APPENDIX A. SAMPLE DESIGN

The major features of sample design are described in this appendix. Sample design features include target sample size, sample allocation, sample frame and listing, choice of domains, sampling stages, stratification, and the calculation of sample weights.

The primary objective of the sample design for the Montenegro Multiple Indicator Cluster Survey was to produce statistically reliable estimates of most indicators, at the national level, for urban and rural areas, and for the three regions of the country: South, Central and North.

A stratified, two-stage random sampling approach was used for the selection of the survey sample.

Sample Size and Sample Allocation

Montenegro is characterised by a very low fertility rate and a small number of household members. For example, one generation of children born makes up less than 1 percent of the population, and the average number of household members is around 3. Owing to these facts the modification of the recommended sample plan had to be made, and that was the stratification of households in selected census block units into two categories: households with children and households without children under 5. The allocation of the sample in the category of households with children was significantly bigger than the allocation of the sample in the category of households without children.

The target sample size for the Montenegro MICS was calculated as 2300 households. For the calculation of the sample size, the key indicator used was the percentage of children aged 0-4 years who had had Acute Respiratory infections. The following formula was used to estimate the required sample size for these indicators:

$$n = \frac{[\,4\,(r)\,(1\text{-}r)\,(f)\,(nr)\,]}{[\,(me)\,2\,(r)2\,(p)\,(nh)\,]} \qquad (1)$$

where

- *n is the required sample size, expressed as number of households*
- *4 is a factor to achieve the 95 per cent level of confidence*
- *r is the predicted or anticipated prevalence (coverage rate) of the indicator*
- *nr is the factor necessary to raise the sample size by 100(nr -1) percent for non-response*
- *f is the shortened symbol for deff (design effect)*
- *me*r is the margin of error to be tolerated at the 95 percent level of confidence, defined as me percent of r (relative sampling error of r)*
- *p is the proportion of the total population upon which the indicator, r, is based*
- *nh is the average household size.*

For the calculation, *r* (percentage of children aged 0-4 years who had had Acute Respiratory infections) was assumed to be 12 percent. The expected non-response rate *nr*, was determined at 15 percent. The value of *deff* (design effect) was taken as 1.5 based on estimates from previous surveys. The maximum relative error allowed (*me*) was 20 percent, *p* (percentage of children aged 0-4 years in the total population) was taken as 6.5 percent and *nh* (average household size) was taken as 3.

The resulting number of households from this exercise was 6478 households. Only a sample of that size would provide a significant number of children under 5 for drawing reliable conclusions. Therefore, in order to cut down the number of households in the sample, but not to lose estimation reliability, the stratification of the sample into categories with and without children aged 0-4 years was needed. For calculation of the necessary number of households in each category, the following formula was used:

$$n = (n_s)\ (n_c)\ (p_s) \qquad (2)$$

where

- n is the required sample size, expressed as the number of households
- n_s is the expected number of households with, or the number of households without children under 5 in a cluster, depending on what category the calculation is used
- n_c is the number of clusters in the sample, and
- p_s is the probability of selection of the household in each category.

Taking into account that the proportion of children under 5 in the total population, *p* was 6.5 percent, and if the average household size is 3, the estimated number of households with children was 19.5 per 100 households (the average number of households in each cluster). So the n_s was assumed to be 19.5 for the category with children, and 80.5 for the category without children. The probability of selection of a household (p_s) with at least one child out of all households with children was assumed to be 0.5, and the probability of selecting a household without children from all households with children in each cluster was 0.1. Supposing that 140 clusters were about to be selected, the total number of households was calculated at 1365 households with, and 1127 of households without children under 5, which makes a total of 2492 households.

The average cluster size in the Montenegro MICS was determined as 18 households, plus 3 backup households. Back-up households were to be interviewed only if some of the first 18 households were not found. In cases where a household refused to be interviewed, the substitution with a back-up household was not possible. The calculation was based on a number of considerations, including the budget available, and the time that would be needed per team to complete one cluster. Dividing the total number of households by the number of households per cluster, it was calculated that the selection of a total number of 141 clusters in Montenegro would be needed in all regions.

Allocation of the total sample size to the three regions was targeted with probability proportional to the regions' size. Therefore, 141 Montenegro sample clusters were allocated across the regions, with the final sample size calculated at 2538 households (141) clusters * 18 households per cluster). In each region, the clusters (primary sampling units) were distributed to urban and rural domains, proportional to the size of the urban and rural populations in that region. The table below shows the allocation of clusters to the sampling domains.

Table SD.1: Allocation of Sample Clusters (Primary Sampling Units) to Sampling Domains

Region	Population (Census 2003)			Number of Clusters		
	Urban	Rural	Total	Urban	Rural	Total
North	24789	16961	41750	19	13	32
Central	65235	18266	83501	50	14	64
South	24789	33922	58712	19	26	45
Total	114814	69149	183963	88	53	141

Sampling Frame and Selection of Clusters

The 2003 Montenegro Population Census framework was used for the selection of clusters. Census enumeration areas (app. 100 households) were defined as primary sampling units (PSUs), and were selected from each of the sampling domains by using systematic pps (probability proportional to size) sampling procedures, based on the estimated sizes of the enumeration areas from the 2003 Population Census. The first stage of sampling was thus completed by selecting the required number of enumeration areas from each of the 3 regions by urban and rural areas separately.

Listing Activities

Since the sample frame (the 2003 Population Census) was not up to date, household lists in all selected enumeration areas were updated prior to the selection of households. For this purpose, listing teams were formed, who visited each enumeration area, and listed the occupied households. The Statistical Office of the Republic of Montenegro (MONSTAT) and The Strategic Marketing Research Agency were responsible for updating household lists. The listing exercise was performed by teams which were the direct implementers of the field work during the course of data collection that came later. The whole territory of Montenegro was divided into 5 districts according to the regional network of institutions responsible for listing and fieldwork. In each district one or two teams of people was selected – one supervisor for the district and the interviewers (whose number depended on the number of clusters in the region). Criteria for the selection of the interviewers and supervisors were their qualifications, communication skills, experience in the field work and knowledge of the region where research was to be conducted. A total of 7 teams were formed. For each team, the list of all households in the selected cluster from the last census was provided. The interviewers' task was to go to the addresses listed and to mark any change that had happened, e.g. the dwelling didn't exist any more, the household had moved away from the dwelling and another household was living there, and to note the number of children under five living in the household. The listing process was performed during September 2005. Besides providing updated information on households, updating household lists made interviewers more acquainted with the field.

Selection of Households

Lists of households were prepared by the listing teams in the field for each enumeration area. The households were then sequentially numbered from 1 to n (the total number of households in each enumeration area) at the Strategic Marketing Research Agency. Selection of 18 plus 3 back-up households with equal probability in each enumeration area was carried out using the method of random start and equal random walk (simulation of the SRSWoR scheme). Before the selection of households, updated census block units were put into two categories: households with children and households without children under 5.

Calculation of Sample Weights

The Montenegro Multiple Indicator Cluster Survey sample is not self-weighted. Sample weights were used essentially because the sample stratification according to region, type of settlement and households with and without children under 5 was made. Calculated sample weights were used in the subsequent analyses of the survey data.

The major component of weight for both samples is the reciprocal value of the sampling fraction employed in selecting the number of sample households in that particular stratum:

$$W_h = 1 / f_h \qquad (3)$$

The term *fh*, the sampling fraction at the *h-th* stratum, is the product of probabilities of selection at every stage in each stratum:

$$f_h = P_{1h} * P_{2h} \qquad (4)$$

where P_{ih} is the probability of selection of the sampling unit in the *i-th* stage for the *h-th* stratum.

where P_{ih} is the probability of selection of the sampling unit in the *i-th* stage for the *h-th* strata.

Since the estimated numbers of households per enumeration area prior to the first stage selection (selection of primary sampling units) and the updated number of households per enumeration area were different, individual sampling fractions for households in each enumeration area (cluster) were calculated. The sampling fractions for households in each enumeration area (cluster) therefore included the probability of selection of the enumeration area in that particular sampling domain, and the probability of selection of a household in the sample enumeration area (cluster).

A second component which has to be taken into account in the calculation of sample weights is the level of non-response for household and individual interviews. The adjustment for household non-response is equal to the inverse value of:

$$RR = \text{Number of interviewed households} / \text{Number of occupied households listed} \qquad (5)$$

After completion of the fieldwork, response rates were calculated for each sampling domain. These were used to adjust the sample weights calculated for each cluster. Response rates in the Montenegro Multiple Indicator Cluster Survey are shown in Table HH.1 in this report.

Similarly, the adjustment for non-response at the individual level (women and under-5 children) is equal to the inverse value of:

$$RR = \text{Completed women's (or under-5's) questionnaires} / \text{Eligible women (or under-5s)} \quad (6)$$

The numbers of eligible women and under-5 children were obtained from the household listing in the Household Questionnaire in households where interviews were completed.

The unadjusted weights for the households were calculated by multiplying the above factors for each enumeration area. These weights were then standardised (or normalised), one purpose of which is to make the sum of the interviewed sample units equal to the total sample size at the national level. Normalisation is performed by multiplying the aforementioned unadjusted weights by the ratio of the number of completed households to the total unadjusted weighted number of households. A similar standardisation procedure was followed in obtaining standardised weights for the women's and under-5s questionnaires. Adjusted (normalised) weights varied between 0.48 and 1.28.

Sample weights were appended to all data sets and analyses were performed by weighting each household, woman or under-5 with these sample weights.