improved grey water disposal method. The options are: 1 Piped sewer; 2 Soak away/cesspit/septic system; 3 Sanitation facility; 4 Open channel	Q39

Note: The numbers referred to in the definition is the option numbers in the relevant question, cf. Appendix 1

2.2.4 Hygiene practices

Three questions are used to assess hygiene practices. They are concerned with:

- access to hand washing supplies;
- use of hand washing supplies at critical times;
- disposal of child faeces.

Hand washing is one of the most important preventive measures for reducing the prevalence of diarrhoeal disease. Whether a household has access to hand washing supplies such as soap can provide an indication of whether appropriate hand washing practices within the household are possible or likely. Along with appropriate hand washing supplies, hand washing *at critical times* is an important element of appropriate hand washing practices.

Knowing how children's faeces are disposed of is a critical aspect of hygiene improvement. Particular focus is placed on children's faeces because their faeces are more likely to be a source of contamination in the household environment than the faeces of other household members - and because in many societies they are considered harmless and therefore are not disposed of properly.

The JMP questions focus on the availability of soap alone, but in the proposed questionnaire soap alone is not considered a sufficient indication of appropriate hand washing practices, as soap may be used for other purposes. In many places/countries, no distinction is made between soap for hand washing and soap for other uses such as washing dishes, washing clothes, cleaning, etc.

It appears more useful to look at the entire "hand washing system". This is also supported by the WASH and EHP surveys, which find that the observed (concurrent) presence of items needed for hand washing, i.e. soap, water and basin, works best as an indicator for "good hygiene practice". For this reason, an observation is included of people's hand washing practices (they are asked to demonstrate how they usually wash their hands) to determine if the three prerequisites are present.

A fourth indicator of hand washing - the presence of a clean towel or cloth[2] - is not used, as assessing cleanliness through the enumerators' judgement is very subjective while at the same time towels or cloths that are not clean constitute a health hazard.

Table 2.4 Hygiene behaviour indicators

Indicator	Definition	Source/ Question
Hygiene behaviour		
I23 - Availability of soap in the household	Percentage of households where soap (or ash, sand or detergent) is availability and observed by the enumerator.	Q43, Q44,
I24 - Adoption of appropriate hand washing practices	Percentage of households where availability of three crucial preconditions for hand washing has been observed: 1 Soap; 2 Water for hand washing; 3 Facility for hand washing.	Q42
I25 - Use of soap at critical times	Percentage of households reporting to have used soap after defecation and at least two more critical times. The most critical times are: 5 Washing hands after defecation; 6 Washing hands after cleaning child; 7 Washing hands before feeding children; 8 Washing hands before preparing food; 9 Washing hands before eating.	Q46
I26 - Awareness of children's hand washing practices	Percentage of households that mention two critical times for hand washing by children. Critical times are: "before eating" and "after defecation".	Q47
I27 - Properly disposal of children's faeces	Percentage of households who report using one of three proper methods to dispose of children's faeces (1 Dropped into toilet facility; 2 Washed away into toilet facility; 9 Buried).	Q50
I28 - Clean surroundings around residence	Percentage of households with surroundings free of faecal matter.	Q72

Note: The numbers referred to in the definition is the option numbers in the relevant question, cf. Appendix 1

2.2.5 Household characteristics

In addition to the core questions on water, sanitation, drainage and hygiene a number of standard questions on household characteristics are included to establish the households' socio-economic situation. Although this is not the main

[2] As suggested by the EHP questionnaire.

objective of the present study, this may allow for identification of the relationship between poverty/well-being levels and water, sanitation and hygiene conditions.

3 Sampling methodology

The purpose of the sampling methodology is to devise a method that will provide reliable data at low cost. Emphasis is placed on obtaining accurate information for policy decisions that can be used at the local level.

The selection of the sample is central in any survey, as only a representative sample will reflect the situation in the area. As an example, it is of no use to have an estimate of households with improved water services with a narrow confidence interval if the sample selected in the survey is biased.

When appropriately used and carefully carried out, the proposed sample selection method results in precise estimates of the water, sanitation and hygiene indicators. The proposed sampling methodology and its applicability to various field settings is presented in the chapter.

A number of abbreviations are used through out the chapter and therefore explained here:

- Primary sampling unit: PSU
- Basic sampling unit: BSU

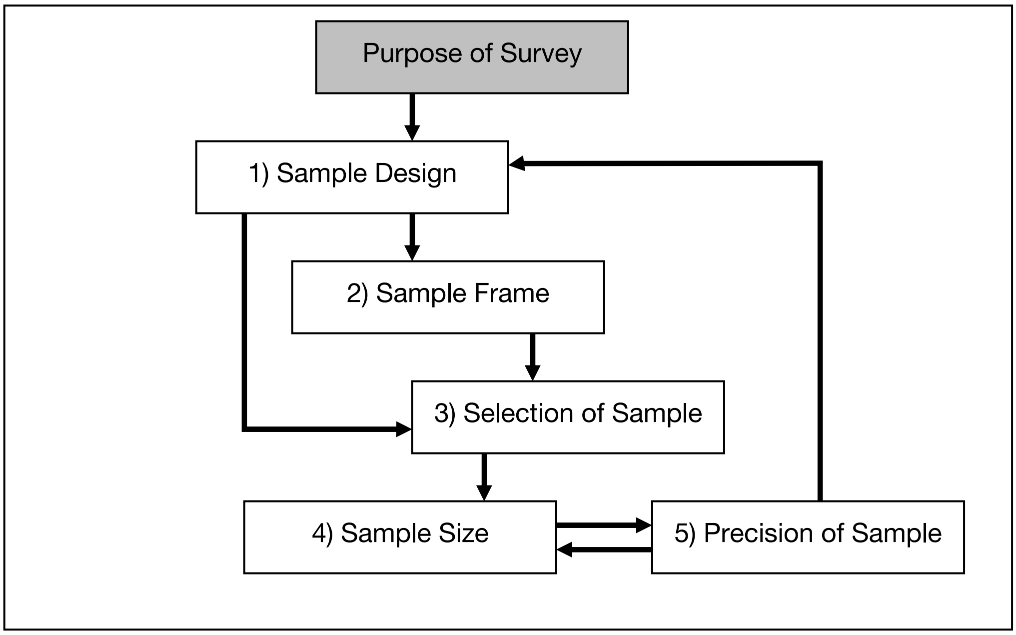
In a two-stage cluster sampling procedure where a number of towns and a number of household in each town are selected for interviews, the PSU is the town and BSU is the household. Using the cluster sampling method, the towns are selected so that the probability that a town is included in the sample is proportional to the size of the town. Hence the towns are selected with

- Probability proportionate to size: PPS

3.1 Overview

The proposed sampling methodology consists of five interdependent steps. The aim of these steps is to define the sample design along with the resulting precision of the sample, cf. Figure 3.1.

Figure 3.1 Overview of sampling methodology



After identifying the scope of the survey, which comprises the areas and population to include, the methodology contains the following steps:

1) Selection of sample design; overall methodological approach for instance random sample selection or two or multi-stage cluster sample selection.

2) Identification of sample frame; identification of data about the population living in the area of the survey.

3) Selection of sample; regions, communities, households included in the survey.

4) Determination of the sample size; number of observation in the survey

5) Calculation of the expected precision of estimates of indicators; estimation of the expected standard deviation of the indicators. The standard deviations of the indicators are mutually dependent on the sample design and the sample size, cf. the upward arrows in the figure above.

In the following sections each of these five steps are described in detail. The description consists of the proposed sampling methodology as well as a description of the practical use of the applied methodology in Ethiopia. The practical applications are shown in boxes.

3.2 Purpose of survey

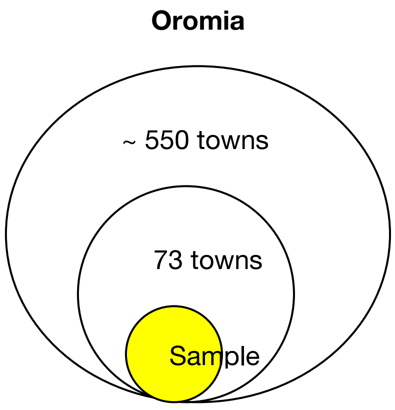
The purpose of the survey determines to a great extent the sampling strategy, so it is therefore of vital importance to set out clearly the aims of the investigation.

Box 3.1 Aim of investigation in Ethiopia

The regional planning office in Oromia and the World Bank decided that the survey should be carried out in urban areas in the Oromia region.

In this region there are approximately 550 towns and out of these 73 towns were chosen as candidates for possible water and sanitation projects.

Thus the population from which the sample should be taken is the 73 towns and not the entire population of towns in Oromia, cf. the inner circle in the figure.



3.3 Sample design

The two- or multiple stage cluster sampling procedure is proposed for household surveys on water and sanitation. The term cluster refers to a natural grouping within the population, such as a neighbourhood, town, district or other community, from which a sub-sample may be selected.

This procedure is based on the selection of a certain number of these clusters, i.e. primary sampling units (PSUs) with a probability proportional to size. Within each of these PSUs a fixed number of households or basic sampling units (BSUs) is selected, either with equal selection probability or - if possible - based on stratification.

The selection of towns with a probability proportional to size (PPS) is carried out by creating a cumulative list of town populations and selecting a systematic sample from a random start. An example of the practical implementation of PPS is given in section 3.5.

Each BSU in the population will have an equal probability of being in the sample. Such a sampling procedure is said to be self-weighting and leads to simplified formulas for analysis.

The selection of the sample may be done in two- or multi-stages. For example a country may be split into regions, a number of districts chosen from each region, a few communities from each district, and a number of households from each community. However, the basic principles for deciding sample size and structure and the methods for estimating rates and their standard deviation are the same.

The basic sample selection procedure is simple random sampling for instance by random selection of households from a complete list. However, this procedure is usually not applied due to lack of such detailed information and the high costs of travelling between the often remotely located households.

Though relatively seldom used in practice for household surveys, simple random sampling is an important baseline for evaluating other sample procedures such as two- or multiple-stage cluster sampling.

Box 3.2 Sample design in Ethiopia

In Ethiopia, the PSUs are the towns in the Oromia region and the BSUs are the households located in these towns. The method used is two-stage cluster sampling, where the clusters - towns - are selected at the first stage and the households within the clusters are selected at the second stage.

3.4 Sample frame

The choice of the cluster sample procedure means that a sample frame is required. A sample frame is a list of the units from which samples are to be selected, along with the unit-specific information required to select a sample from this set of units. For example, a sample frame for the primary sampling unit (PSU) could be a list of all the towns in the region with their populations. Similarly, a sample frame for the basic sampling unit (BSU) could be a list of all households in the town. The purpose of a sampling frame is to enable the selection of PSUs and - if possible - also BSUs within each PSU.

The sample frame can be defined on the basis of existing information such as a census, census background material, relevant surveys or administrative records. Usually, such information will be available, but the application must depend upon the reliability, coverage and age of the data.

3.4.1 Sample frame for PSUs

In most cases the quality of existing data will allow the selection of PSUs. For example, if PSUs are selected from a number of towns and the towns are selected with probability proportionate to size, then the sampling frame will consist of a list of the towns with their population sizes. If data on town population sizes are not available, it will not be possible to carry out sampling with a probability proportional to size. This will necessitate a count of the total number of households in each selected town and thereby increase the costs of the survey considerably. As a less costly alternative, an estimate of the number of inhabitants can be used.

3.4.2 Sample frame for BSUs

The ideal frame for the selection of households, BSUs, would be a list containing all households in each town that allows the analysts to choose a random selection from the list. If such a list does not exist - which it usually does not - and if the PSU is small, then a list can be created by carrying out a quick census.

Instead of listing, a sample frame can be created from enumeration maps, which are almost always available because they are produced for population censuses. Enumeration maps are demarcated geographical areas, each of them consisting of approximately 200 households. The constant number of households means that the maps vary in size and are small in densely populated areas and large in sparsely populated areas. Enumeration maps can typically be obtained from the national statistical office.

A town with 20,000 inhabitants could for instance have 20 enumeration maps. These maps define a sample frame from where it is possible to a) randomly select a number of demarcated geographical areas from the town and b) randomly select households (BSUs) from each demarcated area. Even enumeration maps that are 5-10 years old can be used as a sample frame unless exceptional changes such as resettlement have been taking place.

The use of enumeration maps means that the enumerator should choose the households to interview by an ad hoc random selection method. If it is not possible to establish a sample frame, a way has to be found which ensure that the sample is as representative as possible. These methods are described in the next section.

Box 3.3 Sample frame in Ethiopia

In Ethiopia, a sample frame was defined for PSUs as well as BSUs.

For PSUs, the 73 towns in Oromia that were selected as candidates for water and sanitation improvement constituted the population from which the PSUs is selected**.** A simple sample frame with information about the population in these 73 towns in 2004 was available and applied as the basis for a PPS sample selection.

For BSUs, enumeration maps from population censuses were used as the sample frame. The maps were from 1994 and it was therefore expected that, to some degree, they were inaccurate. In order to determine how inaccurate they were, a linear regression analysis was carried out with the number of inhabitants in the towns in 2004 as the response variable and the number of enumeration maps per town as the explanatory variable.

As each enumeration map includes around 200 households, the total number of maps per town would indicate the total number of inhabitants. The regression analysis showed that the number of enumeration maps in 1994 is highly correlated with the number of inhabitants in 2004 (R^2=0.98) and, more importantly, that there were no extreme observations (no standadized residual above 2). This means that the population size increased consistently in the towns over the period.

The advantage of using enumeration maps is that they clearly limit a small geographical area from which it is easy to take a random selection of households. In addition, it is straightforward to select the maps randomly from each town in order to get a representative sample of geographical areas in each town. The practical application of the enumeration maps is described in more detail in section 0.

3.5 Selection of sample

In a large region or country where an overall estimate is required, it will usually be sensible to select the sample of PSUs in at least two stages. For example, if the region is split into a number of administrative units, one would select a sample of districts by the PPS method. Within each selected unit, communities would be selected, again by the PPS method. The same number of communities must be selected in each unit. Households within the communities would be selected in the usual way, again with the same number selected in each community.

3.5.1 Selection of PSUs

The two-stage cluster sample procedure requires a list of all the PSUs in the area where the survey is to take place in order to enable the sample selection. In addition, some approximate measure of the number of BSUs in each PSU is also necessary. If it the average size of households does not vary significantly from one PSU to another, a general measure of the PSU population size will be sufficient.

The relative size of the PSU is more important than the absolute size, so even an out-of-date census will be adequate if allowance is made for known variations in population growth rate. If some PSUs are too small to provide an adequate sample of BSUs, they should be combined with neighbouring PSUs before making the list.

Random selection of PSU by probability proportionate to the size of PSUs is carried out by creating a cumulative list of PSU populations and systematically selecting a sample from a random starting point. Further, it is recommended to select a constant number of BSUs within each PSU, which will imply that each BSU in the population will have an equal probability of being in the sample. This results in a self-weighted sample and simplified formulas for analysis. In Box 3.4 the selection of a PPS sample in Ethiopia is described.

It should be noted that in selecting a PPS sample it is possible for the same PSU to be selected twice, if the PSU has a population greater than the sampling interval. This can happen if a PSU is much larger than the other PSUs. If this should be the case, the correct procedure to follow would be to select two sub-samples of BSUs from within the PSU. It is *not* appropriate to select another PSU instead, or to repeat the whole sampling procedure until no PSUs are repeated, as either of these approaches invalidates the required probabilities.

Box 3.4 Selection of sample in Ethiopia, selection of PSUs

In Ethiopia, 1,000 household were selected from the 73 towns chosen as candidates for water and sanitation improvements. Using random selection with probanility proportional to size (PPS), 25 towns (PSUs) were selected for the sample. The choice of 25 town was based on a trade-off between expected precision of the indicators and travel cost of the enumerators. The relationship between precision, number of clusters and sample size is further described in section 3.6.

The PPS is carried out by listing the towns and their populations, then calculating the cumulative population.

The towns were not listed in any particular order. An implicit stratification by size of the town, for instance, could be obtained if the towns are ordered by size because the selection procedure implies that the selected towns will be spread over the list. Such an ordering is called implicit stratification because a kind of stratification is obtained but it is not taken into account explicitly when calculating standard deviations of the indicators. The advantage in the case of implicit stratification by size would be that all sizes of towns would be represented in the sample.

With 25 towns selected from a total population of 73, the first step of the PPS is to divide the total population of the 73 towns (797,065) by the number of towns to be selected (25) to obtain the sampling interval (797,065/25 = 31,883).

Then a random number between 1 and 31,883 is chosen. Suppose this number is 12,179. This number is then inserted in the first row in the list to identify the first town in the sample. Since 12,179 lies between 10,093 and 13,062, town 2 is chosen. Now the sampling interval is added to the initial random number: 12,179 + 31,883 = 44,062 and so town 3 is chosen.

Then add the sampling interval again: 44,062 + 31,883 = 75,944 and town 7 is chosen. The procedure continues until 25 towns have been selected.

Town	Population	Cumulative population	Cumulative sampling interval
1	10,093	10,093	12,179
2	2,969	13,062	12,179
3	34,136	47,198	44,062
4	6,394	53,592	75,944
5	16,186	69,778	75,944
6	2,103	71,881	75,944
7	24,542	96,423	75,944
...	…	…	…
73	6859	797.065	809.244

In order to obtain a self-weighted sample, a constant number of households is selected within each of the 25 towns. A total number of household interviews of 1,000 means 40 households within each town.

3.5.2 Selection of BSUs

In section 3.4 it was argued that the ideal procedure for selection of BSUs would be to have a list of all BSUs in each PSU and to choose from the list at random. However, such a list does not often exist, and alternative sources of information about the population in the BSU may be obtained through other sources such as a quick census or a consultation of community leaders. Regardless of the source of information, it is crucial that the local parties such as community leaders only provide information about the population and have no influence on the selection of the sample. A cross-checking may be necessary, and if the information received is considered to be biased, a random selection procedure with not sample frame for BSU is preferable.

Enumeration maps - used as the basis for population censuses - can be considered as an alternative approach. Using enumerations maps requires the selection of households to interview from each enumeration map. This will usually involve two stages: a method of selecting one household to be the starting point, and a procedure for selecting successive BSUs randomly after that.

The first BSU to interview can be on the border of the enumeration map, and the subsequent BSUs can be selected by one of a variety of random selection procedures e.g. flipping a coin, spinning a bottle or choosing every fifth BSU for interview. The practical application of enumeration maps is described in Box 3.5.

If neither a listing nor enumeration maps are available, a strategy could be to choose a central point in the PSU, such as the market if it is a town; then to choose a random direction from that point, count the number of households between the central point and the edge of town in that direction, and select one of these houses at random as the starting point of the survey. The remaining households in the sample should be selected to give as widespread coverage as possible of the community[3].

In this connection it is important to take into consideration variations in the concentration of households. Thus in areas where buildings or compounds have more than one household, attention should be directed towards preventing households in such multi-household dwellings from being underrepresented. If most dwellings contain more than one household, as, for example, is common in some parts of Africa, then such compounds may be treated as a cluster and multistage sampling could be applied.

In large PSUs, the selected sample can have more starting points in different parts of the PSU. This would also reduce the under-representation of BSU in the outer parts of the PSU inherent in having just one central starting point. An example of this approach is described in further detail in the Box below.

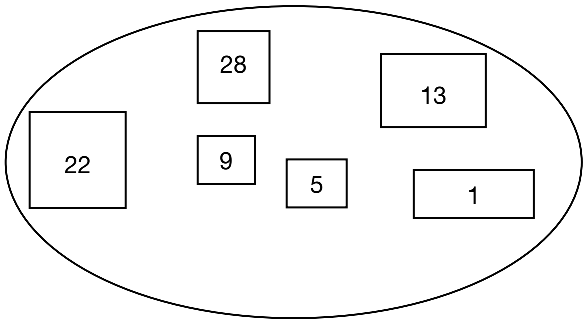
[3] A brief description of the random walk/route approach can be found in Crouch, S. and Housden, M (1996) Marketing research for managers. 2nd edition. Oxford: Butterworth Heinemann.

Box 3.5 Selection of sample in Ethiopia, selection of BSUs

The selection of households within each of the 25 PSUs is carried out using enumeration maps. In order to assure that the households interviewed are spread over the towns, 6-8 enumeration maps were chosen for each town depending on the size of the town. The enumeration maps were randomly selected for each town by the use of a stochastic number generator.

For a town with 30,000 inhabitants and around 30 enumeration maps, the random selection procedure implies the generation of 6 random numbers between 1 and 30. These numbers determine which enumeration maps are used in the subsequent selection of the households. An example of such a selection of enumeration maps is illustrated below.

Selection of enumeration map areas in a town



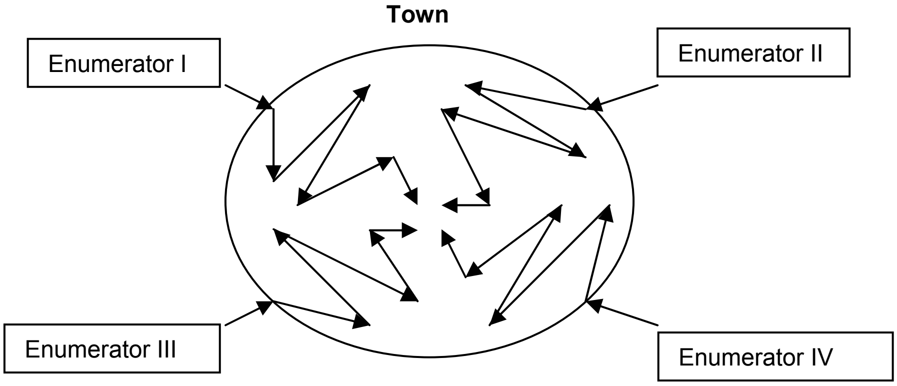
The random selection of enumeration maps generally ensures that the maps are dispersed randomly across each town and that a representative sample is drawn from each town. As illustrated in the figure, the enumeration maps vary in size according to the population density and will typically be larger in the outer areas than in the more densely populated centres of the towns.

The selection of the 5-7 households within each enumeration map area is carried out using one of the techniques for random selection described in the previous section. Whether or not the enumerator has used these techniques properly is easily checked, as interviewed households are marked on each map by the enumerator.

Box 3.6 Selection of sample in Ethiopia, selection of BSUs (continued)

In towns with less than 6,000 inhabitants - and less than 6 enumeration maps - it was decided that the enumeration maps were too large to work in practice, i.e. the enumerators have difficulties using maps that measures 1x1 meter.

As an alternative strategy of selecting households in a random manner, the enumerators may identify four points in the outer area of these towns where each enumerator begins his or her rounds of interview, cf. illustration below.



From these starting points each enumerator makes his or her way towards the centre of the town while randomly selecting households. The random selection could be done by for example spinning a bottle or by closing the eyes and turning the body.

Each enumerator is free to choose his or her strategy for random selection as long as the final pattern of selection is similar to the pattern illustrated in the above figure. Rather than instructing the enumerators to use a specific rule to obtain a randomly selected sample, the concept of randomness was carefully explained along with different ways of making a random selection (flipping a coin, spinning a bottle, taking every third household).

3.6 Precision of the sample

The precision of the sample depends on

1 The size of the sample;

2 The number of clusters in the sample: and

3 The size of the variable whose value is being measured.

The size of the population from which the sample is selected has little effect in practice.

In the following sections each of these factors will be explained in further detail. In addition, a method to assess the precision of a sample before the survey is carried out - a precision forecast - is described.

3.6.1 Determinants of precision

The design of the sample has great influence on the precision of the results and the cost of the survey. If sufficient resources are available for the project, the ideal approach is to decide on a certain level of precision and then design the required sample accordingly.

The cluster procedure is used to reduce the cost of the data collection. The implication of the cluster procedure as compared to the simple random selection is larger confidence intervals and thereby a lower precision than in a simple random sample. This is because the repeated draws from the same PSU will result in BSUs that have similar characteristics. The cluster method will therefore reflect the population's diversity to a lesser extent than a simple random method. On the other hand, for a constant sample size, the precision increases with the number of clusters.

The design effect is the loss of effectiveness by the use of cluster sampling instead of simple random sampling. The design effect is basically measured as the ratio of the actual variance under the cluster sampling method to the variance under the assumption of simple random sampling.

The design effect increases with the number of BSUs selected in each PSU. Thus, for a fixed total sample size, a design with more PSUs and fewer BSUs in each PSU will provide more precise estimates of sample statistics than a design with fewer PSUs and more households in each PSU.

The size of the variable measured by the survey and its distribution across the population also influence the precision of the sample.

For example, suppose the overall (unknown) proportion of households with access to improved sanitation were 40%. If the proportions in each community in the region varied very little (say from 35% to 45%), then a small number of clusters would give a reasonably precise estimate. If, on the other hand, the proportions in each community varied more widely (say from 0% to 80%), one would need a considerably larger sample to be sure of obtaining the same precision. This variability is measured by the rate of homogeneity (σ), which will be discussed in detail below.

3.6.2 Estimation of the precision of the sample

An assessment of the precision of the sample should be made before the survey is carried out. This assessment will indicate how many clusters and observations are needed in order to obtain a certain level of precision or what level of precision can be obtained if the number of clusters and the sample size are dictated by limited resources.

The precision of the sample can be measured by the standard error of the variable that is measured by the survey. The usual way to use the standard error is to construct a 95% confidence interval for the true value of the variable. The confidence interval defines an interval within which there is 95% certainty of finding the true value of the variable. In practical terms, the confidence interval is defined as the mean value of the variable plus/minus two standard errors.

For a cluster sample method, the standard error of an estimated proportion p - for example the proportion of households' access to improved sanitation - can be defined as:

Equation 3.1 $$Std.err = \sqrt{\frac{p(1-p)D}{n-1}}$$

where D is the design effect and n is the number of observations. Equation 3.1 is an extension of the simple formula used when the data are assumed to come from a simple random sample, the binomial formula:

Equation 3.2 $$Std.err = \sqrt{\frac{p(1-p)}{n-1}}$$

The value of the square root of D measures the increase in the standard error of the estimate due to the application of the cluster sample method. D is - as described above - the loss of effectiveness by the use of cluster sampling. The design effect is defined as:

Equation 3.3 $$D = 1 + (b-1)\sigma$$

σ is the rate of homogeneity, a measure of the variability between clusters as compared to the variation within clusters while b is the average number of responses per cluster. The value of D (or equivalently of σ) will be estimated in the light of experience gained from previous surveys of similar design and subject matter.

The design effect is conventionally set at two when calculating the sample size, cf. Bennet et al. (1991)[4]. Based on the data collected in Ethiopia, the design effect is estimated to be 3 for the question on access to improved water services and 4 for access to improved sanitation, cf. table below. Hence these levels of designs effects of 3-4 are proposed for future estimations on sample precision.

[4] A list of literature is found in Appendix 6

Table 3.1 Design effects and rate of homogeneity

Source	Design effect	Rate of homogeneity
Ethiopia (this survey)	3-4	0.08-0.13
Bennet, S *et al*	2	0.2
Kristof Bostoen *et al*	7	0.19

Note: For Ethiopia, the design effect was estimated using *Equation 3.9*, the rate of homogeneity was calculated using *Equation 3.10*. The two numbers were calculated both for access to drinking water and access to sanitation which is the reason for the interval shown in this table.
Source: The design effect and rate of homogeneity for Ethiopia are calculated on the basis of the data collected in this project. Other sources cf. foodnote references in Appendix 6.

Other studies have found even higher design effects including Kristof Bostoen *et al* where the design effect for access to water and sanitation after stratification is estimated to be around 7 and around 13 when no stratification is used (i.e urban and rural households were included in the same sample)

Such high design effects mean that a cluster sampling procedure in some cases needs significantly more observations than simple random sampling in order to achieve the same level of precision - or equivalently, for a constant sample size, that the number of clusters needs to be increased.

The rate of homogeneity, σ, is a measure of the variability between clusters as compared to the variation within clusters. In a two-stage cluster sample σ is equivalent to the "intra-cluster correlation". In a more complex design such as a stratified multistage survey, σ is composed of the components of variability from all stages of the design.

The value of σ will be higher for variables whose values vary more between clusters than within them. For example, because families in the same area tend to have broadly similar water and sanitation standard, variables such as access to improved water supply and access to improved sanitation facilities will be more likely to produce the same response for two individuals in the same cluster than for individuals in separate clusters. These variables are often assumed to have a relatively high value of σ, around 0.20, cf. the results from Bennet et al in the table above.

In the survey in Ethiopia, the rate of homogeneity is estimated to be 0.08 for access to improved water supply and 0.13 for access to improved sanitation facilities, indicating a lower variance between the towns than seen in other studies.

It should be noted that the rate of homogeneity depends on the variable measured and should be reported for each main indicator. For instance, in the case of urban areas in Ethiopia, the rate of homogeneity is higher for water supply than for access to sanitation meaning that the difference in water supply between towns is less than the difference in sanitation facilities.

Number of clusters

The design effect, the rate of homogeneity and the number of clusters in a sample are closely connected and have significant influence on the precision of the sample, i.e. the standard errors of the measured variables in the sample. The relationship is illustrated using Equation 3.1 and Equation 3.3, and the result is presented in Table 3.2.

The assumptions are the following:

- Sample size: 1,000,
- Rate of homogeneity: 0.1,
- Expected level of indicator: 0.4 (eg. 40% has access to improved drinking water)

Increasing the number of clusters from 10 to 80 will decrease the design effect from 21 to 3. In addition, the precision of the sample will increase, as the standard error decreases from 0.051 to 0.023.

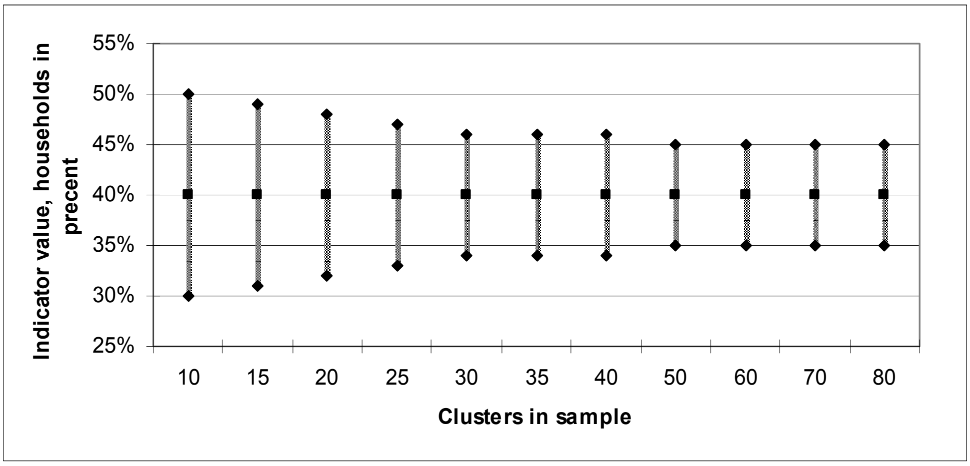
Table 3.2 Relationship between number of clusters and expected standard error, for given rate of homogeneity (0.1) and number of interviews (1,000)

Clusters	Interviews per cluster	Design effect (D)	Variable (p)	Standard error (s)	95 pct confidence interval
10	100	21	0.40	0,051	0,30-0,50
15	67	14	0.40	0,043	0,31-0,49
20	50	11	0.40	0,038	0,32-0,48
25	40	9	0.40	0,034	0,33-0,47
30	33	7	0.40	0,032	0,34-0,46
35	29	7	0.40	0,030	0,34-0,46
40	25	6	0.40	0,029	0,34-0,46
50	20	5	0.40	0,026	0,35-0,45
60	17	4	0.40	0,025	0,35-0,45
70	14	4	0.40	0,024	0,35-0,45
80	13	3	0.40	0,023	0,35-0,45

The decrease in the standard error implies that the 95% confidence interval is narrowed down from the interval of 0.30-0.50 to 0.35-0.45. Thus, on the basis of estimates of the rate of homogeneity and the expected mean value of the variable measured by the survey, it is possible to assess what kind of precision the survey will generate.

The narrowing of the confidence interval is further illustrated in Figure 3.2. The figure indicates that narrowing of the confidence interval becomes more pronounced when the number of clusters increases from 10 to 20 than from 30 to 40 and higher. The marginal gain in precision of increasing the number of clusters and thus the costs of the survey are not constant. There is a greater marginal gain in precision when the number of cluster is increased from a low level than from a high level.

Figure 3.2 Confidence interval for increasing number of clusters



It should be noted that if the frequency of the variable measured by the survey is expected to be quite low, for example 2%, it is not sensible to design a survey to achieve an absolute precision of 5%. In such a case, the standard error desired needs to be considered relative to the expected prevalence rate, and would be much smaller, say 0.5%.

Box 3.7 Number of primary sampling units (PSUs) in Oromia

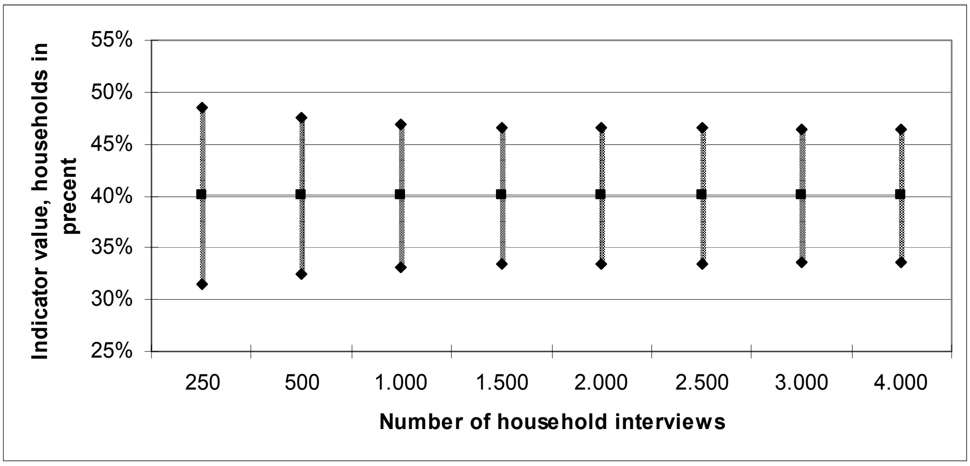
In Ethiopia, the number of primary sampling units was decided on the basis of the available resources and the marginal gain in precision of increasing the number of clusters compared to the additional costs of increasing the number of clusters.

As the sample is to be decided before the estimates of design effect and rate of homogeneity is available, the rate of homogeneity in Bennet et al. of 0.20 was assumed. It was decided that the survey should consist of 25 towns (clusters) with 40 interviewed households in each town. This was decided as a trade of between transportation cost and marginal precision of the estimates.

For a variable with an expected proportion - or frequency - of 40% in the population, this means a standard error of 4.6% and a 95% confidence interval between 31.0 - 49.0%. As it was seen in Table 3.2, the estimated rate of homogeneity in Ethiopia turned out to be lower than expected, with means that the precision of the indicators is ±7% rather than ±9%..

Number of observations
Similarly, the number of observations needed to obtain a certain level of precision can be analysed, cf. the figure below where the confidence intervals illustrated depending on the number of observations in the sample.

Figure 3.3 Confidence interval for increasing number of observations



Note: The number of clusters is assumed to be 25, the remaining parameters are as in Figure 3.2

Comparing the decrease in the standard error and, hence, also the narrowing of the confidence intervals in the two figures above, it can be concluded that the gain in precision in the overall estimate (*p*) is limited, as the number of interviews are increased.

If a certain precision of the estimate households with access to improved water supply is needed, emphasis should be on the number of clusters rather than on the number of interviews. For instance if the results are needed with a precision of ±5%, this can be obtained with 1,000 interviews distributed on 60 clusters. As seen in the figure above such precision cannot be obtained with 25 clusters even with 4,000 interviews.

The above graphs should be used as guidance for decisions regarding sample size and precision before the survey is carried out. Prior estimates of the rate of homogeneity should be obtained from other studies in similar areas in order to be able to estimate the expected level of precision and hence the necessary sample size and number of cluster.

3.7 Analysis of the data

3.7.1 Introduction

The analysis of the sample is carried out by estimating the means of the indicators measured and the precision of these variables in terms of standard errors and the associated 95% confidence intervals.

The statistical analysis can be carried out in any statistical package or in a spreadsheet.

3.7.2 Notation

The formulas for the estimation of the mean and variance operate with the following notations:

- m = The number of clusters (PSUs) in the sample;
- M = The overall population from where the sample is selected;
- n_k = The number of BSUs per cluster (PSU);
- N = Total number of inhabitants in the population (M);
- N_k = BSU population, $k = 1,.., M$.

These sample characteristics for Ethiopia are given in the Box below.

Box 3.8 Sample characteristics in Ethiopia

In Ethiopia, a sample consisting of 25 towns are selected from a population of 73 towns. In addition, in each town a sample of 40 households are selected. The total population in the 73 towns is N, while the population in the individual town is N_k, $k = 1,..,M$. Hence, the following notations:

- *m = 25 towns;*
- *M = 73 towns;*
- *n_k = 40 interviewed households per town;*
- *N = Total population of the 73 towns: 797,065;*
- *N_k = Individual town population.*

3.7.3 Estimation of mean and variance

For each PSU in the population the mean is estimated by:

Equation 3.4 $$\bar{y}_k = \frac{1}{n_k} \sum_{i=1}^{n_i} y_{ki}$$

As a starting point, the first case is where $m = M$, i.e. where all PSUs in the population are investigated.

The use of a sampling procedure where the probability of selecting a town is proportional to the size of that town (PPS) implies that the sample is self-weighting, which means that the overall average is an unbiased mean of the overall population average. The mean is estimated by:

Equation 3.5 $$\bar{y}_{Sample} = \frac{1}{m} \sum_{k=1}^{m} \frac{1}{n_k} \sum_{i=1}^{n_k} y_{ki} = \frac{1}{n} \sum_{1=1}^{n} y_{ki}$$

The variance of this estimator is relatively simple due to the self-weighing qualities of the sample[5]:

Equation 3.6 $$var\left(\bar{y}_{Sample}\right) = \frac{1}{m(m-1)}\sum_{k=1}^{m}\left(\bar{y}_k - \bar{y}_{Sample}\right)^2$$

This formula requires that the sample of *m* towns is randomly selected with PPS with replacement. Replacement means that there is a possibility that towns are represented twice or even more often in the sample. As the selection is made with PPS, this is most likely to happen in large towns.

This means that large towns have a similarly large weight in the sample, which tends to reduce the variance and gives the self-weighting feature of the estimator. This feature of PPS with replacement is in contrast to the case of simple random sampling in which replacement, i.e. the possibility that a unit is drawn twice, increases the variance of the estimator.

In practice, sampling can be done by implicit stratification in order to ensure a geographical representative sample. When selecting towns, this can be done by ensuring that the towns are ordered by size, distance to a regional centre, or other aspects that are considered of importance for the variation of the water service.

Within a town, implicit stratification can be done by selecting households in all parts of the towns as was done using enumeration maps in Ethiopia. Such implicit stratification is likely to reduce the variance but is not taken into account when calculating the variances. This is the case because no information about the subpopulations is available.[6]

The standard error of the measured variable is simply the square root of the variance expression in Equation 3.6:

Equation 3.7 $$std.err.\left(\bar{y}_{Sample}\right) = \sqrt{\mathrm{var}\left(\bar{y}_{Sample}\right)}$$

The 95% confidence interval is approximate and defined as:

Equation 3.8 $$\bar{y}_{Sample} \pm u_{1-\alpha/2} \cdot std.err.\left(\bar{y}_{Sample}\right)$$

$$\approx \bar{y}_{Sample} \pm 2 \cdot std.err.\left(\bar{y}_{Sample}\right)$$

---

[5] Cf. p. 307 in Cochran, W. G. Sampling Techniques. Third edition. Wiley 1977.

[6] If information about the subpopulations is available, stratified random sampling can be used. In this case it may be possible to divide a heterogeneous population into subpopulations, each of which is internally homogeneous, and obtain a gain in the precision of the estimates, cf. Cochran (1977).

$u_{1-\alpha/2}$ is the standard normal quantile for the 95% confidence interval. This 95% quantile is equal to 1.96 and since this value is very close to 2, it is stated that the 95% confidence interval is approximately equal to plus minus two times the standard error.

The results of the survey may be used to estimate the design effects for use in the same or future surveys. The design effect is estimated by:

$$\textit{Equation 3.9} \quad D = \frac{Var(random\ sample)}{Var(cluster\ sample)} = \frac{\bar{y}_{Sample}\left(1-\bar{y}_{Sample}\right)/(n-1)}{\frac{1}{m(m-1)}\sum_{k=1}^{m}\left(\bar{y}_k - \bar{y}_{Sample}\right)^2}$$

And the rate of homogeneity by:

$$\textit{Equation 3.10} \quad \sigma = \frac{D-1}{n_k - 1}$$

The estimates of the design effect and the rate of homogeneity in Table 3.1 are calculated using *Equation 3.9* and *Equation 3.10*.

4 Implementation of the household survey

This chapter describes the practical aspects of the implementation of the survey in three phases: preparatory activities, testing, and data collection. A detailed account of the implementation activities is provided in the following sections. These activities include:

- organising a data collection team;
- adjustment of survey instrument;
- pilot testing;
- training of enumerators;
- quality control;
- coding answers into database.

The implementation team should consist of a supervisor and a number of enumerators and data entry operators.

The questionnaire is adjusted to the local conditions where the survey is to be conducted. This could be done by taking the survey instrument in Appendix 1 as a starting point and testing the questionnaire with a number of households. Part of this work is to identify correct translations of the terms and concepts included in the questionnaire.

The aims of the pilot testing are both to have a final test of the survey instrument and to test if the chosen sample selection procedure works. In addition, the pilot testing is an important part of the training of the enumerators. The training is important, as it is the understanding of the enumerators that determines the quality of the data.

A quality control check is conducted in order to ensure that the interviews have actually been carried out and that they were conducted in a professional manner. Finally, coding answers using a pre-designed entry format should minimise the risk of errors; this is discussed at the end of the chapter.

A prerequisite for conducting any survey is that the local authorities are well informed and that the enumerators are able to document who has authorised the interviews. A letter from the authorising institution, e.g. the Ministry or a regional authority should be obtained, and local authorities where the interviews are conducted should be informed.

4.1 Composition of data collection team

The implementation team should consist of a supervisor and a number of enumerators and data entry operators. Experience market research companies will often have a structure for this type of work that varies from country to country. So the composition described below if one out of many ways of possible constellations.

For a survey of 1,000 households the implementation team could consist of:

- one coordinator/supervisor;
- ten enumerators, divided into teams, each with a Team Leader;
- four back-checkers;
- one data entry programmer;
- one or more data entry operators.

The number of enumerators should be chosen such that each enumerator achieves a high level of efficiency, while the number of interviews does not become too tedious. Between 100 and 200 interviews per enumerator is appropriate. The number of interviews one enumerator can conduct per day vary depending on the geography of the area. In the towns in Oromia one interviewer was able to conduct 8-10 interviews per day.

The Team Leader should either be an experienced interviewer or be part of the team and supervise some interviews in order to obtain hands-on experience with the working conditions of the team.

The number of teams should be decided upon in order to make the data collection practical in terms of geographical coverage. For surveys that are not conducted in large cities, a number of vehicles and drivers are also needed.

The Data Entry Programmer should design the data entry form, e.g. in the statistical package SPSS. This part of the data entry process requires some programming skills. The data entry operators could be the enumerators so that each enumerator enters his or her own data into the database as this would reduce the number of errors. However, access to computer hardware and software are often limited, which means that there are usually only a few data entry operators who can work at any given time.

4.2 Testing and revising the survey instrument

The aim of this activity is to make relevant changes to the questionnaire so that the questions are understood as intended and the subsequent interviews can be conducted as smoothly as possible. Adjustments to suit local conditions are important, because local differences questions to be worded differently in order to be understood, while some questions are only relevant in some locations.

Hence, this activity involves qualitative field testing of the questionnaire to determine if the questions are relevant and easily understood by the respondents.

Qualitative testing comprises two major activities:

1. Obtain a first-hand impression of the present levels of water supply, sanitation and hygiene practices;

2. Validate the questionnaire in terms of relevance, clarity and cultural sensitivity and to check the consistency of the questionnaire.

Prior to qualitative testing of the questionnaire, a first hand-impression of the existing services, hygiene conditions and other relevant background information is obtained during meetings with local authorities or key informants. A checklist can be used for this purpose (see Appendix 5). This information provides a contextual framework for subsequent assessment of the relevance and accuracy of the questions.

The qualitative assessment of the questionnaire should be conducted in households that represent the diversity of water supply, sanitation, and socio-economic and cultural conditions in the survey area. In each location two or three households are selected with varying access to water supply and sanitation facilities and/or service levels and with different socio-economic backgrounds.

The assessment can be a combination of household in-depth interviews and group discussions with members of the family, visitors, neighbours, etc. A "real" interview situation should be simulated where each question is asked. Questions that are not easily answered should be discussed in detail to determine the causes, e.g. lack of clarity, confusing terminology, embarrassment, etc., and a joint effort is to be made to find alternatives. The findings are recorded and the questionnaire revised.

Translation of questionnaire
The questionnaire is translated into the local language. If there are several local languages spoken in the survey area, a version should be made in each language. Translation of the written questionnaire is preferred, because differences in the translations from enumerator to enumerator are inevitable and during the subsequent analysis, it could not then be determined whether the differences are due to conditions of the households or to variations in the wording of the questions.

However, it may be that the local language is seldom used as a written language and is therefore difficult for the enumerators to work with. In this case the enumerators' language skills in the local language should be tested to ensure that correct translations will be made.

The translation of the questionnaire should be discussed with the translators, allowing them to raise relevant issues. This discussion will be assisted by a third party who understands both English and the local language(s).

4.3 Pilot testing

The pilot survey is used to test the sampling methodology and the design of the questionnaire. Therefore, the method envisaged for the main survey should be used in the pilot. The pilot interviews should, if possible, include respondents living under conditions similar to those found in the population, i.e. in terms of socio-economic and geographical characteristics and water and sanitation services. This will test whether the questionnaire is relevant to and correctly understood by all groups of respondents.

Typically, a pilot test comprises 20-50 interviews.

The pilot testing is an important part of the enumerators' training. As discussed further below, one purpose is to test if the enumerators understand the methodology.

The data from the pilot testing should be entered into a database and analysed in order to check that all questions have been well understood. On the basis of the experience gained during the pilot testing and the analysis of the results of the pilot phase, changes can be made in the questionnaire and methodology as well as in the data entry form.

4.4 Training of enumerators

Training enumerators is very important, as the quality of the data collection determines the quality of the results of the survey.

The training comprises the following:

- purpose of the survey;
- pilot briefing;
- pilot debriefing;
- main phase briefing.

In addition, the enumerators should be supervised by the local coordinator during the entire data collection phase.

Purpose of the survey
The purpose of the first meeting with the enumerators is to explain the background and purpose of the survey.

Also, the nature of the work should be explained, including travelling conditions (if any), duration of interviews, and remuneration.

Pilot briefing
The enumerators for the pilot and the main phases should be briefed face-to-face by the consultant in order to make sure that the task is well understood.

The briefing is required in order to describe the questionnaire and the approach for the interviews in terms of:

- how to fill in the questionnaire, including explanation of probes, jumps - "GO TO" instructions;
- what to do if responses cannot be coded, etc;
- manners and approach with respondents;
- return of questionnaires, etc.

The consultant should go through and explain carefully the meaning of each question, so that enumerators understand exactly the type of information asked for. This is especially important regarding technical issues about water and sanitation. Water supply and sanitation technologies should be explained and the enumerators should be familiarized with these by use of the show cards as described in Chapter 2.

The sampling methodology, including how to select the households should also be carefully explained. If no list is available, which will often be the case, the concept of random selection should be explained as well as the rules for selecting households for the interviews.

Finally, the back-check procedure should be explained. As described below, the back-check procedure is a quality control check that consists of revisiting a sub-sample of the households in order to ask a limited number of questions that can verify the results of the main survey.

The purpose of the back-check procedure is to deter the enumerators from filling in the questionnaires by themselves. Hence, the purpose of explaining the procedure is to inform the enumerators about the likelihood that such behaviour will be detected. Back-checks may also be helpful in revealing undisclosed difficulties with question interpretation.

After the briefing, the enumerators should interview each other in pairs, so that each enumerator has interviewed and been interviewed. The completed questionnaires will be checked by the consultant to ensure that the questionnaire has been completed properly (e.g. answers being circled, instructions followed etc.).

If any lack of understanding is perceived, further explanation will be provided and, if necessary, the entire interview team will be informed.

The briefing lasts 3-4 hours. When the briefing is concluded, pilot interviews are initiated.

Pilot debriefing
The pilot survey is used to test the sampling methodology and the design of the questionnaire. The debriefing session is conducted by the consultant with the enumerators. This will establish firsthand whether there are problems with the sampling methodology or the questionnaire. It will also allow the consultant to check that the questionnaires have been filled in correctly.

Main phase briefing
Data collection will begin with a briefing session. This briefing session will focus on lessons learned from the pilot survey and on changes in the final version of the questionnaire, as well as the practical aspects of the sampling strategy.

Supervision and control during the main phase
During this phase, the coordinator should follow the enumerators' work in order to ensure that the data collection is progressing according to the plan and solve any difficulties that may occur.

Supervision involves:

- ensuring that the rules for recruitment of households are followed;
- ensuring that the enumerators ask all questions and that the instructions in the questionnaire are followed;
- discussing issues regarding the conduct of the survey.

Preferably, supervision is done by daily meetings, but daily phone contact is often more practical and works almost equally well.

4.5 Field work quality control

The back-checking procedure should be initiated to ensure the quality of results and to ensure that the questionnaires are valid (and not filled in by the enumerators by themselves).

A sample of 10% of the interviews should be back-checked. This sub-sample should be selected after the termination of the main data collection phase and chosen such that all enumerators are back-checked. For a sample of 1,000 households conducted by 10 enumerators, 10 interviews by each enumerator will be back-checked.

It is important to make sure that addresses are provided by interviewees so that it is possible to locate them. If an enumerator does not fill in addresses or phone numbers, then the coordinator should investigate the issue.

The back-checking questionnaire includes questions that can verify the answers from the main questionnaire. Furthermore, the back-checking questionnaire should include questions on the conduct and length of the interview.

The back-checking should not include enumerators. Training for this group should be similar to the training given to enumerators

4.6 Coding of questions and data entry into database

The data collected is entered into a database using a pre-coded entry form, e.g. an SPSS form. The use of the form ensures that only data within the correct range can be coded. For instance, if the possible answer categories are *yes*, *no* and *don't know,* the variable can only have values between 1 and 3. If the operator codes 13 by mistake, this is rejected by the form. The pre-coded form reduces the number of errors in data entry.

While the programming of the pre-coded entry form requires programming skills, the entry of the data does not require computer skills.

Both the results from the main questionnaire and the back-check questionnaire should be entered into the database. This will enable cross-checking results from the main survey with those from back-checking.

The supervisor will check if the data are coded correctly, both by checking the data file and by checking randomly chosen questionnaires.

4.7 Information and documentation to local authorities

Two prerequisites for conducting any survey are that the:

- households are confident that the enumerator actually represents the institution that he or she claims to represent;

- local authorities are well informed so that no obstructions are encountered.

To facilitate data collection, it is important for the enumerators to have a letter from the authorities explaining the purpose of the survey. This letter would typically be obtained from the relevant Ministry or regional authority.

The purpose of the letter is to provide documentation and legitimacy to the enumerators so that they can show the letter to both households and local authorities as needed.

The local authorities should be informed about the survey. This should be done by a courtesy visit to the local authorities where the purpose of the survey is explained. However, it is crucial that the local authorities do not take part in selecting of households for interviews, because this could lead to bias.

5 Case Study in Ethiopia

This chapter presents the methodology implemented in the case study in Ethiopia. The questionnaire was prepared before the field trip and was further developed by the use of qualitative interviews with households in small towns in Oromia. The final version of the questionnaire is reported in Appendix 1.

A subset of towns in Oromia was selected as the target of the survey. This means that the result of the survey is representative for the 73 towns selected for projects aiming at improving the households' water and sanitation facilities.

A two stage cluster sample procedure was applied to select the sample. From a total of 73 towns, 24 towns were selected randomly with probability proportionate to size (PPS). The number of towns was chosen in order to have a reasonable low standard deviation of a central parameter while keeping the number of interviews per enumerator in each town practical.

At the second stage of the sample selection we used enumeration maps. These maps were used by the statistical authorities in the 1994 for the last census. The enumerators were asked to select households spread out in each enumeration map. and to mark on the map where the households interviewed can be relocated. For small towns the enumeration maps are too large to be practical. Therefore, a different approach was taken for the four smallest towns. Here the enumerators were instructed to start interviewing from the outer bound of the town and select randomly chosen household equally spread between the outer bourn and the centre.

The methodology was successfully tested by a pilot in Mojo in order to test both the questionnaire and if the enumerators were able to use enumeration maps as envisaged. The main phase of the data collection was initiated on Saturday 23 October, and a total of 1013 households were interviewed.

The result of the survey is 28 carefully selected indicators (cf. Chapter 2). The aim has been to identify a limited set of well defined indicators that can be comparable between projects and over time. For each indicator a mean and standard deviation is estimated and for the presentation, confidence intervals are presented.

The survey indicates that 46% (±7%) of the households have access to improved water services and 38% (±5%) of the households have access to an improved sanitation facility.

More than four out of every five households have soap, while few households (5%, ±3%) are observed to have adopted appropriate hand washing practices, which means that one or more of the necessary preconditions for hand washing was unavailable: soap, water and sink, bucket or equivalent.

This contradicts the finding that 30% (±9%) claims to wash their hands at critical times indicating that the observation approach may not result in reliable results. This is discussed further below.

5.1 Survey instrument

The questionnaire was developed by taking the questionnaire recently developed and harmonised by the Joint Monitoring Programme (JMP) and other questionnaires as starting point, and though discussions with the World Bank both in Washington and in Addis Ababa. Furthermore, the questionnaire was tested in Ethiopia by conducting qualitative interviews. The questionnaire is reported in Appendix 1.

In addition, a back-check questionnaire was developed for the quality control. This back-check questionnaire is reported in Appendix 3.

A set of picture cards were used to support the enumerators in understanding and asking questions on the different water supply and sanitation facilities, cf. Appendix 4. The show cards were based on those used in the WHS survey, supplemented by materials obtained from the Loughborough University Image Library and other sources.

The quality of the illustrations can, to some extent, be improved in order to more clearly present the characteristics of a given technology and to take into consideration the variations in the design of the different technology types across countries and continents. Such improved illustrations are currently being procured by the WHO.

5.2 Sampling methodology[7]

Through discussion with the regional planning office in Oromia and the World Bank it was decided to carry out the survey among households in urban areas in the Oromia region. There are approximately 550 towns in this region, and 73 towns had been chosen as candidates for possible water and sanitation projects. A number of selection criteria provided the background for this selection and, generally, the selected towns have a low level of water and sanitation services

[7] The description of the sample selection in this section follows the boxes in Chapter 3 and is included here to make the chapter on the results from Ethiopia self-contained.

while there is some potential for cost recovery. The towns vary in size between 2,000 and 45,000 inhabitants.

Thus the population from which the sample was taken is the 73 towns and not the entire population of towns in Oromia.

5.2.1 Sample design

The sample design used was the two-stage cluster sampling procedure with a number of towns selected in the first stage (primary sampling units, PSUs) and a number of households selected in the second (basic sampling units, BSUs).

First stage - selection of towns
Information about the population in the 73 towns in 2004 was applied as the basis for a sample selection based on probability proportionate to size (PPS). A total of 25 clusters were selected resulting in 24 towns, because two clusters were chosen from one large town (Waliso).

The number of primary sampling units was decided on the basis of the available resources and the marginal gain in precision of increasing the number of PSUs compared to the additional costs of increasing the number of clusters. With the assumption of a rate of homogeneity of 0.10 for the measured variable, it was decided that the survey should consist of 25 towns with 40 interviewed households in each town. For a variable with an expected proportion - or frequency - of 40% in the population, this means a standard error of 4.3% and a 95% confidence interval of 33% - 47%. If this interval is too wide to be acceptable vis-à-vis the object of the survey, the number of towns should be increased in order to reduce the standard deviation of the estimator.

The 25 primary sampling units were drawn with a probability proportionate to the size of the towns from the population of 73 towns. In practice, the PPS procedure is carried out by creating a cumulative list of town populations and then selecting a systematic sample from a random start.

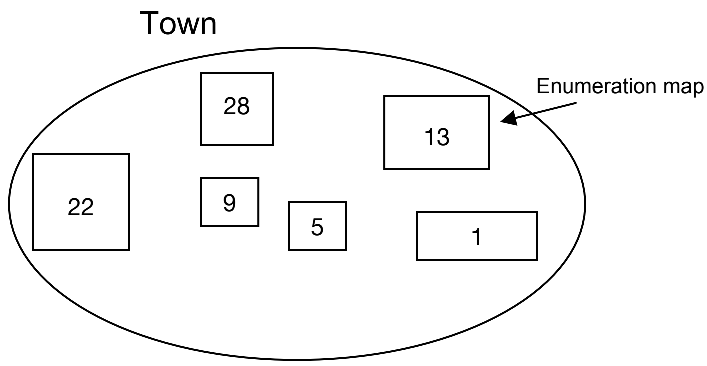
In order to obtain a self-weighted sample, a constant number of households is selected within each of the 25 PSUs. Using this method, one town can be selected twice or more and this is the reason for distinguishing between towns and clusters. In the town that was chosen twice, Waliso, 80 households were interviewed.

Second stage - selection of households
For the selection of households in each town, enumeration maps from the last population census were used. This was to ensure that the households selected were spread all over the town. Enumeration maps are particularly helpful when selecting a sample for a water and sanitation survey (and other infrastructure services), because households that are geographically close are likely to have the same level of service. Hence, the application of enumeration maps gives a sample, which is implicitly stratified and this lowers the variance of the estimates.

Enumeration maps were obtained for the 20 medium-sized towns and, hence, not for the smallest four towns, as enumeration maps for very small towns are large and unwieldy.

For example, in a town with 30,000 inhabitants there are around 30 enumeration maps. A random selection procedure was used to select 6 maps out of the 30. An example of such a selection of enumeration maps is illustrated below.

Figure 5.1 Selection of enumeration map areas in a town



As illustrated in the figure, the enumeration maps vary in size according to the population density and will typically be larger in the outer areas than in the more densely populated centres of the towns.

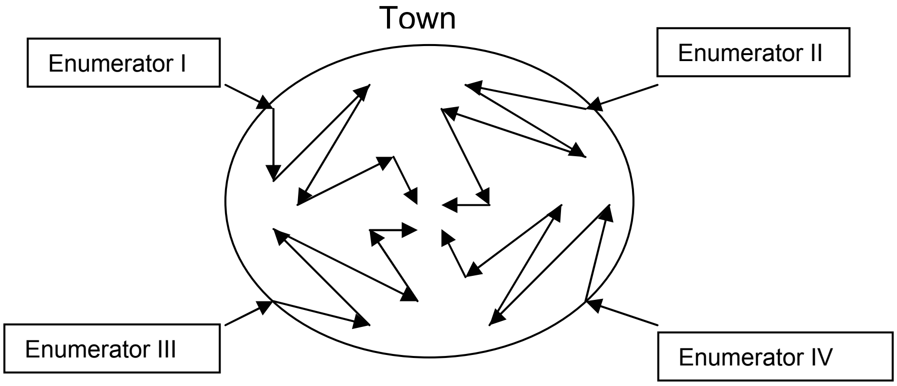
The selection of the 5-7 households to interview within each enumeration map area is carried out using a method of random selection. Rather than instructing the enumerators to use a specific rule to obtain a randomly selected sample, the concept of randomness was carefully explained along with different ways of making a random selection (flipping a coin, spinning a bottle, taking every third household).

Whether the enumerator interviews households spread out on the enumeration map is easily checked, as interviewed households are marked on each map by the enumerator.

In small towns with less than 6,000 inhabitants it was decided that the size of the enumeration maps become too large in order to work in practice, as the enumerators would have difficulties operating maps that measures 1x1 meter.

As an alternative strategy of selecting households, the enumerators identified four points in the outer area of these towns where each enumerator begins his or her rounds of interview, c.f. illustration below.

Figure 5.2 Selection of households in the small towns



From these starting points each enumerator made his or her way towards the centre of the town while randomly selecting households. As above, the random selection of households could be done in several ways. Each enumerator was free to choose his or her strategy for random selection, choosing between flipping a coin, spinning a bottle etc. as long as the final pattern of selection was similar to the pattern illustrated in the above figure. The approach taken here was based on the idea to give the enumerators an understanding of the importance and meaning of random selection rather than giving an explicit rule. In the survey in Peru, more precise rules were given for the selection of households.

5.3 Implementation

The survey methodology was tested thoroughly before the main data collection. This was done both by conducting qualitative interviews with households in order to test the relevance and wording of the questions, and by conducting pilot interviews where the instrument and the sample selection procedure were tested.

The implementation followed the time schedule below:

- Test of questionnaire, qualitative interview: 13-14 September 2004
- Pilot interviews: 15 September 2004
- Main phase interviews: 18-31 September 2004
- Coding answers into database 4-13 October 2004
- Back-checking interviews 20-25 October 2004

The time schedule for the implementation, test of questionnaire, and pilot interviews was tight. The coding of the answers was expected to take one week, but took instead one and a half weeks. The back-checking was initiated after the data file was received by the consultant.

The total duration of the process from initiation to termination of back-checks was two months.

5.3.1 Qualitative work

Prior to the testing of the questionnaire and the household selection procedure by the enumerators in the field, a qualitative assessment and test was conducted by the consultant's team with the aim to:

- obtain a first-hand impression of the present levels of services and hygiene conditions;
- validate the questionnaire in terms of relevance, clarity, and the cultural sensitivity of the questions;
- check the consistency of the questionnaire.

A first hand-impression of the existing services, hygiene conditions and other relevant background information was obtained from the local authorities (Wereda, Kebele, Water Office) during courtesy visits. This background information was useful in assessing the relevance and effectiveness of the questionnaire. A uniform check list was prepared specifying relevant water supply, sanitation and hygiene topics (see Appendix 5).

The qualitative assessment of the questionnaire was conducted in three towns. The towns were selected in such a way that that small towns as well as medium towns were included. Care was taken to ensure that different technical options were represented. Due to time constraints, highland towns close to Addis Ababa were chosen. Therefore, technologies, socio-economic, and socio-cultural conditions that may be prevalent in other parts of Oromia region, in particular lowland towns, were not represented. Preferably, the locations for the qualitative assessment tests should be selected carefully so as to capture the full range of diversity.

The assessment approach was a combination of household in-depth interviews and group discussions conducted in a number of households that were selected to cover a variety of water supply and sanitation technologies and different socio-economic characteristics. The English version of the questionnaire was used and the interpreter translated all questions into the appropriate language. Questions that were not easily answered were discussed in detail with the interviewee and other members of the family, neighbours who had dropped in etc, to identify the causes of misunderstandings (lack of clarity, imprecise terms, etc) and to find improved or alternative formulations.

The interpreter also provided feed-back on the usefulness and clarity of the questions, the observation parts of the questionnaire, and the enumerator instructions.

The questionnaire was subsequently revised and adjusted in accordance with the findings of the qualitative assessment and translated into the local language.

5.3.2 Training of enumerators

The training of the enumerators was initiated Tuesday and terminated Friday so that the data collection could start Saturday morning. Twelve enumerators were employed for the interviews and they were all graduate students. They were divided into three teams and a team leader was appointed for each team. In addition, each team was provided with a car with driver.

The training sessions were as follows:

Tuesday
The purpose of this initial meeting with the enumerators was to give them a brief introduction to the purpose of the survey as well as the working conditions.

Wednesday
Wednesday morning the enumerators were trained in a classroom session where we carefully explained the purpose of all questions in the questionnaire and described the sampling method as well as the concept of randomness. This was followed by test interviews where the enumerators interviewed each other.

In the afternoon all enumerators and the entire consultant team went to Mojo to conduct the pilot survey, 60 km from Addis Ababa. This was to test both the questionnaire and if the enumerators were able to find the area on the enumeration maps. A total of 21 pilot interviews were conducted, and some enumerators chose to conduct the interviews in pairs.

Thursday
Thursday the enumerators were debriefed in order for the consultant's team to learn from their experience and to receive comments to the Oromo version of the questionnaire. The purpose of the debriefing was to obtain information for the final corrections of the questionnaire and to ensure that the methodology envisaged was working and well understood by all enumerators.

Friday
The Friday training session had two purposes. First, to update the enumerators on the changes made from the pilot questionnaire to the final questionnaire. There were only minor changes regarding the translation into Oromo. Secondly, the enumerators were divided into teams and the selected towns for the survey were distributed among the teams so that one team was going west, one east, and one south. Also, the enumerators were instructed about their routes.

Supervision during the main phase
The enumerators were supervised from a distance during the main phase. Each team leader was instructed to call the coordinator every day to discuss progresses and to report any difficulties encountered.

5.3.3 Data entry

The answers were typed in using an SPSS form pre-coded for the questionnaire. Such SPSS format was used in order to avoid misunderstandings and errors in the coding

Both the results from the main phase questionnaire and the back-check questionnaire were coded into the database using the SPSS form.

5.4 Results

The estimated indicators for water supply, sanitation, and hygiene are given in the tables below. The indicators in the tables are average indicators for the entire region, i.e. they give a representative view of the situation among the towns selected for water and sanitation projects (73 towns).

Water supply

Continuous access to improved water services as main drinking water source is obtained by 46% (±7%) of the households, cf. Table 5.1. The figure in parenthesis represents the limits of the 95% confidence intervals, meaning that the true estimate most likely lies within the range spanned by the low and high estimates. Hence, the data indicates that continuous access is likely to be available to between 39% and 53% of the households.

More than half of the households, 62% (±7%), have access part of the year. The difference in accessibility is due to seasonality.

Table 5.1 Water supply indicators, mean and 95% confidence intervals

	Indicator	Mean	Obs.
I1	Continuous access to improved water sources as main drinking water source	46%(±7%)	935
I2	Access to improved water supply as their main drinking water source at least part of the year	62% (±7%)	998
I3	Time to collect water is 30 minutes or less	92% (±4%)	805
I4	Daily availability	81% (±4%)	945
I5	Continued supply of main drinking water source/ limited irregular interruptions during the last six months	73% (±6%)	954
I6	Availability of 20 l/c/d	33% (±7%)	492
I7	Reasonable access	11% (±9%)	388

Note: The interval given (±x%) indicates the 95% confidence interval around the mean value.

Most households spend less that 30 minutes making a round trip to collect the water, and most households have access to water on a daily basis.

The indicator for continued supply reflects interruptions during the last six months, with limited interruptions defined as water being interrupted less than a total of 30 days during the last 6 months. 73% (±6%) of the households did not experience so many interruptions.

An adequate quantity of water was defined, in this survey, as 20 litres per household member per day. One third of the households consumes more than 20 litres per day per household member. With reasonable access defined as daily access to improved water all year around, continuous water supply and use of an adequate quantity of water, approximately every ninth household has reasonable access to water.

Appropriate and hygienic drinking water storage facilities demand a container with a narrow neck or cover and that the container and the dipper to take water are both clean. 29% (±8%) of the households have appropriate and hygienic drinking water storage facilities, cf. the table below.

Table 5.2 Indicators for water storage, quality and treatment, mean and 95% confidence intervals

	Indicator	Mean	Obs.
I8	Appropriate and hygienic drinking water storage facilities.	29% (±8%)	337
I9	Satisfaction with the quality of their main drinking water services.	43% (±9%)	1003
I10	Water treatment within the last week.	3% (±0.1%)	968

The number of households included for the calculation of I8 is lower than for the other indicators because only a little more than a third of the households has a water storage facility.

Almost every second household is satisfied with the quality of their drinking water services and few treat the water before drinking.

The initial expenses or contribution to get water services are estimated to 904 Birr (±108 Birr), and the average monthly expenses to 15 Birr (±3 Birr). The households' expenses for a drinking water container is approximately 46 Birr (±5 Birr), cf. Table 5.3.

Table 5.3 Indicators for water cost, Ethiopian Birr (ETB), mean and 95% confidence intervals

	Indicator	Mean	Obs.
I11	Average initial costs of water services	904 Birr (±108)	408
I12	Average monthly costs of water services	15 Birr (±3)	445
I13	Average cost of storage facility	46 Birr (±5)	186
I14	Average costs of water treatment	0.18 Birr (±0.27)	71

Note: 12 Ethiopian Birr (ETB) = 1 USD

Most of the households that treat the water before drinking indicate that they don't pay anything for the water treatment. This is why it is not possible to find a reasonable estimate of water treatment cost (the mean given above is not significantly different from zero).

5.4.2 Sanitation

The results of the sanitation indicators are given in the table below.

Table 5.4 Indicators for access to improved sanitation facility, mean and 95% confidence intervals

	Indicator	Mean	Obs.
I15	Access to improved sanitation facilities	33% (±5%)	982
I16	Access and actual use of improved sanitation facility	17% (±5%)	982
I17	Sanitation facility within less than 5 min. from household	94% (±2)	667

A third of the households has access to improved sanitation facilities.

Less than a fifth of the households has improved sanitation facilities with signs of actual use. These are households where none of the following was observed: dense vegetation in front of the toilet, waste or debris on path, major crevice or potholes on path, or entrance to toilet obstructed. These observations are subjective to the interviewers judgment and hence more uncertain that the first indicator. In order to cross-check the results a similar indicator can be constructed where the interviewer states whether the path is clear to the toilet. This gives an indicator of actual use of the same magnitude.

Most households use a sanitation facility that is less than 5 minutes from the home, cf. table above. This indicator does not include the households that have no toilet facility.

9% (±2%) of the households has an improved sanitation facility kept in good hygienic condition cf. Table 5.5. This means that in most improved facilities, either visual faeces were observed around the hole, surface flow of sewage or the facility smelt badly.

Almost half of households are satisfied with the facility they are using. They rate the following aspects either good or acceptable: effective operation, convenience, cleanliness, smelling distance to toilet, and ease of repair.

Table 5.5 Indicators for quality of sanitation, mean and 95% confidence intervals

	Indicator	Mean	Obs.
I18	Improved sanitation facility in good hygienic condition	9%(±2%)	926
I19	Satisfaction with the quality of the sanitation facility	46%(±8%)	825

The average cost of building a toilet facility and the annual cost of maintenance and repair are given in Table 5.6. The average cost of building a facility is 580 Birr (±140 Birr), and the average annual cost for maintenance and repair is 251 Birr (± 95 Birr). These are average costs for all types of technologies.

Table 5.6 Indicators for cost of sanitation, mean and 95% confidence intervals

	Indicator	Mean	Obs.
I20	Average cost for building facility	580 Birr (±140)	401
I21	Average annual costs for maintenance and repair	251 Birr (±95)	180

Note: 12 Ethiopian Birr (ETB) = 1 USD

It should be noted that the number of households that were able to answer these two questions are lower than for other questions, which indicates that these questions are difficult for the respondents to answer.

5.4.3 Grey water discharge

Improved grey water disposal methods include piped sewer, a soak away system, an open channel or use of the sanitation facility, and 4% of the households have access to such improved grey water disposal methods, cf. the table below.

Table 5.7 Indicators of grey water disposal, mean and 95% confidence intervals

	Indicator	Mean	Obs.
I22	Access to improved grey water disposal methods	4%(±2%)	984

Most households dispose of the used water in the street or another space outside the household's premises or in the yard or the garden of the premises. This is the case for more than four out of every fifth household.

5.4.4 Hygiene practices

Six indicators are used to assess hygiene practices. They are concerned with access to hand washing supplies, use of hand washing supplies at critical times, awareness, disposal of child faeces, and cleanliness of the surroundings near the residence. The results are given in the table below.

Table 5.8 Hygiene behaviour indicators, mean and 95% confidence intervals

	Indicator	Mean	Obs.
I23	Household with soap (or ash, sand or detergent)	78%(±5%)	944
I24	Adoption of appropriate hand washing practices	5%(±3%)	838
I25	Use of soap at critical times	30%(±9%)	824
I26	Awareness of children's hand washing practices	68%(±4%)	824
I27	Proper disposal of children's faeces	52%(±7%)	378
I28	Clean surroundings around residence	70%(±7%)	957

Most all households have soap available for hand washing. Two households use sand or ash as an alternative to soap.

According to the observed hand washing facilities, almost no households have adopted appropriate hand washing practices, which means that not all of the necessary preconditions for hand washing were observed: soap, water, and facility for hand washing (e.g. sink or bucket).

Taken separately, in most households soap was observed, in 16% water was observed and in 19% of the households a facility for hand washing was observed[8]. Hence the inappropriateness of the hand washing facilities is the lack of water or facility rather than the lack of soap.

On the other hand, 30% (±9%) of the households state that they use soap at critical times. This contradicts the observations, and hence either the indicator based on observation underestimate the occurrence of appropriate hand washing practice or the indicator based on the stated use of soap represent and overestimation.

When making observations, the interviewers may not have been probing sufficiently to see the water and the facility for hand washing, for instance in the case that these items were stored away. So in order to avoid this ambiguity in the results it is very important to train the enumerators to probe to see each of the three items.

If the respondents have been educated about appropriate hand washing without actually adopting these, the indicator based on claimed hand washing practice may represent an overestimation. In the questionnaire, the options stating when to wash hands are in the questionnaire for use in coding by the enumerator, but are not read out loud. There is however a risk of some enumerators reading out loud the options so that the respondents need only to say yes to the critical times. This situation should be avoided by appropriate training and supervision, but may still occur.

The availability of soap combined with the question on hand washing at critical times is believed to be the most reliable indicators for hand washing. This is because it is very simple for the enumerator to ask and observe whether the soap is available, whereas it is less obvious for the enumerator to ask for a description of the hand washing practice and even less obvious to ask for a demonstration.

There is a high level of awareness of correct hand washing practices for children. Further, 52% (±7%) of the households with children aged less that 4 dispose of the children's faeces in a hygienically correct way; according to the results of the survey, either the faeces are disposed of in the toilet facility or buried.

[8] These percentages are among households who agreed to show the enumerator the place for hand washing; 13% of the households did not wish to show the hand washing facility. If it is assumed that the reason for not showing the facility is non-existence, the percentages are 75%, 14% and 17% for soap, water and facility (basin), respectively.

This may indicate that the observation regarding necessary preconditions for hand washing does not work as intended, for instance because the bucket used for hand washing is stored away when not in use.

Finally, clean surroundings around the residences with no faecal matter from children or animals were observed in more than two thirds of the households.

5.5 Field work quality control

A back-checking procedure was initiated to ensure the quality of results and to check that the questionnaires were genuine and not filled in by the enumerators at home.

A team of three back-checkers was employed for the work and they were trained by explaining the purpose of the research, the importance of their job as well as the questions in the back-check questionnaire. It was emphasised that the difficult part of the back-checking was to relocate the households interviewed in the main phase.

5.5.1 Back-check sample

A sample of 10% of the households was back-checked and these households were chosen so that interviews conducted by all twelve enumerators could be checked. This was ensured by selecting a limited number of towns and back-checking a large fraction of households interviews in these towns.

The towns were selected by the consultant after the termination of the main interview phase: Gedo, Meki, Agere Marayam and Yabelo.

The back-check questionnaire was in Oromo and in order to ease the job of finding the households, which participated in the survey, the back-checkers received the relevant enumeration maps and addresses of the households (though the addresses contain very limited information).

Ideally the back-checking is conducted during the main phase of the survey, because this may allow early correction of quality flaws. However, in Ethiopia this was not possible, because the data entry was done in Addis Ababa only after the return of the enumerators.

5.5.2 Back check results

The quality assurance is carried out by comparing the result of two identical questions. These are the key questions about main source of drinking water and toilet facility/toilet practice of the household. In addition, a question about the number of household was included in the back-check questionnaire.

The back-checking procedure is in this project used for two purposes:

- to check if the interviews were actually carried out

- to check the reliability of the results technology

The result of the back-checking is analysed below.

Number of household members
To check if there is reasonable accordance between the result in the back-check and in the main questionnaire, the back-check question about the total number of households living in the household is considered. In the main questionnaire this questions is spilt up into adults, children and infants and it is the sum of these three that adds up to the number of household member referred to in table below. The answers from the main phase are given in the columns and the corresponding answers from the back-checking are given in the rows.

Table 5.9 Comparison between results from main and back-checking: Number of household members

	Main phase answer							
Back-check answers	1	2	3	4	5	6	7	8 or more
1 household member	1	0	0	0	0	0	1	0
2 household members	1	5	1	0	0	0	0	1
3 household members	1	5	5	2	1	1	0	2
4 household members	0	0	5	6	4	0	1	1
5 household members	0	4	1	1	7	2	1	1
6 household members	1	0	2	0	5	3	3	2
7 household members	1	2	0	3	2	2	3	6
8 or more	1	2	2	2	2	2	2	6

The dark shaded cells are those where there is accordance between the number of household members; this is the case for 31%. If the light shaded are included in the observations where there is assumed to be accordance, 65% of the observations in the back-checking is in accordance with the results from the main questionnaire.

The level of accordance is lower than expected but not sufficient to conclude that the interviewers did not fill out questionnaires themselves without interviewing the households.

Main drinking water source
The back-checking of the main drinking water source indicates to what degree the result depends on the enumerator collecting the data. Comparing the answers for main drinking water source for the main phase and for the back-check, shows that 57% indicates that they have an improved drinking water source in the main survey[9], whereas 66% indicates this in the back-checking.

The specific answers from the two questionnaires are given in the table below.

[9] This is indicator I2 - where the questions is asked without mentioning seasonality .

Table 5.10 Comparison between results from main and back-checking, main drinking water source

	Main phase answer										
Back-check answer	Piped water into dwelling	Piped water to yard/plot	Public tap/standpipe	Tubewell /borehole	Protected dug well	Unprotected dug well	Protected spring	Unprotected spring	Small-scale vendor	Tanker-truck	Surface water
Piped water into dwelling	0	0	0	0	0	0	0	0	0	0	0
Piped water to yard/plot	8	13	8	0	0	0	0	0	5	0	1
Public tap/standpipe	3	3	18	0	1	0	2	1	12	0	1
Tubewell/borehole	0	0	0	0	0	0	0	0	0	0	0
Protected dug well	0	0	0	0	0	0	0	0	0	0	0
Unprotected dug well	0	1	1	1	1	0	0	0	0	0	0
Protected spring	0	0	6	0	0	0	0	0	0	0	1
Unprotected spring	0	0	3	0	0	0	1	0	0	1	1
Small-scale vendor	0	0	3	0	1	0	0	0	17	0	1
Tanker-truck	0	0	0	0	0	0	0	0	0	0	0
Surface water	1	1	0	0	0	0	0	0	0	0	9

Less than half of the observations are in accordance. This may be due to uncertainty regarding the definition of the technology both among the enumerators and the back-checkers as well as the respondents. This may indicate that there is need for more training or better defined categories than used in the present survey.

Toilet facility/toilet practice

Similarly, the results of the back-checking are compared to the result from the main phase of the data collection. Here, whereas 40% of the households has improved toilet facilities according to the main phase interviews, only 17% of the household has these facilities according to the back-checking.

For 37% of the households, the same answer is obtained in the main phase as in the back-checking.

Table 5.11 Comparison between results from main and back-checking, toilet facility/toilet practice

	Main phase answer							
Back-check answer	Flush to piped sewer system	Flush to septic tank	Flush/ pour flush to pit	Flush/ pour flush to elsewhere	Composting toilet	VIP/pit latrine with slab	Pit latrine without slab/open pit	No facilities/bush/ field
Flush to piped sewer system	0	0	0	0	0	0	0	0
Flush to septic tank	0	0	0	0	1	0	0	0
Flush/pour flush to pit	0	0	0	0	0	0	0	0
Flush/pour flush to elsewhere	0	0	0	0	0	0	1	0
Composting toilet	0	0	0	0	0	0	0	0
VIP/pit latrine with slab	1	0	2	0	5	4	10	0
Pit latrine without slab/open pit	3	0	7	2	7	15	30	15
No facilities/bush/field	1	0	1	0	2	4	3	15

The results of the back-checking may indicate that more work needs to be done in describing the various technologies and in training the enumerators to distinguish between the technologies.

5.6 Design effect and rate of homogeneity

The design effect and the rate of homogeneity have been calculated for each indicator. Both vary depending on the question forming the basis of the calculation. For instance for an indicator where the difference across towns is limited, the design effect and the rate of homogeneity will both be low.

This table has been included because it may be useful for future surveys. The design effect should be used for the questions which are most important regarding the purpose of the survey. If the survey focused on drinking water, the design effect for I1 should be used to calculate expected standard deviation; if the survey is focused on toilet facilities, the design effect for I15 should be used.

Table 5.12 Design effects and rate of homogeneity for all indicators

No	Indicator	Design effect	Rate of homogeneity
I1	**Continuous access to improved water sources as main drinking water source**	**4.25**	**0.09**
I2	Access to improved water supply as their main drinking water source at least part of the year	5.49	0.12
I3	Time to collect water is 30 minutes or less	4.88	0.12
I4	Daily availability	2.46	0.04
I5	Continued supply of main drinking water source/ limited irregular interruptions during the last six months	4.41	0.09
I6	Availability of 20 l/c/d	3.08	0.11
I7	Reasonable access	5.29	0.15
I8	Appropriate and hygienic drinking water storage facilities.	2.67	0.13
I9	Satisfaction with the quality of their main drinking water services.	8.95	0.20
I10	Water treatment within the last week.	0.53	-0.01
I11	Average initial costs of water services	2.01	0.07
I12	Average monthly costs of water services	6.96	0.35
I13	Average cost of storage facility	1.21	0.03
I14	Average costs of water treatment	0.87	-0.07
I15	**Access to improved sanitation facilities**	**2.91**	**0.05**
I16	Access and actual use of improved sanitation facility	4.03	0.08
I17	Sanitation facility within less than 5 min. from household	1.84	0.03
I18	Improved sanitation facility in good hygienic condition	1.55	0.02
I19	Satisfaction with the quality of the sanitation facility	4.84	0.12
I20	Average cost for building facility	1.82	0.05
I21	Average annual costs for maintenance and repair	1.74	0.12
I22	Access to improved grey water disposal methods	3.72	0.07
I23	Availability of soap in the household	3.92	0.08
I24	Adoption of appropriate hand washing practices	3.51	0.08
I25	Use of soap at critical times	8.48	0.23
I26	Awareness of children's hand washing practices	1.39	0.01
I27	Properly disposal of children's faeces	1.84	0.06
I28	Clean surroundings around residence	6.09	0.14

Some of the estimated values of the rate of homogeneity are less the zero (I10, I14). This is due to low values of the indicator and sampling variation. The rate of returns less than zero should be interpreted as 0, i.e. that there is not more correlation between the households living in the same town than household living in different towns. Water treatment is an example of this; very few household treat the water and almost no households are able to give an estimate of the amount spent on the water treatment.

On the other hand, the largest variation across towns is found for reasonable access (I7) which means that there is a correlation between household with reasonable access and the town that they live in.