

DIGNITY – Sampling Document

The DIGNITY survey is representative of low-income areas and slums of the Dhaka City Corporations (North and South, from here on referred to as Dhaka CCs), and an additional low-income site from the Greater Dhaka Statistical Metropolitan Area (SMA), following a two-stage stratification design. In the first stage primary sampling units (PSUs) were selected using probability proportional to size (PPS). In the second stage, all of the households in the PSU were listed, and 20 households were selected for interviewing. Strata were used in both stages.

First Stage: Selection of the PSUs

Low-income PSUs were defined as nonslum census enumeration areas (EAs), in which the small-sample area estimate of the poverty rate is higher than 8 percent (using the 2011 Bangladesh Poverty Map). The sampling frame for these low-income areas in the Dhaka CCs and Greater Dhaka is based on the population census of 2011. For the Dhaka CCs, all low-income census EAs formed the sampling frame. In the Greater Dhaka area, the frame was formed by all low-income census EAs in specific *thanas*¹ where World Bank project locations were located.

Three strata were used for sampling the low-income EAs. These strata were defined based on the poverty head-count ratios. The first stratum encompasses EAs with a poverty headcount ratio between 8 and 10 percent; the second stratum between 11 and 14 percent; and the third stratum, those exceeding 15 percent. The number of EAs selected from each stratum in the Dhaka CCs and the greater Dhaka areas is shown in Table 1.

Slums were defined as informal settlements that were listed in the Bangladesh Bureau of Statistics' slum census from 2013/14. This census was used as sampling frame of the slum areas. Only slums in the Dhaka City Corporations are included. Again, three strata were used to sample the slums. This time the strata were based on the size of the slums. The first stratum comprises slums of 50 to 75 households; the second 76 to 99 households; and the third, more than 100 households. Small slums with fewer than 50 households were not included in the sampling frame. Very small slums were included in the low-income neighborhood selection if they are in a low-income area. The number of slums selected from each stratum is shown in the bottom panel of Table 1.

Table 1 Sample Allocation

Strata	Number of EAs	Selection criteria (poverty headcount ratio or number of households)	Criteria 1 (male and female working-age member)	Criteria 2 (only female working-age member)	Criteria 3 (only male working-age member)
<i>Low-income areas in the Dhaka CCs</i>					
1	20	8%-10%	328	57	15
2	18	11%-14%	181	33	7
3	5	15% +	82	15	4
<i>Low-income area in greater Dhaka</i>					
3	5	15% +	180	31	9

¹ Thana, or *Upazila*, is an administrative unit in Bangladesh, functioning as a subunit of a district.

<i>Slums in the Dhaka CCs</i>					
1	3	50 to 75	52	6	2
2	8	76 to 99	137	22	1
3	7	100+	116	21	3

Altogether, the DIGNITY survey collected data from 67 PSUs. Locations of the PSUs are shown in **Error! Reference source not found..**

Second Stage: Selection of the Households

In each sampled PSU a complete listing of households was done to form the frame for the second stage of sampling: the selection of households. When the number of households in a PSU was very large, smaller sections of the neighborhood were identified, and one section was randomly selected to be listed. The listing data collected information on the demographics of the household to determine whether a household fell into one of the three categories that were used to stratify the household sample:

- i) households with both working-age male and female members;
- ii) households with only a working-age female;
- iii) households with only a working-age male.

Households were selected from each stratum with the predetermined ratio of 16:3:1. In some cases there were not enough households in categories (ii) and (iii) to stick to this ratio; in this case all of the households in the category were sampled, and additional households were selected from the first category to bring the total number of households sampled in the PSU to 20.

Sampling Weights and Probability of Selection

The probability of selection of the sample was calculated separately for each stage and enumeration areas (EA) using the following specification:

$$P_{ij}^1 = n_j \frac{M_{ij}}{\sum M_{ij}}$$

where the probability of the i^{th} EA being selected from the j^{th} stratum is presented by P_{ij}^1 . n_j represents the number of selected EAs within each stratum. The number of households in the i^{th} EA in the j^{th} stratum is represented by M_{ij} while $\sum M_{ij}$ represents the total number of households in the stratum.

The probability of the second stage of the selection process is estimated using the following specification:

$$P_{ij}^2 = \frac{h_{tij}}{H_{tij}}$$

where the probability of a household being selected for the sample is represented by P_{ij}^2 . h_{tij} is the number of households of type t in EA i in stratum j selected to be surveyed from a total number of households (H_{tij}) within each EA and the particular category.

The overall likelihood of a particular household being selected from a particular stratum is therefore represented as the product of the two aforementioned probabilities, calculated as follows:

$$P_{ij} = P_{ij}^1 \times P_{ij}^2$$

The weight is subsequently constructed as the inverse of the likelihood of a particular household being selected ($1/P_{ij}$).