

Design of a Multi-Stage Stratified Sample for Poverty and Welfare Monitoring with Multiple Objectives

A Bangladesh Case Study

Faizuddin Ahmed

Dipankar Roy

Monica Yanez-Pagans

Nobuo Yoshida



WORLD BANK GROUP

Poverty and Equity Global Practice Group

March 2017

Abstract

This paper describes the design of a multi-stage stratified sample for the Bangladesh Household Income and Expenditure Survey 2016/17. This survey instrument will be used by the Government of Bangladesh to estimate reliable poverty and welfare statistics at three different levels: (i) annual estimates at the district level, (ii) quarterly estimates at the national level, and (iii) annual estimates at the division level for urban and rural areas. The sample for this

survey was designed to achieve these three objectives. The paper explains how the three objectives are prioritized and how inconsistencies in achieving more than one objective can be reconciled. Further, the paper modifies the standard formulas to estimate the optimal sample size and the allocation of the sample across strata by explicitly taking into consideration the effect of clustering in the sample.

This paper is a product of the Poverty and Equity Global Practice Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The authors may be contacted at nyoshida@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Design of a Multi-Stage Stratified Sample for Poverty and Welfare Monitoring with Multiple Objectives: A Bangladesh Case Study

Faizuddin Ahmed
Dipankar Roy
Monica Yanez-Pagans
Nobuo Yoshida

JEL Classification: C81, C83, I32

Keywords: stratified clustered sampling, household surveys, poverty measurement

*Corresponding author: Nobuo Yoshida (nyoshida@worldbank.org), The World Bank. Faizuddin Ahmed (fahmed3@worldbank.org) and Monica Yanez-Pagans (myanezpagans@worldbank.org) are also at The World Bank and Dipankar Roy (dipankar.roy@bbs.gov.bd) is at the Bangladesh Bureau of Statistics. The authors are very grateful to Juan Munoz and Kristen Himelein for their advice and suggestions during the preparation of this paper. All errors and mistakes remaining in this paper are solely ours.

1. Introduction

In recent decades, the need for detailed demographic and socioeconomic data to measure living standards in developing countries has increased dramatically. Data used for measuring living standards are traditionally collected through household income and expenditure surveys or through integrated household surveys. The data these surveys generate are critical for monitoring welfare and informing the design of policies to reduce poverty and promote equitable distribution of economic growth. In most developing countries, integrated household surveys are a central component of national statistical systems, which provide key inputs in social and development policy making and national accounts.

Collecting expenditure or income data for welfare monitoring is expensive and time consuming. Generally, households are visited several times and need to report on their food and non-food expenditures for often more than 50 items over the past days or weeks either using a recall or a diary method. The complexity of the data collection can cause households to make mistakes when reporting their incomes or consumption or enumerators to make mistakes when recording their responses. As a result, collecting consumption or income data from an entire population in a country to monitor trends of key welfare statistics such as poverty headcount rates or mean income is very time-consuming and resource intensive to carry out.

To overcome these challenges, National Statistical Offices (NSOs) conduct household income and expenditure surveys for only a sample of the population. The selection of a sample, which is called “sampling,” is an important subject in statistics and needs to be done carefully so that the select sample of individuals or households represent the entire population and data collected from the sample can therefore be used to estimate characteristics of the population with a reasonable level of precision.

Sampling provides two clear advantages – substantially reducing the cost and time needed to collect data and producing statistics that are representative of the entire population. For example, a sample of a couple of thousand households or individuals is often enough to estimate nationally representative statistics of a country with a population of one hundred

million. However, the risk of sampling is also large – if the selection of households or individuals is not properly done, statistics constructed from the sample do not represent the characteristics of the population of interest. This is the reason why sampling is so critical when NSOs are designing a survey.

This paper describes the sampling design for the 2016/17 Bangladesh Household Income and Expenditure Survey (HIES). The HIES, a large household survey collected by the Bangladesh Bureau of Statistics (BBS), is the main government instrument for monitoring welfare and living conditions in Bangladesh. The HIES is a well-established sample household survey in Bangladesh and has been conducted about every five years since 1991-92, with the latest three rounds collected in 2000, 2005, and 2010.¹

BBS decided to redesign the sample of the HIES 2016/17 to produce more frequent and more disaggregated poverty estimates than the previous rounds. Specifically, the samples of the latest three rounds of the HIES survey are designed to provide reliable annual poverty estimates at the division level for both urban and rural areas.² But in the HIES 2016/17, BBS decided to produce reliable annual district poverty statistics and quarterly national poverty statistics in addition to annual poverty estimates at the division level for urban and rural areas. As a result, BBS needed to change the design of sampling for the HIES 2016/17 significantly. For example, the sample size for the HIES 2016/17 needed to increase almost fourfold compared to the HIES 2010. This substantial increase in the sample size also forced a revision of the sampling frame to accommodate the larger number of household field interviews.

The main challenge when designing the HIES 2016/17 was how to achieve multiple goals with one sampling design. As mentioned above, the objective of the HIES 2016/17 is three-fold: (i) producing reliable annual poverty and welfare statistics at the district level; (ii) producing reliable quarterly poverty and welfare statistics at the national level; and (iii) producing reliable

¹ In addition to collecting detailed household consumption and income data, the HIES collects detailed information on household demographics, education, health, livelihoods and economic activities, quality of housing, asset ownership, migration, social safety-nets, micro-credit, access to services, and household shocks and coping strategies.

² Divisions are the higher level of geographical divisions in Bangladesh. As of 2010, there were a total of seven divisions in the country, including Dhaka, Chittagong, Barisal, Khulna, Sylhet, Rangpur, and Rajshahi.

annual poverty and welfare statistics at the division level for urban and rural areas. A potential challenge is that the optimal sample for objective (i) is not necessarily identical to that needed to achieve objectives (ii) or (iii). Therefore, the design of the sample requires careful thinking on how to achieve the multiple objectives, which are not necessarily mutually consistent, within a single sample. This paper describes how the sample of the HIES 2016/17 was designed to achieve the aforementioned three objectives.

This paper also reviews the formulas that are often used by NSOs and development practitioners to calculate the sample size and the optimal allocation of the sample across strata. The HIES 2016/17 follows a stratified clustered sampling. Clustered sampling is used when the sample is selected in multiple stages. That is, in a first stage, clusters of households or individuals are selected randomly and then households or individuals are selected from the clusters in a second stage. This is less efficient than simple random sampling (selecting household or individuals in one stage directly) in terms of minimizing sampling errors, but it is more cost-effective as selected households or individuals are physically located relatively close to each other in the same cluster. This paper shows that the standard formulas used by practitioners for defining the sample size and the allocation of the sample across strata assume stratified random sampling rather than stratified clustered sampling. As a result, the sample size and the allocation of the sample can be too optimistic in terms of minimizing the sampling errors.

This paper is organized as follows. Section 2 reviews key literature on optimal sample allocation for stratified random samples and stratified clustered samples. Section 3 describes objectives of the survey, data used to design the HIES 2016/17 sample, sampling frame, stratification and domains of study, selection of target variables and accuracy statistics for sample performance evaluation, determination of the sample size, and allocation of Primary Sampling Units (PSUs) across strata. Section 4 summarizes evaluation of the proposed sampling design. Section 5 analyzes robustness of the proposed sample design. Section 6 describes survey sampling weights computation. Section 7 presents the main conclusions of the paper.

2. Background

2.1. HIES rounds sampling design prior to the HIES 2016/17

The HIES rounds have been central for monitoring poverty and welfare indicators in Bangladesh. Prior to the HIES 2016/17, the HIES rounds were designed to produce annual poverty estimates at the division level for urban and rural areas and Statistical Metropolitan Areas (SMAs) (see Munoz 1998 and 2004). The sample size of the HIES gradually increased over time and, as of the latest round completed in 2010, the sample size was 12,240 households. To avoid the effects of seasonality, the sample was proposed to be randomly split across quarters – a full round of the HIES data collection takes 12 months – but there was never an explicit attempt by BBS to produce quarterly poverty estimates before the HIES 2016/17.

All of the previous HIES rounds follow a stratified clustered sampling design – each of urban, rural, and SMA areas in a division forms a stratum. Primary Sampling Units (PSU) are constructed based on Census of Population and Housing Enumeration Areas and represent a cluster. PSUs are selected with probability proportional to size (PPS) and 20 households from each PSU are selected using systematic sampling.

It is worth noting that PSUs for all of the previous rounds of the HIES were selected from the Integrated Multiple-Purpose Sample (IMPS). The IMPS is a master sample updated after each Census of Population and Housing, and which has been used in Bangladesh as the sampling frame for the selection of Primary Sampling Units (PSUs) in all previous rounds of the HIES and other surveys such as the Labor Force Surveys and Demographic and Health Surveys.

The HIES 2016/17 deviates from the sampling design used in the previous rounds of the HIES in several ways. First, the objectives of the HIES 2016/17 have changed significantly from the one used to design the sample in previous rounds. Second, the HIES 2016/17 did not use the IMPS for the selection of PSUs. Third, the formulas used for the estimation of the optimal sample size and the allocation of the sample across strata were derived to explicitly take into account the clustered design of the sample.

2.2. Formulas for the optimal sample size and the optimal allocation of samples across areas

In his influential book on survey sampling, Cochran (1977) presents two key formulas for designing any stratified simple random sampling. The first formula determines the minimum sample size required to achieve a certain level of statistical reliability regarding the relationship between one statistic and another targeted variable. The second formula, based on Neyman's (1934) original allocation rule, determines the optimal allocation of the sample across strata.³

These two key formulas, presented in Cochran's (1977) book, assume that the sample follows a simple stratified random sampling design.⁴ However, most surveys that we are interested in—that is, household income and expenditure surveys and integrated household surveys designed to measure living standards and monitor poverty and welfare patterns—are not a simple stratified random sample, but rather use multi-stage stratified clustered sampling designs.⁵ Therefore, the original Cochran (1977) formulas for designing a sample for a living standards household survey is generally not appropriate and would result in an inadequate determination of the sample size and a sub-optimal allocation of the sample across strata. Nevertheless, the formulas are widely used for the selection of the sample size and allocations across strata.

In this paper, we propose a modification to Cochran's sample size formula to estimate adequate sample sizes for multi-stage stratified clustered surveys. Cochran's (1977) original formula, building on work by Cornfield (1951) and Kish (1965), estimates sample size by calculating the minimum sample size required to achieve a certain level of precision (plus or minus x

³ The Neyman's (1934) original allocation rule minimizes the variance of the sample mean of a targeted variable given a fixed pre-established sample size assuming that the cost of sampling in each stratum is constant.

⁴In sampling, stratification is the process of dividing members of a population into homogeneous subgroups before selecting the sample. This provides two main advantages over simple random sampling (SRS). First, it ensures that different subgroups of the population are properly represented in the sample. Second, it allows to gain statistical precision in the estimation of the variance of the mean of a targeted variable because the variability within the subgroups is lower compared to the variations across the entire population.

⁵ A stratified clustered sampling design consists in selecting the sample using multiple stages. In the first stage, the population is divided into groups (or clusters). In the second stage, one or more clusters are chosen at random and for everyone within the chosen cluster a second sampling unit (or ultimate sampling unit) is selected.

standard deviations) when comparing one variable to another. To modify this formula, we first calculate the sample size required to achieve the desired level of precision for a simple random sample (SRS). We then adjust this formula by the design effect (DEFF), which in this application measures the efficiency of a stratified clustered sample compared to SRS. We also derive two different formulas for estimating the sample size based on whether the targeted variable being studied is “continuous” or “dichotomic”.

To determine optimal allocation of the sample across sampling strata, we also modify Neyman’s (1939) original allocation rule, which incorporates the survey design effect and can therefore optimally allocate for multi-stage stratified clustering. The derivation of this formula follows Neyman’s (1939) original closely and consists in solving an optimization problem that minimizes the stratified clustered sampling variance of the targeted variable of interest subject to a fixed pre-established sample size that is determined using the modified Cochran (1977) formula described above. The solution to this optimization problem gives us a more general optimal rule for the allocation of the sample across strata. This rule can be used both for multi-stage stratified clustered samples and for stratified simple random samples.

3. Sampling strategy

3.1. Objective of the survey

As mentioned before, the main objective of the HIES 2016/17 is to produce reliable poverty statistics, measured by the poverty headcount ratio, at three different levels: (i) annual poverty estimates at the district level; (ii) annual poverty estimates at the division level for urban and rural areas; and (iii) quarterly poverty estimates at the national level.⁶ A secondary objective of the HIES 2016/17 is to understand the rapid urbanization occurring in Bangladesh, and to

⁶ The poverty headcount ratio is defined as the proportion of people living below the official national poverty line. In Bangladesh, there are two official government poverty measures: the poverty headcount ratio and the *extreme* poverty headcount ratio. The poverty headcount ratio is based on a set of division-level poverty lines, defined separately for urban and rural areas, called “upper poverty lines”. In contrast, the extreme poverty headcount ratio is based on a set of division-level poverty lines, also defined separately for urban and rural areas, called “lower poverty lines”. Both upper and lower poverty lines include a food and non-food component, but they differ in how the non-food component is defined. For details on how the poverty lines and poverty estimates are produced, see World Bank (2013).

assess how urbanization is linked to poverty and welfare measures. The sample for the HIES 2016/17 is explicitly designed to produce reliable poverty estimates at the three levels, and is implicitly designed to produce an urban sample large enough to understand Bangladesh's urbanization patterns.

3.2. Data used for the sampling design

The data used for the design of the sample for the HIES 2016/17 come from the latest round of the HIES collected in 2010 and include 12,240 households. The HIES 2010 sample was designed to provide reliable poverty and welfare statistics only at the division level for both urban and rural areas.

Designing a household survey sample often entails using data from a previous household survey to estimate the key parameters needed to determine the sample size and how to allocate it across strata. In the case of Bangladesh, the latest round of the HIES was collected six years ago. Given the rapid growth observed in the country over the past decade, however, it is not clear whether these data provide a good representation of the actual living standards in the country or not. Still, previous household survey rounds usually provide the best source of information to design a sample.

The main goal of Bangladesh's HIES 2016/17 survey was to estimate reliable poverty rates at more disaggregated levels than the HIES 2010 survey.⁷ Standard errors of mean poverty rates estimated at the district-level poverty using the HIES 2016/17 might end up being a little bit higher than what is predicted under the proposed sampling design because reliable district-level poverty rates cannot always be constructed based on the HIES 2010. That is, some of the parameters estimated from the HIES 2010 to design the sample for the HIES 2016/17 might not provide accurate information of living standards in 2016/17. To address this concern, we analyze the robustness of the proposed sampling design for the HIES 2016/17 in several ways. The results of this robustness analysis are presented in Section 5.

⁷ The HIES 2016/17 survey instrument (i.e. questionnaire) is very similar to the one used in the HIES 2010 and field logistics are expected to follow similar protocols in terms of the survey implementation.

3.3. Sampling frame

In the HIES 2016/17, the primary sampling units (PSUs) are the Census Enumeration Areas (EAs) used for Bangladesh's 2011 Census of Population and Housing. EAs are small geographic areas consisting of an average of 110 households. The sampling frame for the selection of PSUs consists in the list of all EAs covering people residing in non-institutional dwelling units (98.5 percent of the population in Bangladesh).⁸ This is an important departure from the HIES 2010, where the sampling frame for the selection of PSUs consisted in the list of all EAs in the Bangladesh IMPS that excluded some geographic areas, such as urban slums.⁹ For the HIES 2016/17, the IMPS could not be used because the most recent version of IMPS based on the 2011 Census of Population and Housing included only 2,012 EAs, an insufficient number to serve as a sampling frame for this survey. Table 1 summarizes the main characteristics of the sampling frame for the HIES 2016/17:

Table 1
Sampling frame for the HIES 2016/17

Area	Number of households	Number of EAs	Mean number of households in EA
Rural	24,588,843	228,202	108
Urban	5,020,815	43,872	114
City Corporation	2,253,737	21,266	106
Total	31,863,395	293,340	109

Source: Bangladesh Bureau of Statistics, 2011 Census of Population and Housing

3.4. Stratification, sub-stratification, and domains of study

In Bangladesh, the eight divisions are the first-level administrative geographical partitions of the country. As of 2016, the country had eight divisions: Barisal, Chittagong, Dhaka, Khulna,

⁸ The list of all EAs in the 2011 Census of Population and Housing includes 296,718 clusters, of which 293,340 EAs cover non-institutional dwelling units.

⁹ The IMPS is a master sample design that is updated after each Census of Population and Housing, and which has been used in Bangladesh as the sampling frame for the selection of PSUs in the HIES 2000, 2005, and 2010. Using a master sample frame for multiple surveys makes regular survey data collection more efficient. For details on how the Bangladesh IMPS Design was produced, see Maligalig and Martinez (2013).

Mymensingh, Rajshahi, Rangpur, and Sylhet. Each division is subsequently divided into 64 districts, or *zilas*. Each district is further subdivided into smaller geographic areas with clear rural and urban designations. In addition, urban areas in the main divisions of Chittagong, Dhaka, Khulna and Rajshahi are classified into City Corporations (CCs), or Metropolitan Areas and Other Urban areas. Table 2 shows the distribution of the population by division and across urban, rural, and CCs areas.

Table 2
Distribution of population by division (%)

Division	Number of districts	Rural	Urban	City Corporation	Total (%)
Barisal	6	5.3	1.1	-	6.3
Chittagong	11	14.3	1.6	3.2	19.1
Dhaka	13	13.6	1.7	10.6	25.9
Mymensingh	4	6.1	0.8	-	6.8
Khulna	10	9.2	1.4	1.3	11.9
Rajshahi	8	10.3	1.6	0.6	12.5
Rangpur	8	9.8	1.5	-	11.3
Sylhet	4	5.2	1.0	-	6.2
Total	64	73.7	10.6	15.7	100.0

Source: Household Income and Expenditure Survey (2010)

The sampling design for the HIES 2016/17 includes two different levels of stratification:

- (i) *Division-Rural/Urban/CC*: The eight administrative divisions in Bangladesh are initially used to construct 16 sampling strata: one urban and one rural for each division. In addition, each urban sampling stratum in the four main divisions of Chittagong, Dhaka, Khulna and Rajshahi is further divided into two sampling strata: City Corporation and Other Urban. This brings the total number of sampling strata to 20: eight in rural divisions, eight in urban divisions, and four main CCs.
- (ii) *District- Rural/Urban/CC*: As explained in detail below, PSUs in the HIES 2016/17 are allocated at the district level. Therefore, the sample is implicitly sub-stratified at the district level. Since there are a total of 64 districts in Bangladesh, the sample design includes a total of 132 sub-strata: 64 urban, 64 rural, and four main CCs.

The poverty and welfare domains of study for the HIES 2016/17 are slightly different from the sampling strata defined for the sampling design. Although the sample includes 20 sampling strata, official poverty statistics will only be produced at the following levels: (i) annual district level; (ii) quarterly national level; and (iii) annual division level for urban and rural areas (but not for CCs separately).

3.5. Selecting target variables and accuracy statistics

One of the first decisions when designing a sample is determining key target variables to study. That then determines the sample size and allows an assessment of the ability of the sample to achieve a certain level of precision to estimate statistics related to the key target variables.¹⁰ In the case of the HIES 2016/17, the main objective is to generate reliable poverty estimates for Bangladesh at different levels of disaggregation. Poverty in Bangladesh is estimated based on per capita nominal consumption expenditures. As a result, three different target variables are used to design the HIES 2016/17 sample: (i) nominal per capita expenditure; (ii) nominal total household expenditure; and (iii) poverty headcount ratio.¹¹ Table 3 presents summary statistics for the three target variables by division.

¹⁰ In the case of integrated household surveys with multiple objectives, a decision needs to be made as to what the key target variables are for design of the sample. The number of target variables needs to be kept at a minimum, and must consist of closely related indicators with relatively similar levels of prevalence in the population of interest. Otherwise the proposed sample size would grow dramatically.

¹¹ Nominal per capita expenditure is used as a target variable rather than real per capita expenditure in order to adhere to the traditional convention followed by BBS for the design of the samples for the previous three rounds of the HIES collected in 2000, 2005, and 2010.

Table 3
Summary statistics of target variables by division

Division	Nominal per capita expenditure (in 2010 taka)	Nominal total household expenditure (in 2010 taka)	Poverty headcount ratio (%)
Barisal	2,154 (78)	10,708 (425)	39.35 (3.33)
Chittagong	2,887 (144)	15,765 (797)	26.20 (1.99)
Dhaka	2,870 (129)	14,008 (703)	26.08 (1.64)
Mymensingh	1,827 (60)	9,310 (334)	47.36 (3.43)
Khulna	2,182 (65)	10,335 (323)	32.07 (2.28)
Rajshahi	2,231 (59)	10,588 (345)	29.67 (2.11)
Rangpur	1,937 (63)	9,434 (343)	42.32 (3.22)
Sylhet	2,184 (93)	13,936 (727)	28.10 (3.02)

Note: Data for the estimation of these statistics comes from the HIES 2010. All estimates reported in this table have been produced using population level weights to match official statistics. For the proposed sampling design, however, we use household level weights. Standard errors adjusted for survey design are reported in parenthesis.

A second important decision when designing a sample is selecting accuracy statistics. These are used to evaluate the performance of the proposed sample design to achieve a certain level of precision in the estimation of these statistics for the target variables of interest. The most common accuracy statistics are “standard errors” (SE) of the mean and “relative standard errors” (RSE) of the mean; most household surveys measuring living conditions try to achieve SEs of the mean of the target variable(s) that are no more than 5 percent, and RSEs of the mean that are no more than 10 percent.¹² For HIES 2016/17, we use three accuracy statistics to assess

¹² The RSE of a sample mean is the SE of the mean divided by the mean and then expressed as a percentage.

the performance of the sample: (i) RSEs of less than 10 percent from the mean nominal per capita expenditure for each district and division (urban/rural), as well as for each quarter at the national level; (ii) RSEs of less than 10 percent from the mean nominal household expenditure for each district and division (urban/rural), and for each quarter at the national level; and (iii) SEs of less than 5 percent of the poverty headcount ratio for each district and division (urban/rural), and for each quarter at the national level. Final sample size determination, described in the next section, is based on the nominal per capita expenditure target variable, but we evaluate performance of the sample in terms of these three accuracy statistics to confirm that the final sample provides reliable poverty and welfare estimates at all these levels.

It is important to highlight that RSEs are not used to evaluate the performance of the sample in terms of poverty rates. Instead, we use SEs for poverty rates. This responds to two reasons. First, the value of the RSE can go up very quickly when the poverty rates are small as it is the case in several districts in Bangladesh. For instance, in 2010 poverty rates of Kushtia and Noakhali were as low as 3.6 percent and 9.6 percent, while the poverty rates of the poorest districts of Kurigram, Barisal, and Shariatpur were as high as 63.7 percent, 54.8 percent, and 52.6 percent respectively. This large heterogeneity across districts in terms of poverty rates results in misleading results when interpreting RSEs. The poorest districts with elevated poverty rates will have relatively low RSEs, while better-off districts with low poverty rates will have relatively large RSEs. Therefore, the RSE is not a good accuracy statistic to use for target variables whose distributions show high levels of dispersion. Second, poverty measurement has traditionally focused on SEs to evaluate the precision of the poverty rates, which aligns more closely with the standard literature on poverty measurement and its practical applications.

3.6. Estimation of the sample size

The HIES 2016/17 uses a stratified, two-stage clustered sampling design. In the first stage, PSUs are equally allocated across the 64 districts in Bangladesh: 36 PSUs are allocated to each of the 64 Bangladesh districts, bringing the total number of PSUs in the sample to 2,304 (36 PSUs*64 districts). An equal number of PSUs are allocated across districts because the HIES

2016/17 aims to produce poverty and welfare statistics with similar accuracy across all districts. In addition, we expect that achieving low RSEs and SEs in the estimation of district-level poverty and welfare will be more difficult than at the division level, as well as at the quarterly level. Therefore, allocating an equal number of PSUs across districts ensures that the final sample is evenly distributed across all districts.

In the second stage, we randomly select 20 households from each PSU, resulting in 720 households allocated to each of the districts (36 PSUs*20 households). This brings the total targeted sample size to 46,080 households (2,304 PSUs*20 households). Assigning 36 PSUs to each district was determined as follows. First, this is the minimum sample size needed to achieve an RSE of less than 10 percent for the mean nominal per capita expenditure for each district and division (urban/rural), and for each quarter at the national level with 95 percent confidence (our selected accuracy statistics to determine sample size). To arrive at this number (36), we use a formula (derived in Annex 3) that is a simple extension of the original formula proposed by Cochran (1977) to determine sample size for continuous target variables under simple random sampling when controlling for a desired RSE.¹³ This simple extension, which introduces survey design effects to the original formula, results in the following expression:

$$n = \left(\frac{Z_{\alpha/2} * CV_{SRS}(\bar{y})}{r(\bar{Y})} \right)^2 * DEFF \quad (1)$$

where n is the minimum sample size required for allocation to each district in order to achieve a certain level in the accuracy statistic $r(\bar{Y})$ associated with the targeted variable \bar{y} ; $CV_{SRS}(\bar{y})$ is the coefficient of variation of the targeted variable estimated under the assumption of simple random sampling; $DEFF$ is the design effect of the target variable; and $Z_{\alpha/2}$ is the critical value of a standard normal distribution with $\alpha\%$ confidence level.¹⁴

¹³ Derivation of the formula to determine optimal sample size for proportions when one is interested in controlling for SE instead is provided in Annex 3.

¹⁴ This formula is derived under the assumption that the targeted variable \bar{y} is normally distributed.

In the case of the HIES 2016/17, $r(\bar{Y})$ is the 10 percent RSE desired for the mean total household expenditure estimated at the district level; $DEFF$ is the average design effect of the target variable across all districts; $CV_{SRS}(\bar{y})$ is the coefficient of variation of total household expenditure estimated at the national level; and $Z_{\alpha/2}$ is set at 1.96, which corresponds to a significance level of 5 percent.¹⁵ We use a national estimate of the CV – necessary to estimate the sample size needed to achieve the desired accuracy – because we cannot always estimate district level CVs with enough precision based on the HIES 2010 data. Therefore, the above calculations assume that all districts have the same DEFFs and CVs for the target variable of interest. It is possible, that the final HIES 2016/17 sample might estimate district level welfare estimates with RSEs slightly larger than 10 percent, depending on the how close the CVs and DEFFs of the target district variables are to their actual values.

It is also important to note that we focus on district-level subsamples to estimate sample size, because, as mentioned earlier, we expect that achieving accuracy statistics at the district level will be more difficult than at the higher division and quarterly levels of aggregation. By substituting all values in equation (1), we find that the sample size needed is 715 households for each district in order to obtain RSEs of mean total household expenditure below 10 percent. This brings the preliminary total sample size to 45,760 households (715*64 districts) for the HIES 2016/17. However, we allocate 720 households to each district to fit pre-established field protocols used by BBS in the HIES 2010, and to facilitate field work and survey implementation management. More specifically, 720 households is divisible by 36 (i.e. the number of PSUs ultimately allocated to each district), which allows easy allocation of 20 households per PSUs. Also, two enumeration teams can easily cover the 36 PSUs in each district over 12 months without having to move to a different district (each team of two enumerators will visit 18 PSUs each over one year). Lastly, the number PSUs, 36, allocated to each district is a multiple of 4, which allows dividing into quarters.

¹⁵ To estimate $DEFF_h$, for each district, we use the Stata `'svy: mean, over(district)'` command combined with the `'estat effects, srssubpop'` and `'estat sd, srssubpop'` routine. An important note to mention here is that Stata uses a linear approximation to estimate survey design effects, which opens the possibility that for some district-stratum pairs we might observe $DEFF$ values that are lower than 1. When this happens, the strategy adopted was to replace the estimated $DEFF$ values with $DEFF=1$.

Although we can only focus on one accuracy statistic to estimate sample size, we also assess whether the sample also achieves the other proposed accuracy statistics in Section 4.

Table 4 summarizes the key parameters of the HIES 2016/17 sample design. The final sample includes all 64 districts in Bangladesh and 20 strata (eight divisions*urban/rural=16, plus four main CCs). Each district is assigned 36 PSUs, an equal number for each district that we derived by modifying Cochran’s formula to estimate the sample size for a target continuous variable for a stratified clustered sample design. The target variable used to estimate the sample size for the HIES 2016/17 is nominal total household expenditure, which follows the convention used in the three previous rounds of the HIES. The accuracy statistics used to assess the performance of the sample is an RSE within 10 percent for each district. The cluster is set at 20 households per PSU, which also follows the convention used in previous HIES rounds. The total number of PSUs is 2,304 and the targeted sample size is 46,080. In terms of field logistics, these numbers suggest that a total of 128 teams will be needed to collect field data, with each team consisting of two people (enumerators) that will visit 18 PSUs over 12 months.

Table 4
HIES 2016/17 summary of sampling design

Description	Number
Number of districts	64
Number of strata	20
Number of PSUs in each district	36
Number of households in each PSU	20
Total number of PSUs in sample	2,304
Total sample size	46,080
Total number of teams	128
Total number of enumerators	256

3.7. Allocation of PSUs across sub-strata within districts

In the second stage, we allocate the 36 PSUs assigned to each district across rural, urban, and CC using a modified version of the Neyman’s allocation rule (“optimal allocation rule”) that

minimizes the standard error of a target variable while explicitly accounting for the stratified clustered sample design.¹⁶ The standard Neyman allocation rule provides an optimal allocation of PSUs when the sample follows a stratified random sample design (i.e. squared root of the design effect of $DEFT_h=1$ for each stratum h), and when finite sample correction is zero (i.e. $f_h=0$). For stratified clustered samples, the standard Neyman allocation requires modification to achieve an optimal allocation. The modified Neyman allocation rule is given by the following expression:

$$n_h^* = \frac{W_h S_h DEFT_h \sqrt{(1-f_h)}}{\sum_{h=1}^L W_h S_h DEFT_h \sqrt{(1-f_h)}} * n \quad (2)$$

where h is a sub-stratum or stratum within a district; n_h^* is the optimal sample allocation in sub-stratum h ; n is the total number of households in a district; W_{hd} is the population share in each sub-stratum h ; S_{hd} is the standard deviation of the target variable of interest; DEFT is the square root of the design effect of the target variable of interest; and $(1 - f_h)$ is a finite sample correction factor. For the sample design of the HIES 2016, we assume finite sample correction is zero.¹⁷

All parameters required to estimate the optimal allocation of PSUs across sub-strata are estimated based on the HIES 2010 data. Since the cluster size in the HIES 2016/17 is identical to the cluster size used in the HIES 2010 (20 households), the intra-cluster correlation coefficients used to estimate the *DEFTs* can be computed directly from the HIES 2010 data.¹⁸

After deriving the modified Neyman allocation to produce reliable poverty statistics at the district level, the following five additional adjustments are made to achieve the other two goals

¹⁶ The Neyman allocation rule is derived from an optimization problem consisting in minimizing the variance of a target variable subject to a pre-determined fixed sample size. The full derivation of the modified Neyman allocation rule used for the allocation of PSUs in the HIES 2016/17 is presented in Annex 2.

¹⁷ The finite sample correction factor must be taken into account when the sample size corresponds to a large percentage of the size of the population of interest. This condition rarely holds in practice.

¹⁸ The performance of the sample is directly affected by the decision to fix *ex-ante* the cluster size in each PSU to 20 households. This follows the decision of the BBS to follow field data collection protocols closely aligned with those used in the three previous rounds of the HIES collected in 2000, 2005, and 2010. This paper does not assess the performance of the sample under the selection of different cluster sizes.

of the sample design – producing accurate quarterly poverty estimates at the national level and producing accurate annual estimates at the division level for urban and rural areas. First, we round up numbers calculated using the modified version of the Neyman equation. Second, although the modified Neyman allocation is supposed to minimize standard errors of poverty rates at the district level given the original allocation of the sample across districts, this would fail to assign sub-samples of sufficient size to some CCs. Since producing reliable poverty statistics at the stratum level (urban, rural and CC areas of each division) is also a goal of the HIES 2016/17, we deviate from the original modified Neyman allocation and move some of the PSUs from urban and rural areas to CCs. This deviation will slightly increase the standard errors of poverty statistics for the district level including the CCs, but it will overall reduce poverty statistics at the stratum level.¹⁹ Third, we allocate a minimum of four PSUs to each stratum of a district so that at least one PSU is allocated to each quarter. Fourth, the HIES 2010 sample did not assign any PSU to the urban Khulna district, but the minimum of four PSUs are assigned to this stratum in the HIES 2016/17.²⁰

Lastly, to produce nationally representative poverty rates for each quarter, the sample needs to be allocated to the four quarters of the data collection uniformly. Munoz (2004) suggested that to avoid seasonality, the entire sample should be randomly split into four quarters. However, randomly allocating the sample across quarters does not guarantee a full removal of seasonality. In particular, if the number of PSUs allocated to a stratum in a specific district is a multiple of four, then we can allocate the same number of PSUs to all quarters. Yet, if the number of PSUs allocated to a stratum in a specific district is not a multiple of four, there is

¹⁹ Note that we are not changing the sample size of districts. This is certainly suboptimal. Ideally, we can shift samples from districts with very low projected standard errors to districts with the problematic strata; however, how to design such an allocation requires more careful analysis and is beyond the scope of this paper.

²⁰ The HIES 2010 sample frame was based on the IMPS, which in urban areas was constructed using the concept of Statistical Metropolitan Areas (SMA). The concept of SMA was replaced by the concept of Rural/Urban/CC (RUC) in the 2011 Census of Population and Housing. Of the 64 districts, only in three does the old SMA concept not match perfectly with the new RUC: in the districts of Gazipur and Narayanganj in the Dhaka division (districts 33 and 67, respectively), and the district of Khulna in the Khulna division. For Gazipur and Narayanganj, a perfect match can be achieved by replacing all SMA areas to Other Urban areas. For Khulna district, however, a match is not straightforward as the SMA area was divided into CC and Other urban areas. In addition, all of the PSUs from the Khulna district available from the HIES 2010 come from SMA areas, and there is therefore no baseline for Other Urban Khulna district. To conduct all the analysis for the design of the sample for the HIES 2016/17, we use the same sub-stratum values for both CC and Other Urban areas.

no way we can allocate the same number of PSUs to all quarters. As a result, the distribution of the sample across strata in these districts might differ across quarters.

Now, let’s look at Table 5, which illustrates the challenges faced when trying to allocate PSUs across urban and rural strata within a district. For example, 26 PSUs are assigned to the rural stratum and 10 to the urban stratum. 26 is not a multiple of four, so we need to assign the closest integers to each of the quarters (6.5) which are 6 and 7. For example, Table 5 assigns 7 PSUs to quarters 1 and 2 and 6 PSUs to quarters 3 and 4. We repeat the same procedure for the urban stratum. This allocation in Table 5 is problematic for two reasons: (i) the share of rural areas differs across quarters; and (ii) the total sample size of quarters 1 and 2 is different from that of quarters 3 and 4. The first issue is problematic because poverty rates can differ across quarters if poverty rates in urban areas are different from those in rural areas even if true poverty rates do not actually change. The second issue is problematic because the precision of the poverty statistics across quarters will differ.

Table 5
Example of challenge for allocating samples over quarters

Division	District	Stratum	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Total
40	1	Rural	7	7	6	6	26
40	1	Urban	3	3	2	2	10
Sample for district by quarter			10	10	8	8	36
Share of rural sample			70%	70%	75%	75%	72%

To address these issues, we first tackle the second problem – assigning the same number of PSUs to all quarters at the district level. Given that 36 is a multiple of four, we make sure we assign 9 PSUs to each of four quarters. This means for quarter 1 and 2, we assign 7 PSUs for the rural stratum of this district; so we need to assign 2 PSUs to the urban stratum of this district (Table 6). A problem of this adjustment is that the share of the rural sample within this district differs across quarters, even more than in the previous allocation (Table 5). In Table 5, the share of the rural sample ranges between 70 percent and 75 percent, but in Table 6, the share ranges between 67 percent and 78 percent.

Table 6
Example of a modified allocation of samples over quarters

Division	District	Stratum	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Total
40	1	Rural	7	7	6	6	26
40	1	Urban	2	2	3	3	10
Sample for district by quarter			9	9	9	9	36
Share of rural sample			78%	78%	67%	67%	72%

To maintain balance in the distribution of PSUs across strata at the national level over quarters, we randomly select quarters with more PSUs than other quarters. More specifically, the allocation is done as follows. The selection of quarters with more PSUs than others is made only at the first stratum of each district. In the example in Table 6, the first stratum is always the rural stratum of each district. PSUs in other substrata are allocated so that the sum of PSUs for each quarter is 9. For each of the first stratum within a district, we randomly select a quarter from which we start a sequential allocation of PSUs over quarters. This is how the sequential assignment works. For example, for the rural stratum of district 4 in division 10 in Table 7, we randomly select quarter 1 as the starting point for the sequential allocation. Then, we assign one PSU for each quarter in order. Since we start from quarter 1 and the total number of PSUs allocated to this stratum of the district is 25, the sequential allocation of PSUs ends with quarter 1 and as a result, only quarter 1 has 7 PSUs and all others have 6 PSUs. For the rural stratum of district 6 in division 10, we select quarter 3 as a start of this sequential assignment of PSUs. Given that the total number of PSUs allocated to this area is 26, the sequential allocation ends with quarter 4 and thus quarters 3 and 4 receive 7 PSUs while quarters 1 and 2 receive 6 PSUs. Since the selection of starting quarters is random, if we repeat this allocation, we can achieve a similar distribution of PSUs across urban, rural and CC areas at the national and division levels. The example illustrated in Table 7 shows that with just three districts, the share of rural sample is already converging to 54 percent.

Table 7
Example of an allocation of samples over quarters and across districts

Division	District	Stratum	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Total
10	4	Rural	7	6	6	6	25
10	4	Urban	2	3	3	3	11
10	6	Rural	6	6	7	7	26
10	6	Urban	3	3	2	2	10
10	9	Rural	1	2	2	2	7
10	9	Urban	8	7	7	7	29
Sample for district by quarter			27	27	27	27	108
Share of rural sample			52%	52%	56%	56%	54%

Note: Numbers highlighted represent the first quarter of the sequential allocation of PSUs across quarters.

While incorporating these five manual adjustments to the original Neyman sample allocation implies deviating slightly from the optimal allocation, our assessment of the sample performance presented later in the paper shows that the final sample achieves the specified levels of accuracy for the targeted variables. Table 8 shows the final allocation of PSUs across sub-strata after these four manual adjustments.

Table 8
HIES 2016/17 allocation of PSUs by district and quarters

Division	District	Stratum	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Total
40	1	9	7	7	6	6	26
40	1	10	2	2	3	3	10
20	3	3	8	8	8	8	32
20	3	4	1	1	1	1	4
10	4	1	7	6	6	6	25
10	4	2	2	3	3	3	11
10	6	1	6	6	7	7	26
10	6	2	3	3	2	2	10
10	9	2	1	2	2	2	7
10	9	1	8	7	7	7	29
50	10	14	7	7	8	8	30
50	10	15	2	1	1	2	6
20	12	3	8	7	7	7	29

Division	District	Stratum	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Total
20	12	4	1	2	2	2	7
20	13	4	2	2	1	2	7
20	13	3	7	7	8	7	29
20	15	3	1	1	1	1	4
20	15	5	7	7	7	7	28
20	15	4	1	1	1	1	4
40	18	9	7	7	6	7	27
40	18	10	2	2	3	2	9
20	19	4	2	1	1	1	5
20	19	3	7	8	8	8	31
20	22	4	1	2	2	1	6
20	22	3	8	7	7	8	30
30	26	8	7	7	7	7	28
30	26	6	1	1	1	1	4
30	26	7	1	1	1	1	4
55	27	17	8	8	8	8	32
55	27	18	1	1	1	1	4
30	29	7	2	2	2	2	8
30	29	6	7	7	7	7	28
20	30	3	7	7	6	6	26
20	30	4	2	3	3	2	10
55	32	18	2	2	1	2	7
55	32	17	7	7	8	7	29
30	33	6	1	1	1	1	4
30	33	7	8	8	8	8	32
30	35	6	6	6	6	6	24
30	35	7	3	3	3	3	12
60	36	19	7	6	6	6	25
60	36	20	2	3	3	3	11
50	38	14	7	7	8	7	29
50	38	15	2	2	1	2	7
35	39	12	7	7	7	7	28
35	39	13	2	2	2	2	8
40	41	9	6	6	6	6	24
40	41	10	3	3	3	3	12
10	42	2	4	4	4	5	17
10	42	1	5	5	5	4	19
40	44	9	4	3	4	4	15
40	44	10	5	6	5	5	21

Division	District	Stratum	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Total
20	46	3	8	8	8	8	32
20	46	4	1	1	1	1	4
40	47	9	2	2	3	3	10
40	47	10	1	1	1	1	4
40	47	11	6	6	5	5	22
30	48	7	1	1	1	1	4
30	48	6	8	8	8	8	32
55	49	18	2	2	1	2	7
55	49	17	7	7	8	7	29
40	50	10	2	2	2	2	8
40	50	9	7	7	7	7	28
20	51	3	8	8	8	8	32
20	51	4	1	1	1	1	4
55	52	18	1	1	1	1	4
55	52	17	8	8	8	8	32
30	54	7	1	1	1	1	4
30	54	6	8	8	8	8	32
40	55	9	5	5	5	5	20
40	55	10	4	4	4	4	16
30	56	6	7	7	7	7	28
30	56	7	2	2	2	2	8
40	57	9	8	8	8	8	32
40	57	10	1	1	1	1	4
60	58	20	2	2	1	2	7
60	58	19	7	7	8	7	29
30	59	6	6	6	5	5	22
30	59	7	3	3	4	4	14
35	61	13	4	4	4	4	16
35	61	12	5	5	5	5	20
50	64	14	8	8	8	8	32
50	64	15	1	1	1	1	4
40	65	10	2	3	2	2	9
40	65	9	7	6	7	7	27
30	67	6	1	1	1	1	4
30	67	7	8	8	8	8	32
30	68	7	4	4	4	4	16
30	68	6	5	5	5	5	20
50	69	15	1	1	2	2	6
50	69	14	8	8	7	7	30

Division	District	Stratum	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Total
50	70	14	8	8	8	8	32
50	70	15	1	1	1	1	4
35	72	12	7	7	7	7	28
35	72	13	2	2	2	2	8
55	73	18	3	2	2	3	10
55	73	17	6	6	7	7	26
20	75	4	1	2	1	1	5
20	75	3	8	8	8	7	31
50	76	15	3	3	2	3	11
50	76	14	6	6	7	6	25
55	77	17	7	6	6	6	25
55	77	18	2	3	3	3	11
10	78	1	8	8	8	8	32
10	78	2	1	1	1	1	4
10	79	2	1	1	1	2	5
10	79	1	8	8	8	7	31
50	81	15	1	1	1	1	4
50	81	14	1	1	1	1	4
50	81	16	7	7	7	7	28
30	82	6	4	5	5	5	19
30	82	7	5	4	4	4	17
20	84	3	8	8	8	8	32
20	84	4	1	1	1	1	4
55	85	18	3	2	3	3	11
55	85	17	6	7	6	6	25
35	86	13	3	3	3	3	12
35	86	12	6	6	6	6	24
40	87	10	3	2	2	2	9
40	87	9	6	7	7	7	27
50	88	15	2	2	2	1	7
50	88	14	7	7	7	8	29
30	89	7	3	2	3	3	11
30	89	6	6	7	6	6	25
60	90	19	6	6	6	6	24
60	90	20	3	3	3	3	12
60	91	19	7	8	7	7	29
60	91	20	2	1	2	2	7
30	93	7	4	4	4	4	16
30	93	6	5	5	5	5	20

Division	District	Stratum	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Total
55	94	18	4	3	4	4	15
55	94	17	5	6	5	5	21

Table 9 shows the final allocation of PSUs across the 20 sampling strata. The four main CCs each have less than 36 PSUs allocated as these areas have only one sub-stratum. This means that the final sample might not provide reliable poverty and welfare statistics at the division level for urban, rural, and CCs, but the sample will provide reliable poverty and welfare estimates at the district level. This illustrates the trade-offs that household survey sampling experts must assess when designing a sample for a survey with multiple objectives. For the HIES 2016/17, BBS decided to prioritize the production of reliable annual district-level poverty and welfare statistics over annual division-level poverty and welfare statistics for rural, urban, and CC. This decision stems from the fact that district-level statistics correspond to clear administrative policy making units, while division-level statistics do not.

Table 10 shows the final allocation of PSUs for urban and rural areas.²¹ The share of urban PSUs in the HIES 2016/17 sample is slightly higher than 30 percent, while only 23 percent of Bangladesh’s population lives in cities, according to the 2011 Census of Population and Housing. This is explained by the manual adjustments conducted during the allocation of PSUs. As described above, we purposely designed the survey to oversample urban PSUs to help BBS study the relationship between living standards and rapid urbanization.

²¹ BBS has traditionally published urban and rural, but not City Corporation, welfare statistics. Therefore, we only report in this paper the allocation of PSUs across urban and rural areas.

Table 9
HIES 2016/17 allocation of PSUs by stratum

Stratum no	Stratum name	No. of PSUs
1	Barisal Rural	162
2	Barisal Urban	54
3	Chittagong Rural	308
4	Chittagong Urban	60
5	Chittagong City Corporation	28
6	Dhaka Rural	262
7	Dhaka Urban	178
8	Dhaka City Corporation	28
9	Khulna Rural	236
10	Khulna Urban	98
11	Khulna City Corporation	26
12	Mymensingh Rural	100
13	Mymensingh Urban	44
14	Rajshahi Rural	211
15	Rajshahi Urban	49
16	Rajshahi City Corporation	28
17	Rangpur Rural	219
18	Rangpur Urban	69
19	Sylhet Rural	107
20	Sylhet Urban	37
	Total:	2,304

Table 10
HIES 2016/17 allocation of PSUs for urban and rural areas

Sector	HIES 2016/17 sample		Population share, Census of Population and Housing 2011
	No. of PSUs	Percentage	
Urban	699	30.3%	23.1%
Rural	1,605	69.7%	76.9%
Total	2,304	100.0%	100.0%

4. Evaluating the performance of the sample

In this section, we evaluate the performance of the HIES 2016/17 sample in terms of accuracy. More specifically, we evaluate the sample's ability to estimate accuracy statistics for the targeted variables at the level of precision we selected. As mentioned previously, the key parameters used for the estimation of the sample size and the allocation of PSUs across strata are all based on the HIES 2010 data. Table 11 summarizes district level RSEs of mean nominal per capita and total household expenditure. The table also reports SEs of mean poverty headcount ratios (i.e. the three accuracy statistics selected for the HIES 2016/17) estimated using Equation (1) but using $r(\bar{Y})$ as the left-hand side variable and n as the final sample allocation.

All district level nominal per capita consumption expenditure and total household expenditure show RSEs of less than 9 percent, thus below the targeted threshold of 10 percent. SEs of estimated district level poverty rates are also all lower than the acceptable threshold of 5 percent.

Table 11
Annual district level summary accuracy statistics

Accuracy statistic	Observations	Mean	Std. Dev.	Min	Max
RSE of mean nominal per capita expenditure < 10 percent, district level	64	3.34%	1.28%	1.70%	7.89%
RSE of mean nominal total household expenditure < 10 percent, district level	64	3.54%	1.44%	1.85%	8.58%
SE of mean poverty headcount ratios < 5 percent, district level	64	2.67%	0.89%	0.72%	4.78%

Figure 1 shows the spatial distribution of district level RSE estimates of mean nominal per capita and total household expenditure, and the SE estimates of the poverty headcount ratios.²² These figures suggest that the HIES 2016/17 sample generates accurate estimates of poverty.

²² The magnitude of all district-level statistics shown in Figure 1 is reported in Annex 3.

That is to say, poverty estimates derived from the sample are within our selected accuracy statistical range for all three targeted variables estimated at the district level.

Figure 1
Annual district level accuracy statistics

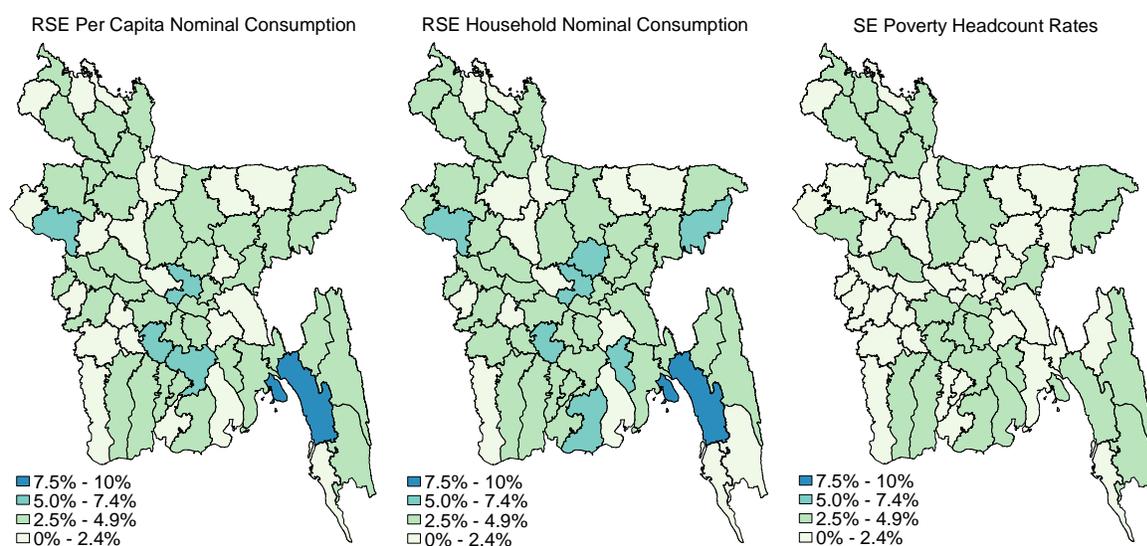


Table 12 shows division-level RSE estimates of mean nominal per capita and total household expenditure, and SE estimates of poverty headcount ratios for urban and rural areas. We have also aggregated the lower district-level statistics.²³ Our analysis suggests that the HIES 2016/17 sampling design will meet all accuracy statistics, thus providing reliable and accurate measures of urban and rural poverty and welfare.²⁴

Lastly, Table 13 shows a summary of quarterly national RSE estimates of mean nominal per capita and total household expenditure. The table also shows SE estimates of poverty headcount ratios. All accuracy statistics are far lower than targeted thresholds. Moreover, the

²³ The underlying assumption for aggregation of district-level statistics to stratum level is that the sample is drawn independently from each of the sub-strata.

²⁴ In Section 5, we also discuss an alternative, direct way to construct division-level accuracy statistics for urban and rural areas; that is, without aggregating from the district-level statistics.

similarity of magnitude of all accuracy statistics across quarters indicates that the sample has been divided adequately across quarters – the division yields poverty and welfare statistics of similar levels of accuracy across quarters.

Table 12
Annual division level accuracy statistics for urban and rural areas

Division	Rural			Urban		
	RSE of mean nominal per capita expenditure	RSE of mean nominal total household expenditure	SE of mean poverty headcount ratio	RSE of mean nominal per capita expenditure	RSE of mean nominal total household expenditure	SE of mean poverty headcount ratio
Barisal	1.8%	1.9%	1.4%	3.5%	3.2%	2.7%
Chittagong	2.2%	2.0%	1.4%	7.4%	8.3%	1.0%
Dhaka	1.3%	1.5%	1.1%	4.0%	4.5%	1.2%
Mymensingh	2.2%	1.8%	2.6%	5.4%	5.3%	3.8%
Khulna	1.0%	1.2%	0.8%	2.3%	1.9%	1.9%
Rajshahi	1.8%	1.8%	1.1%	2.6%	3.0%	2.0%
Rangpur	1.4%	1.4%	1.4%	3.4%	3.3%	2.0%
Sylhet	2.3%	2.4%	1.8%	4.1%	3.8%	1.6%

Table 13
Quarterly national level accuracy statistics

	Q1	Q2	Q3	Q4
RSE of mean nominal per capita expenditure	2.08%	2.08%	2.08%	2.08%
RSE of mean nominal total household expenditure	2.28%	2.28%	2.28%	2.28%
SE of mean poverty headcount ratio	0.83%	0.83%	0.83%	0.83%

5. Robustness analysis

All analysis above is based on the HIES 2010 data. As discussed earlier, the HIES 2010 sample was designed to provide reliable division-level poverty and welfare indicators for urban and rural areas, but not district-level indicators. Therefore, baseline parameters used to design the

HIES 2016/17 sample might be prone to larger margins of error than estimated due to the small district-level sample sizes.

In this section, we use a simple test of robustness to assess how the relatively small district level sample sizes in the HIES 2010 might have affected accuracy statistics for the HIES 2016/17. First, we estimate different sets of division-level statistics for urban and rural areas using both a “bottom-up” and then a “top-down” approach. The bottom-up approach, the methodology we have used to produce the accuracy statistics reported in Table 12, constructs stratum-level statistics by aggregating “up” from district-level accuracy statistics. This approach produces aggregated statistics at the division level for urban and rural areas assuming that the sample is selected from each district independently, as is the case for the HIES 2016/17. The top-down approach instead estimates directly all division-level performance statistics for urban and rural areas, which is the lowest level at which the HIES 2010 sample produced accurate poverty statistics. Comparing the top-down approach to division-level aggregated estimates for urban and rural areas, gives a good test of the accuracy of our bottom-up approach as the latter are estimated based on a sample that has been designed to produce reliable estimates at this level.²⁵

Tables 14, 15, and 16 summarize the main results from this exercise. RSE estimates of mean nominal per capita and total household expenditure, and SE estimates of poverty headcount ratios derived using the bottom-up approach are lower than estimates produced using the top-down approach. Therefore, the bottom-up approach produces more precise estimates in our application than the top-down. Notably, however, even the top-down approach in most cases produces accuracy estimates within the desired levels.²⁶ These results imply that the performance of the HIES 2016/17 sample achieves the desired level of accuracy for poverty

²⁵ For allocation of PSUs across sub-strata, however, we cannot use the bottom-up approach because we cannot fine tune the estimates of the key sampling parameters to allocate PSUs at the sub-strata level.

²⁶ Except for the RSE estimates of nominal per capita expenditure for urban Chittagong and urban Sylhet, 12.5 percent and 11.1 percent, respectively; the RSEs estimates of nominal total household expenditure for urban Chittagong and urban Mymensingh, 13.8 percent and 11.4 percent respectively; and for the SE estimates of the poverty headcount ratios for urban Barisal, urban Mymensingh, urban Khulna, and urban Rangpur, 7.5 percent, 7.2 percent, 5.3 percent, and 5.3 percent respectively.

and welfare monitoring. While accuracy is better for rural areas, the sample design is also accurate for most urban areas.

Table 14
Alternative approaches for estimating RSE of annual per capita consumption
at the Division level (Rural/Urban)

Division	Rural		Urban	
	top-down	bottom-up	top-down	bottom-up
Barisal	4.0%	1.8%	8.5%	3.5%
Chittagong	3.6%	2.2%	12.6%	7.4%
Dhaka	2.6%	1.3%	7.2%	4.0%
Mymensingh	3.2%	2.2%	10.2%	5.4%
Khulna	3.6%	1.0%	7.4%	2.3%
Rajshahi	3.1%	1.8%	6.3%	2.6%
Rangpur	3.3%	1.4%	9.6%	3.4%
Sylhet	4.6%	2.3%	11.1%	4.1%

Note: Urban estimates for the four main divisions of Chittagong, Dhaka, Khulna, and Rajshahi are population weighted averages for Other Urban and City Corporations estimates.

Table 15
Alternative approaches for estimating RSE for annual total expenditure
at the Division level (Rural/Urban)

Division	Rural		Urban	
	top-down	bottom-up	top-down	bottom-up
Barisal	4.3%	1.9%	8.0%	3.2%
Chittagong	3.5%	2.0%	13.8%	8.3%
Dhaka	3.0%	1.5%	8.2%	4.5%
Mymensingh	3.3%	1.8%	11.4%	5.3%
Khulna	3.5%	1.2%	6.1%	1.9%
Rajshahi	3.1%	1.8%	7.0%	3.0%
Rangpur	3.6%	1.4%	9.3%	3.3%
Sylhet	5.4%	2.4%	9.8%	3.8%

Note: Urban estimates for the four main divisions of Chittagong, Dhaka, Khulna, and Rajshahi are population weighted averages for Other Urban and City Corporations estimates.

Table 16
Alternative approaches for estimating SE for annual poverty rates
at the Division level (Rural/Urban)

Division	Rural		Urban	
	top-down	bottom-up	top-down	bottom-up
Barisal	3.7%	1.4%	7.5%	2.7%
Chittagong	2.5%	1.4%	2.6%	1.0%
Dhaka	2.3%	1.1%	2.6%	1.2%
Mymensingh	3.7%	2.6%	7.2%	3.8%
Khulna	2.7%	0.8%	5.3%	1.9%
Rajshahi	2.5%	1.1%	4.6%	2.0%
Rangpur	3.6%	1.4%	5.3%	2.0%
Sylhet	3.5%	1.8%	5.0%	1.6%

Note: Urban estimates for the four main divisions of Chittagong, Dhaka, Khulna, and Rajshahi are population weighted averages for Other Urban and City Corporations estimates.

6. Calculation of sample probability weights

Technically, sampling weights are defined as “the inverse of the probability of being selected into the sample”. In mathematical notation, sampling weights are given by the following expression:

$$w_{hij} = \frac{1}{p_{hij}}$$

where w_{hij} is the household level sample weight for household j , in PSU i , and stratum h ; and p_{hij} is the probability of being included in the sample.

Using sampling weights allows the survey sample to represent the appropriate number of households in the population of interest; using sampling weights inflates the contribution of households who are likely to be under-represented in the sample, and deflates the contribution of those likely to be over-represented, in order to more accurately reflect the reference population the survey is attempting to study.

In multi-stage, stratified clustered designs, the probability of being selected must be reflected in each of the different sampling stages. In the case of a two-stage, stratified clustered design, such as HIES 2016/17, the probability of being selected into the sample is a function of: 1) the probability of a PSU being selected in the first stage plus; and 2) the probability of a household being selected within each PSU in the second stage. This can be calculated as follows:²⁷

$$p_{hij} = p_1 * p_2 = \frac{k_h * m_{hi}}{N_h} \quad (3)$$

where p_{hij} is the probability of household j , in stratum h , and PSU i to be included in the sample; p_1 is the probability of the PSU to be selected in the first stage; p_2 is the probability of a household being selected in the second stage; k_h is the number of PSUs selected in stratum h ; m_{hi} is the number of households selected in PSU hi ; and N_h is the total number of households in stratum h .

The probability sampling weights in Equation (3) are called *ex-ante* weights—that is, done before the survey field work—and closely follow the sampling design. It is common, however, to adjust sampling weights *ex-post*—that is, after survey data collection—to correct for imperfections and bias in the sample identified during field work. If not corrected, these imperfections can lead to errors that detract from understanding the reference population. These adjustments usually include the following: (i) corrections for household non-response conducted at the PSU level; (ii) corrections for sample frame errors generally linked to outdated information, which are usually take place at the PSU level; and (iii) post-stratification to match official projected population statistics at a chosen level of aggregation.²⁸ The final expression below accounts for the *ex-post* adjusted sampling weights:

²⁷ Since the probability of being selected in each stage is a function of the cluster size and the number of units sampled in each stage, the overall probability of being selected in most living standards household surveys is constant at the level in which PSUs are allocated (i.e. sub-stratum in the case of the HIES 2016/17). This is because PSUs are generally allocated across stratum using proportional allocation based on population size or number of households, which results in a constant probability in the first stage as long as no adjustments are made, such as post-stratification or update of the population based on the listing exercise.

²⁸ Depending on the sampling frame used for the sampling design, and whether a listing exercise or post-stratification actually takes place or not, the sum of the household-level probability weights in the survey will

$$p_{hij} = p_1 * p_2 = \frac{k_h n_{hi}}{N_h} * \frac{m_{hi}}{n'_{hi}} \quad (4)$$

where p_{hij} is the probability of household j , in stratum h , and PSU i to be included in the sample; k_h is the number of PSUs selected in stratum h ; m_{hi} is the number of households selected in PSU hi ; N_h is the total number of households in stratum h ; n_{hi} is the number of households to be selected in PSU i and stratum h based on the sampling frame; and n'_{hi} is the number of households in PSU i and stratum h interviewed, listed, or consistent with the official population statistics.

For non-response corrections, the n_{hi}/n'_{hi} term corresponds to the number of households to be selected divided by the actual number of households interviewed in each PSU. For corrections to the PSU size based on the listing exercise, the n_{hi}/n'_{hi} term corresponds to the number of households in the PSU based on the sample frame divided by the number of households in the PSU based on the listing exercise.²⁹ Lastly, for post-stratification adjustments to make sampling weights add up to official population statistics, the n_{hi}/n'_{hi} term corresponds to the number of households based on the official population statistics divided by the number of households based on the sample frame.

The sampling probability weights for the HIES 2016/17 have been adjusted *ex-post* for two things: (i) to match the official share of urban population reported in all official publications produced by BBS after the 2011 Census of Population and Housing, and (ii) to match the official 2016 population projections estimates produced at the district (*zila*) level (BBS 2015). More specifically, the 2011 Census of Population and Housing calculated the share of urban population in Bangladesh at 23 percent. A month after completion of the Census, the Bangladesh Institute of Development Studies (BIDS), the independent agency in charge of

equal the reference population of interest at the time of the survey data collection; and/or weights will vary across PSUs or be constant at the level at which PSUs are allocated.

²⁹ The listing exercise consists of the enumeration of households in a PSU before a survey is implemented. This list is used to select households randomly within a PSU, and it is only conducted in PSUs that were selected in the first stage of a multi-state clustered household survey.

assessing the quality of the data, conducted a Post-Enumeration Check Survey (PECS) that found that there was under coverage both in urban and rural areas, but this was more prevalent in urban areas. BBS thus used a two-step approach to adjust the 2011 Census of Population and Housing estimates. First, it reclassified urban and rural areas using the concepts of: (i) growth centers, (ii) urban agglomerations, and (iii) other urban areas. Second, it inflated all urban and rural counts from the 2011 Census of Population Areas to align with the PECS results. These two adjustments estimated the share of the urban population at 28 percent, which is the number that BBS has been using since then in order to produce official population projections and statistics.³⁰

To address this mismatch between the original HIES 2016/17 sampling and the official population statistics, we adjust our *ex-ante* sampling weights in two ways. First, we re-classify all selected urban/rural PSUs using the new “growth centers” categorizations, which affects 24 out of the 64 *zilas* in the sample. This adjustment brings the urban share from 23 percent to 26.6 percent. Second, we use the official 2016 population projections estimated at the *zila* level to inflate the adjusted sampling weights; that is, for each *zila* we divide the official 2016 population projections by the adjusted 2011 population projections used for the original sampling design.³¹ For this exercise, we use the number of households in the original 2011 census frame (i.e. enumerated households) instead of the adjusted number of households (i.e. the number adjusted after the PECS). After all these adjustments, the final sampling weights for the HIES 2016/17 yield an urban share of 28 percent, and the sum of the sampling weights by *zila* match the official BBS 2016 population statistics.

³⁰ The BBS official population projections are available at the following link:

<http://www.bbs.gov.bd/WebTestApplication/userfiles/Image/PopMonographs/PopulationProjection.pdf>

³¹ To convert the adjusted number of households to adjusted population counts for estimating the final post-stratification factor by *zila*, we use the official household size reported in the 2011 Census for Population and Housing for rural, urban, and City Corporations areas – 4.49 (rural), 4.45 (urban), 3.98 (City Corporation).

7. Estimation of standard errors

The field protocols used by BBS to collect the HIES 2016/17 data are based on the concept of terms. Each term consists of 20 days, the time needed for a team of two enumerators to cover an entire PSU (20 households). The 12-months of data collection are divided across 18 terms (18*20 days=360 days). As the 18 terms of data collection planned cannot be divided exactly into quarters, PSUs in terms 5 and 14 will need to be randomly split to avoid sampling bias so that each quarter corresponds to exactly 4.5 terms. This random split is conducted by randomly allocating PSUs in term 5 to quarters 1 and 2, and those in term 14 to quarters 3 and 4.

The HIES 2016/17 sample was designed so that each of the 132 sub-strata includes at least 4 PSUs during the entire four quarters of planned data collection. The allocation of PSUs was conducted in such a way that each quarter includes at least one PSU per sub-strata to ensure that each of the quarterly poverty estimates includes observations from each of the 132 sub-strata. This implicitly means, however, that standard errors for quarterly poverty rates and welfare statistics cannot be estimated for sub-strata that have only one PSU assigned. This is not an issue when estimating annual poverty rates or annual welfare statistics, as each of the sub-strata includes at least four PSUs, but this poses a challenge when trying to estimate standard errors for quarterly statistics.

To address this issue, we consider four alternative ways of estimating standard errors for quarterly poverty rates and welfare statistics based on the HIES 2016/17. In the first option, we propose to use the method called the “centered approach”. This method calculates the standard error for the sub-strata based on the distance from the grand mean rather than the sub-strata means. In the second method, we propose to use the “certainty method”. This method assumes that a sub-strata with a single PSU do not contribute to the overall standard error and is equivalent to specifying a finite population correction of 1. The third method we propose is called the “scaled one”. This method assumes that the contribution to the standard error for a sub-stratum with a single PSU is equal to the mean estimate for all of the other (multiple PSU) strata and it yields essentially the same results as those that would be obtained using mean imputation. Lastly, in order to maintain full comparability across years in the estimation of

poverty and welfare statistics, we also propose to deviate slightly from the sampling design and estimate standard errors using the 16 strata originally defined for the HIES 2010. Using this approach to estimate standard errors of poverty rates and welfare statistics will allow for full consistency in the way these statistics are produced across time.

8. Conclusion

This paper describes the different steps involved in the design of the sample for the Bangladesh HIES 2016/17, whose main objective is to produce reliable poverty and welfare statistics at different levels. Designing a sample to achieve multiple objectives is not a simple exercise as objectives might not necessarily be mutually consistent. How to prioritize the objectives of a survey and address any potential inconsistencies that might arise between multiple objectives is important to ensure that the final selected sample does not compromise statistical precision for all levels of interest to policy makers.

The HIES 2016/17 sample was designed to provide reliable poverty and welfare statistics at three different levels: (i) annual poverty estimates at the district level; (ii) annual poverty estimates at the division level for urban and rural areas; and (iii) quarterly poverty estimates at the national level. The sampling follows a two-stage clustered design. In the first stage, we allocate an equal number of PSUs to each of the 64 districts in Bangladesh to ensure that the sample estimates district-level poverty and welfare statistics with similar levels of reliability across districts. In the second stage, we allocate the sample across three sampling strata (urban, rural, and City Corporation) for each district independently and select households from each of the selected PSUs. To allocate the sample across strata in each district, we modify the standard Neyman allocation rule that allows the achievement of an optimal allocation for clustered stratified samples (the standard Neyman allocation rule only allows an optimal allocation for stratified samples). In the process of selecting the final sample, we also discussed how field data collection protocols are taken into consideration and illustrate their importance in fine tuning the final allocation of the sample.

To determine the sample size, we use several target variables: (i) nominal per capita expenditure; (ii) nominal total household expenditure; and (iii) poverty headcount ratios. We conducted an assessment of how the sample changes using these different targeted variables but our final estimates are based on total nominal household expenditure in order to be consistent with the sampling design of the HIES 2010.

We use three accuracy statistics to assess the performance of the sample: (i) RSEs of less than 10 percent from the mean nominal per capita expenditure for each district and division (urban/rural), as well as for each quarter at the national level; (ii) RSEs of less than 10 percent from the mean nominal household expenditure for each district and division (urban/rural), and for each quarter at the national level; (iii) SEs of less than 5 percent of the poverty headcount ratio for each district and division (urban/rural), and for each quarter at the national level.

Given that the sample for the HIES 2016/17 was designed to provide reliable poverty and welfare statistics at three different levels, accuracy in the estimation of statistics for all levels cannot be obtained simultaneously without compromising on the cost of the survey (i.e. on the sample size). The final allocation of the sample was selected to minimize the sampling errors at the district level, produces reliable and unbiased estimates of poverty for each quarter, and achieves the level of precision of poverty statistics for most of the sampling strata.

For the HIES 2016/17, BBS decided to prioritize the production of reliable annual district-level poverty and welfare statistics over annual division-level statistics for rural, urban, and CC. This decision stems from the fact that district-level statistics correspond to clear administrative policy making units and estimation of reliable statistics at the district level is in general more challenging. As a result, we first compute the optimal allocation of the sample at the district level and then modify it slightly for cases in which poverty and welfare estimates at the other two levels are not reliable. After these modifications, the final proposed sample successfully achieves an acceptable level of accuracy for poverty and welfare statistics at the three levels – annual district level, quarterly national level, and annual division rural level. Even after these modifications, however, there is some suggestive evidence that poverty and welfare estimates for some of the urban strata in the divisions of Chittagong and Mymensingh

might not be measured with adequately high levels of accuracy. We will be able to assess this once the final survey data are available for analysis.

Lastly, we also discuss the different options available to estimate standard errors of poverty and welfare statistics to be produced using the proposed sampling design, as well as some challenges that arise in estimating standards errors for quarterly statistics.

References

Bangladesh Bureau of Statistics (2015). Population Projection of Bangladesh: Dynamics and Trends 2011-2061. Statistics and Informatics Division (SID). Ministry of Planning. Dhaka, November 2015.

Cochran, W. (1977). Sampling Technique. Third Edition. John Wiley & Sons. New York. USA.

Cornfield J. (1951). The determination of sample size. *Am J Public Health Nations Health*, Vol. 41(6), 654–661.

Kish, L. (1965). Survey Sampling. Wiley Classics Library Edition Published 1995. John Wiley & Sons. New York. USA.

Maligalig, D. S and Martinez, A. Jr (2013). Developing a Master Sample Design for Households Surveys in Developing Countries: A Case Study in Bangladesh. *Survey Methods: Insights from the Field*.

Munoz, Juan (1998). Bangladesh Household Income and Expenditure Survey 2010 Sampling Design. Bangladesh Bureau of Statistics. Mimeo.

Munoz, Juan (2004). Bangladesh Household Income and Expenditure Survey 2005 Sampling Design. Bangladesh Bureau of Statistics. Mimeo.

Neyman, J. (1934). “On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection”, *Journal of the Royal Statistical Society*, Vol. 97, 558-606.

World Bank (2013). Bangladesh Poverty Assessment: Assessing a Decade of Progress in Reducing Poverty 2000-2010. Bangladesh Development Series Paper No. 31. The World Bank Office, Dhaka.

Annex 1: Derivation of the formula to estimate sample size under a stratified clustered sample design

In this annex, we derive the formula to determine the minimum sample size needed to achieve a certain level of precision (estimating a selected accuracy statistic) of a targeted variable for stratified clustered sampling. This formula modifies the derivation presented in Cochran (1977).³²

In a stratified clustered sampling design, we can write the confidence interval of a ratio of the sample mean of a targeted variable we want to achieve as:³³

$$\bar{y} * r(\bar{y}) = Z_{\alpha/2} * SE_{clustered}(\bar{y}) \quad (A1.1)$$

where \bar{y} is the sample mean of the targeted variable, $r(\bar{y})$ is a ratio of the sample mean that we desire to achieve (RSE or SE in most applications), $Z_{\alpha/2}$ is the critical value of a standard normal variable at a α % confidence level, and $SE_{clustered}(\bar{y})$ is the clustered standard error of the ratio of the sample mean. Since $SE_{clustered}(\bar{y}) = \left(\frac{SD_{SRS}(\bar{y})}{\sqrt{n}}\right) * DEFT$, Equation (A1.1) can be written as follows:

$$\bar{y} * r(\bar{y}) = Z_{\alpha/2} * \left(\frac{SD_{SRS}(\bar{y})}{\sqrt{n}}\right) * DEFT \quad (A1.2)$$

where n is the minimum sample size needed to achieve a certain level of reliability in the estimation of a selected accuracy statistic for the targeted variable, SD_{SRS} is the standard deviation of the mean of the targeted variable under simple random sampling (SRS), and $DEFT$ is the square root of the average design effects across clusters ($DEFF$).

³² The derivation of the formula to estimate the sample size for a stratified sampling design is provided in Cochran (1977), pages 77-78.

³³ The use of this formula implicitly assumes that the targeted variable of interest is normally distributed. In addition, the critical value used to estimate the confidence interval is normally set up at 95 percent and therefore $Z_{\alpha/2}$ is replaced by a value of 1.96.

Re-arranging (A1.2), we obtain the following expression:

$$n = \left(\frac{Z_{\alpha/2} * SD_{SRS}(\bar{y}) * DEFF}{\bar{y} * r(\bar{y})} \right)^2 \quad (A1.3)$$

Since $SD_{SRS}(\bar{y})/\bar{y} = CV_{SRS}(\bar{y})$ (coefficient of variation), Equation (A1.3) can be written as follows:³⁴

$$n = \left(\frac{Z_{\alpha/2} * CV_{SRS}(\bar{y})}{r(\bar{y})} \right)^2 * DEFF \quad (A1.4)$$

For continuous targeted variables, Equation (A1.4) is normally written as follows:

$$n_{continuous} = \left(\frac{Z_{\alpha/2} * CV_{SRS}(\bar{y})}{RSE(\bar{Y})} \right)^2 * DEFF \quad (A1.5)$$

where $RSE(\bar{Y})$ is the desired relative standard error of the sample mean of the continuous targeted variable.

For dichotomic targeted variables, Equation (A1.4) can be written as follows:

$$n_{proportion} = \left(\frac{Z_{\alpha/2} * CV_{SRS}(\bar{y})}{Var(\bar{Y})} \right)^2 * DEFF \quad (A1.6)$$

where $Var(\bar{Y})$ is the desired variance of the sample proportion of the dichotomic targeted variable.

³⁴ The final formula for determining the minimum sample size to achieve certain accuracy statistic is generally expressed in terms of the coefficient of variation of the target variable. This is generally easier to estimate *ex-ante* than the standard deviation of the targeted variable.

Annex 2: Optimal allocation of PSUs across strata for stratified clustered samples (modified Neyman allocation rule)

For a stratified random sample, the variance of the sample mean of a targeted variable y can be written using the following expression:³⁵

$$V(\bar{y}_{st}) = \sum_{h=1}^L W_h^2 * V(\bar{y}_h) \quad (\text{A2.1})$$

where W_h is a stratum level weight for stratum h constructed based on a measure of size (population size or number of households more often); and $V(\bar{y}_h)$ is the variance of the sample mean of the targeted variable in stratum h .

For a stratified sample, the variance of the sample mean of the targeted variable in each stratum can be written as follows:

$$V(\bar{y}_h) = \frac{S_h^2}{n_h} (1 - f_h) \quad (\text{A2.2})$$

where S_h is the standard deviation of the targeted variable, n_h is the stratum h sample size, and f_h refers to the stratum sampling fraction or finite population correction (i.e. $f_h = n_h/N_h = \text{stratum } h \text{ sample size} / \text{stratum } h \text{ population size}$).

For a stratified clustered sample, the variance of the sample mean of the targeted variable in each stratum can be written as follows:

$$V(\bar{y}_h) = \frac{S_h^2 * DEFF_h}{n_h} (1 - f_h) \quad (\text{A2.3})$$

where $DEFF_h$ is the stratum h design effect and it is estimated using the following expression: $DEFF = 1 + (m-1)\rho$ where ρ is the intra-cluster correlation and m the number of observations in each cluster.

³⁵ In stratified random sampling, the sample is assumed to be drawn independently from each strata.

By substituting Equation (A2.3) into (A2.1), we obtain the following expression:

$$V(\bar{y}_{st}) = \sum_{h=1}^L W_h^2 * \frac{S_h^{2*DEFF_h}}{n_h} (1 - f_h) \quad (\text{A2.4})$$

Therefore, to obtain the optimal allocation of PSUs in each stratum for a stratified clustered sample, we need to solve the following optimization problem:³⁶

$$\min_{n_h} V(\bar{y}_{st}) = \sum_{h=1}^L W_h^2 * \frac{S_h^{2*DEFF_h}}{n_h} (1 - f_h) \quad (\text{A2.5})$$

$$\text{st. } n = \sum_{h=1}^L n_h$$

where n_h refers to the sample allocated to stratum h and n refers to the total sample size. Therefore, we need to minimize the variance of the sample mean of the targeted variable in each stratum subject to a fixed pre-established sample size of n .

Assuming that the cost of collecting an interview in each PSU is constant across clusters, and assuming that all clusters are of equal sample size, the expression to calculate the optimal size to allocate to each stratum h is:³⁷

$$n_h^* = \frac{W_h S_h DEFF_h \sqrt{(1-f_h)}}{\sum_{h=1}^L W_h S_h DEFF_h \sqrt{(1-f_h)}} * n \quad (\text{A2.6})$$

The Neyman allocation rule is a special case of Equation (A2.6) when $DEFF_h = 1$ and $f_h = 0$) and it is optimal only when the sampling design follows a stratified sample (i.e. $DEFF_h = 1$)

³⁶ The derivation of the optimal allocation rule for a stratified sample is available in Cochran (1977), pages 96-99.

³⁷ The assumption that the cost per unit is the same across all clusters can easily be relaxed if good information exists on the costs of collecting data per cluster to inform the sampling design. This would result in a slightly more complicated expression for the optimal allocation of PSUs across strata, but it is straightforward to derive. Similarly, the assumption of equality of clusters can easily be relaxed and an analog expression derived to account for the unequal cluster sizes.

and when there is no finite sample correction (i.e. $f_h=0$). However, for stratified clustered samples, the Neyman allocation rule is not optimal; the modified more general formula provided in equation (A2.6) will provide an optimal allocation of PSUs across strata.

Annex 3: Assessing the performance of the HIES 2016/17 sample to achieve selected district level accuracy statistics

Table A3.1
District level estimates

District	Division	Sample size	RSE of nominal per capita expenditure	RSE of nominal total household expenditure	SE of poverty headcount ratio
1	40	160	4.0%	3.7%	3.5%
3	20	140	2.8%	2.3%	2.9%
4	10	160	2.3%	2.7%	1.8%
6	10	200	5.1%	3.9%	3.8%
9	10	160	1.7%	2.4%	2.8%
10	50	240	2.6%	2.5%	1.9%
12	20	200	3.9%	3.1%	3.3%
13	20	180	1.9%	2.3%	2.2%
15	20	480	7.9%	8.6%	3.3%
18	40	160	2.0%	2.3%	2.0%
19	20	300	2.5%	2.8%	2.4%
22	20	120	1.7%	2.5%	2.4%
26	30	400	6.7%	7.5%	2.3%
27	55	200	3.5%	4.6%	3.7%
29	30	160	2.6%	3.5%	3.2%
30	20	160	3.2%	3.3%	2.2%
32	55	160	3.7%	3.5%	3.7%
33	30	240	5.0%	5.3%	1.5%
35	30	160	5.2%	5.6%	4.0%
36	60	180	2.7%	2.9%	2.4%
38	50	120	4.4%	4.7%	3.2%
39	35	200	2.1%	2.4%	1.9%
41	40	240	2.4%	2.8%	2.0%
42	10	160	4.0%	4.7%	2.5%
44	40	180	3.2%	3.0%	2.3%
46	20	140	3.8%	3.3%	1.8%
47	40	260	3.1%	3.1%	2.7%
48	30	200	3.5%	3.9%	2.2%
49	55	180	3.5%	4.1%	4.3%
50	40	180	3.3%	3.2%	0.7%
51	20	140	5.0%	5.3%	2.9%
52	55	120	1.8%	1.8%	1.9%
54	30	140	3.7%	3.3%	3.7%

District	Division	Sample size	RSE of nominal per capita expenditure	RSE of nominal total household expenditure	SE of poverty headcount ratio
55	40	160	2.4%	2.5%	2.2%
56	30	160	1.8%	2.2%	2.3%
57	40	140	3.7%	4.2%	0.9%
58	60	220	4.8%	5.3%	3.9%
59	30	120	2.8%	2.8%	2.6%
61	35	440	4.0%	3.4%	4.6%
64	50	180	3.0%	2.9%	2.5%
65	40	140	2.3%	3.3%	1.6%
67	30	240	2.9%	3.5%	2.4%
68	30	200	2.4%	3.8%	2.2%
69	50	160	2.0%	3.6%	1.9%
70	50	120	2.0%	2.7%	1.8%
72	35	160	1.8%	2.2%	2.2%
73	55	140	1.8%	1.9%	2.9%
75	20	200	4.5%	4.6%	1.0%
76	50	200	4.0%	3.8%	3.1%
77	55	140	3.0%	2.8%	2.9%
78	10	160	4.8%	5.4%	3.6%
79	10	140	3.9%	3.9%	2.5%
81	50	360	6.7%	7.2%	3.5%
82	30	160	2.8%	2.7%	2.2%
84	20	140	3.3%	3.5%	3.9%
85	55	220	4.7%	4.1%	3.8%
86	35	120	3.5%	3.1%	4.8%
87	40	180	2.1%	2.3%	2.0%
88	50	200	2.2%	2.3%	1.9%
89	30	160	2.5%	2.8%	2.7%
90	60	220	2.5%	2.5%	2.0%
91	60	240	4.1%	4.0%	3.3%
93	30	280	3.8%	3.5%	4.0%
94	55	120	2.5%	3.2%	2.3%