Note

Bangladesh Development Studies
Vol. XLII, June-September 2019, Nos. 2 & 3

Official Methodology Used for Poverty Estimation Based on the Bangladesh Household Income and Expenditure Survey 2016/17: A Technical Note

FAIZUDDIN AHMED[*]
MARIA EUGENIA GENONI[**]
DIPANKAR ROY[***]
ABDUL LATIF [****]

This paper describes the official methodology to estimate poverty based on household per capita consumption in Bangladesh using the Household Income and Expenditure Survey 2016/17. The analysis discusses changes in the survey that may affect the comparability of poverty estimates across time and a series of robustness checks to the estimation. It also includes a short description on how the official income aggregate is estimated.

Keywords: Poverty Line, Measurement, Survey, Methodology
JEL Classification: I30, I32

I. BANGLADESH'S HOUSEHOLD INCOME AND EXPENDITURE SURVEY

The *Household Income and Expenditure Survey* (HIES) is a comprehensive, nationally representative survey used to measure monetary poverty in Bangladesh. The HIES 2016/17 is the fourth round in the series of HIES conducted by the Bangladesh Bureau of Statistics (BBS) in 2000, 2005, and 2010. Before 2000, BBS monitored poverty using a smaller survey, the *Household Expenditure Survey* (HES), which was limited to expenditure data. The World Bank provided technical

[*] The World Bank.
[**] The World Bank.
[***] Bangladesh Bureau of Statistics (BBS).
[****] Bangladesh Bureau of Statistics (BBS).
The authors acknowledge the contributions of Jose Joaquin Endara Cevallos, Yurani Arias-Granada, Nobuo Yoshida, and Monica Yanez in the preparation of earlier versions of this paper. Valuable comments and advice were received from Ruth Hill, Dean Jolliffe, and Benu Bidani.

assistance to the BBS in the development of the HIES 2016/17 questionnaire, sampling design, data collection protocols, and poverty estimates.

1.1 Sampling Design

A stratified, two-stage sample design was adopted for the HIES 2016/17, with 2,304 Primary Sampling Units (PSUs) selected from the list of the *2011 Housing and Population Census* enumeration areas. Within each PSU, 20 households were selected for interviews. The final sample size was 46,080 households (Ahmed, Roy, Yanez-Pagans and Yoshida 2017).

In Bangladesh, divisions are the first-level administrative geographical partitions of the country. As of 2016, the country has eight divisions: Barisal, Chittagong, Dhaka, Khulna, 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), and other urban areas.

PSUs in the HIES 2016/17 were allocated at the district level. Therefore, the sample was stratified at the district level. Since there were a total of 64 districts in Bangladesh, the sample design included a total of 132 sub-strata: 64 urban, 64 rural, and four main CCs. The sample was also implicitly stratified by month.

Table I presents a summary of sample design and PSU allocation.[1]

TABLE I

HIES 2016/17 SUMMARY OF SAMPLING DESIGN

Description	Number
Number of districts	64
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

Departures from the previous HIES. The samples of the latest three rounds of the HIES were designed to provide reliable annual poverty estimates for the country's divisions by urban and rural areas separately and the Statistical

[1] There was a replacement strategy for households that were not found or refused to answer. However, the households that were replaced were not identified during the fieldwork.

Metropolitan Areas (SMAs).[2] However, the HIES 2016/17 was designed to produce reliable poverty estimates at three different levels: (i) annual poverty estimates at the division level for urban and rural areas; (ii) annual poverty estimates for the country's 64 districts; and (iii) quarterly poverty estimates at the national level. This change implied quadrupling the sample size of HIES 2016/17 compared to previous rounds – from 12,240 in 2010 to 46,080 households.

The substantial increase in the sample size also required using a different sampling frame to accommodate the larger number of PSUs. The PSUs for all the previous rounds of the HIES were selected from the Integrated Multiple-Purpose Sample (IMPS) – a master sample updated after each *Housing and Population Census.* In the HIES 2016/17, the PSUs come from the list of Enumeration Areas (EAs) used for Bangladesh's *2011 Population and Housing Census*. The IMPS could not be used because the most recent version, based on the 2011 Census, included only 2,012 EAs, an insufficient number to serve as a sampling frame for this new round of the survey. Importantly, the Bangladesh IMPS excluded some geographic areas, such as urban slums. Therefore, the HIES 2016/17 has a higher likelihood of capturing slum areas.[3]

1.2 Period of Data Collection

The HIES 2016/17 was in the field for an uninterrupted period of 12 months. The survey was launched on April 1, 2016, and field operations were completed on March 31, 2017. Data were collected over a year to capture seasonal variations in expenditure, expenditure patterns, and income. The one-year period was divided into 18 terms of 20 days. A term is the time needed for a team of two enumerators to cover the 20 households selected within a PSU.

1.3 Questionnaires

The 2016/17 HIES consisted of nine major modules, covering various aspects of household activities and characteristics (household roster, education, health, economic activities, non-agricultural enterprises, housing, agriculture, other assets and income, and consumption). The 2016/17 HIES redesigned and expanded the social safety net questions. The final questionnaire reflects several technical discussions on questionnaire design and content.

---

[2] In 2017, the country had seven divisions: Dhaka, Chittagong, Barisal, Khulna, Sylhet, Rangpur, and Rajshahi.

[3] For details of the sampling design, see Ahmed *et al*. (2017).

1.4 Data Entry and Management

The data collection, entry, and transfer process for the HIES 2016/17 was conducted using Paper and Pencil (PAPI) combined with CAFE (Computer-Assisted Field-Based Data Entry). The data were collected by interviewers using PAPI and later entered or digitized using laptops while interviewers were still in the field. The data entry application was developed in CSPro and was paired with a cloud-based data transferring system, which allowed teams to transfer data to the BBS headquarters and monitor data in almost real time using a mobile internet connection. After the data was transferred to BBS headquarters, it was compiled and exported to a readable version by standard statistical software using an automatized routine.

The data entry and transfer system was combined with a data monitoring system for a selected set of variables important for poverty measurement. This data monitoring system fed from the compiled data to create a set of key indicators that were tracked on a continuous basis. The indicators that were tracked by team, term, division, and district included: number of households, household size, number of households with incomplete food and non-food consumption, number of households with incomplete durable items, number of daily food items consumed by households, number of weekly food items consumed by households, and number of non-food items and durables consumed by households. This information supported the supervision of fieldwork and ensured that consumption data was complete and high quality for poverty estimation.[4]

II. METHODOLOGY TO ESTIMATE POVERTY

2.1 Welfare Aggregate

Poverty estimates in Bangladesh were based on household per capita consumption. The consumption section of the HIES questionnaire was divided into five parts:

A. Daily food consumption: Information on daily food consumption for 130 items was collected for 14 consecutive days. Interviewers registered consumption in quantities and corresponding values with sources of receipts.

[4] It is important to note that other variables collected were not monitored, including income-related information. Ex-post analysis of the data indicates that the data entry of income-related variables suffered weaknesses due to lack of range checks and merging issues in the CSPro data entry programme.

B. Weekly food consumption for around 19 items
C. Monthly non-food consumption for about 49 items
D. Annual non-food expenditure for more than 177 items
E. Inventory of durable goods

The consumption aggregate for the HIES 2016/17 was constructed by adding all food and non-food consumption expenditures reported by households, except for taxes and fees, lumpy-cycle expenditures such as expenses for weddings, and interest and insurance expenses. Non-food expenditures included: fuel and lighting, cosmetics and hygiene items, transport and travel, readymade garments, clothing materials, footwear, household-use textiles, health treatment expenses, housing-related expenses, education, recreation, and leisure. The non-food expenditure component also included housing rent, imputed rent (i.e., the amount that homeowners report they would like to get if they could rent their house), or predicted rent, depending on the homeownership status of each of the households.[5] For renters, the reported rent was included as part of the non-food consumption aggregate. For homeowners, the reported imputed rent was included as part of the non-food consumption aggregate. For households that did not report rent or imputed rent, a predicted rent was estimated using a regression model on the subsample of renters and added to the non-food consumption aggregate. This regression model was estimated using the (log of) reported rent on the left-hand side and was regressed against a set of housing characteristics, including number of rooms, wall materials, access to electricity and tap water, kitchen, dining room, telephone connection, dwelling's land size, and a vector of the 16 original strata dummy variables.

The construction of the consumption aggregate followed the 2010 methodology as closely as possible. However, there was one important departure in the methodology related to the computation of education expenditures. Education expenditures were collected in Sections 2 and 9 of the survey. Traditionally, for the computation of the consumption aggregate, education expenditures are added using the information from Section 9. In the 2016/17 round, it was found that 5.6 per cent of households had reported zero or missing education expenditures in Section 9, but had positive expenditures reported in Section 2. In 2010, this was true for only 1.2 per cent of households. Therefore,

[5] The rent and imputed rent variables were cross-tabulated against house ownership, and a few observations were cleaned to ensure full consistency between these two variables.

the 2016/17 consumption aggregate used information from Section 2 to replace the zero and missing values in Section 9. Section 5 shows that this departure in the computation of the aggregate does not significantly change poverty estimates.

Finally, the consumption aggregate was divided by the household size to obtain a per capita measure. The HIES survey defines a household as a group of people who eat from the same pot and sleep in the same dwelling. Household members are defined as people who have eaten and slept in the dwelling for at least six months during the past 12 months (not necessarily continuous), or members who have been in the dwelling for less than six months over the past year, including any of the following: (i) the head of the household; (ii) a major provider of economic support; (iii) infants under six months old; or (iv) a new bride who joined the household less than six months ago. In addition, servants are counted as household members.

The average household size for the HIES 2016/17 was 4.06 members. This implies a significant reduction in the average household size compared to the latest HIES 2010 (average household size was 4.5), which is not explained by differences in the definition of households. Recent national representative surveys collected by BBS show consistent large reductions in household size for the past years in Bangladesh. For example, the *Quarterly Labour Force Survey 2015/16* shows an average household size of 4.2. Annex 2 summarises an analysis that compares the HIES household size estimate with other surveys and projections and concludes that the estimations from HIES 2016/17 are in line with trends in fertility and population change.

Table II presents the average household consumption per capita for 2010 and 2016/17.

TABLE II

HOUSEHOLD CONSUMPTION PER CAPITA, HIES 2010 AND 2016/17 IN MONTHLY TAKAS OF 2016

Consumption per capita	2010		2016/17	
	Average	Standard error	Average	Standard error
Total	3,431	62	3,760	41
Food	1,901	22	1,809	13
Non-food	1,536	47	1,951	34

Note: All expenditures are deflated across space and expressed in 2016 prices.

2.2 Estimation of the Poverty Lines

The official methodology used in Bangladesh to estimate poverty numbers was based on the Cost of Basic Needs (CBN). The CBN method calculates the cost of obtaining a consumption bundle considered to be adequate to satisfy basic consumption needs. If a person cannot afford the cost of this bundle, then this person is considered poor. Therefore, poverty lines under the CBN method represent the minimum per capita expenditure that a person needs to meet his basic needs.

The first step for computing a poverty line involved estimating the cost of a basic consumption food basket. In Bangladesh, the food basket included eleven items (coarse rice, wheat, pulses, milk, oil, meat, fish, potatoes, other vegetables, sugar, and fruits), as recommended by Ravallion and Sen (1996) following Alamgir (1974). This food bundle provided the minimal nutritional requirements corresponding to 2,122 kcal per day per person. The price for each item in the bundle was estimated using unit-values (price per unit) from the HIES. The price for each item was the median of the unit-values reported by a reference group of households calculated separately for each stratum of the survey. The food poverty line was then computed for each stratum by multiplying the estimated prices with the quantities in the food bundle.[6]

Starting in 2000, the HIES defined 16 different geographical strata that have been used since then to estimate the cost of the basic consumption bundle. The estimation of this bundle at different geographical levels allows analysts to account for cost-of-living differences across areas and therefore provides a more accurate picture of living standards after accounting for price differences across geographic areas. These 16 original strata include urban and rural areas in the six divisions that existed in 2005 (Barisal, Chittagong, Dhaka, Khulna, Rajshahi, and Sylhet) and the four main SMAs (Chittagong, Dhaka, Khulna, and Rajshahi). Out of the 16 original strata, six are classified as rural, and ten are classified as urban.[7]

---

[6] The reference groups are the households belonging to the 2^{nd} to 6^{th} deciles of the per capita consumption distribution that fall within the strata and reflect the median prices that are faced by households located within a reasonable range around the level of consumption where the poverty line is expected to be.

[7] In the HIES 2000, 2005, and 2010, the large cities were defined based on the concept of Statistical Metropolitan Areas (SMA), following the IMPS sampling frame. This concept of SMA was replaced by the concept of Rural/Urban/City Corporation (RUC) in the *2011 Census of Population and Housing*. Of the 64 districts, only in three did the old SMA

Once the food poverty lines were estimated for each stratum, the second step consisted of computing non-food allowances using two different methods. In the first method, the non-food allowance was estimated by taking the median amount spent for non-food items by a reference group of households whose *total* per capita expenditure was close to the food poverty line. The non-food allowance estimated using this method is called the "lower non-food allowance." In the second method, the non-food allowance was estimated by taking the median amount spent for non-food items by a reference group of households whose *food* per capita expenditure was close to the food poverty line. The non-food allowance estimated using this method is called the "upper non-food allowance." Lastly, the food poverty lines were added to the lower and upper non-food allowances, and this yielded the official upper and lower poverty rates at the stratum level (16 upper poverty lines and 16 lower poverty lines). Table III shows a summary of when poverty lines were estimated for Bangladesh for the latest four rounds of the HIES. It is important to note that the update of the poverty lines across time involved a combination of re-estimation of lines in some years and inflation updates in other years.[8] .

TABLE III

POVERTY LINES IN HIES

Year	2000[9]	2005	2010	2016/17
Food PL	Updated from 1991/92	Re-estimated (CBN)	Updated from 2005	Updated from 2010
Non-food PL	Updated from 1991/92	Re-estimated (CBN)	Re-estimated (CBN)	Updated from 2010

2.3 Updating the Poverty Lines

The 2016/17 poverty lines took the 2010 poverty lines and adjusted them by inflation to keep them in real terms. The upper and lower poverty lines for each quarter were estimated by updating the official upper and lower poverty lines available for the HIES 2010 using price indices constructed for each quarter. The annual upper and lower poverty lines were updated using a set of price indices constructed with the full HIES 2016/17.

concept not match perfectly with the new RUC. Section IV discusses the comparability of strata across HIES and its implications for poverty measurement.

[8] For a detailed discussion on how the lines were updated across time from 2000 to 2010, see Jolliffe *et al*. (2013)

[9] The 2005 poverty lines were also back-casted to 2000.

For each quarterly and annual poverty line, a set of composite price indices was constructed for each of the 16 original strata[10] using a combination of the Törnqvist food price index and the non-food Consumer Price Index (CPI) for urban and rural areas.[11] The stratum-specific Törnqvist food price indices were constructed using a set of 13 food expenditure groups, including coarse rice, pulses, meat, potatoes, milk, fruits, sugar, fish, eggs, cooking oil, salt/spices, soft drinks, and betel/cigarette.[12] These food expenditure groups were selected because they represented some of the most frequently consumed items by households but also because they allowed minimizing the inherent issue of differences in item quality. For each of the food expenditure groups and stratum, the median unit-values were calculated, as well as the average budget shares using the 2010 and the 2016/17 data.[13]

Before calculating the median unit-values, outliers were identified and replaced.[14] An outlier was identified if the unit-value was above 2.5 standard deviations of the distribution within the strata. Those cases were replaced using median values from the lowest level (household) to the highest level (national) distribution. If the household reported more than nine observations for the item, the median of those values was used to impute the outlier at this level. If the household did not have enough observations, then the outlier was replaced by the median of the PSU, district, stratum, area (urban/rural), or national, with the condition that there were enough observations to compute the median at that level.

The *Törnqvist* food price indices for each of the food expenditure groups and each stratum *k* were calculated as follows:

---

[10] Section IV of this paper discusses the comparability of strata across time and the implications for poverty measurement.

[11] The *Törnqvist* price index was selected instead of the *Laspeyres* or *Paasche* indexes because it uses budget shares averaged between consecutive years, and therefore allows for changes in consumption patterns over time.

[12] Traditionally, the group of 13 food items used in the HIES to update the poverty lines does not perfectly overlap with the 11 food items used to estimate the poverty lines.

[13] Using the median unit-values instead of the mean unit-values for each group allows minimizing the issue of the difference in item qualities which is inherently present in the estimation of all unit values and also the effect of outliers.

[14] The replacement was done for 1.94 per cent of unit values reported in the daily consumption section and 2.36 per cent of unit values reported in the weekly consumption section.

$$\ln P_{10}^{Tk} = \sum_{j=1}^{n} \frac{w_{1j}^{k} + w_{0j}^{k}}{2} \ln\left(\frac{p_{1j}^{k}}{p_{0j}^{k}}\right) \quad (1)$$

where P^{Tk} denotes the *Törnqvist* price index for region *k*, *1* and *0* denote the two years of comparison (2010 and 2016/17 in this case), w^{k}_{1j} and w^{k}_{0j} are the respective budget shares, and p^{k}_{1j} and p^{k}_{0j} are the respective prices for good *j* in the two years of comparison.

Once the HIES-based *Törnqvist* food price indices were computed for each stratum, a set of stratum-specific composite price indices were constructed to update the poverty lines. These composite price indices were constructed by creating a weighted average of the non-food CPI inflation rate for urban and rural areas between 2010 and 2016/17 and the *Törnqvist* food price indices for each stratum. The relative weights used for this calculation of the composite price index were the stratum-level average food budget shares for 2010 and 2016/17. The non-food CPI inflation rate was computed using the average CPI from February 2010-January 2011 (data collection for the HIES 2010) and the average non-food CPI for each quarter in 2016/17 (e.g., April-June 2016/17 for Q1, July-September 2016/17 for Q2, October-December 2016/17 for Q3, and January-March 2017 for Q4), separated for urban and rural areas. The annual non-food CPI for 2016/17 was computed taking the average from April 2016 to March 2017. These composite price indices are used to update the 2010 lower and upper poverty lines to 2016/17. Quarterly poverty lines are presented in Annex 1 and annual poverty lines in Table IV.

TABLE IV
ANNUAL POVERTY LINES 2016/17

Stratum	HIES 2016/17	
	Lower	Upper
Barisal Rural	1778	2056
Barisal Urban	1993	2756
Chittagong Rural	2030	2439
Chittagong Urban	2135	2606
Chittagong City Corp.	2097	2660
Dhaka Rural	1835	2152
Dhaka Urban	1947	2657
Dhaka City Corp.	2020	2929
Khulna Rural	1677	2019
Khulna Urban	1817	2419
Khulna City Corp.	1942	2360
Rajshahi Rural	1716	2065
Rajshahi Urban	1864	2251
Rajshahi City Corp.	1764	2244
Sylhet Rural	1764	1865
Sylhet Urban	1911	2315

Source: Authors' calculations using HIES 2016/17.

III. POVERTY ESTIMATES

The latest HIES 2016/17 annual poverty estimates show that Bangladesh is continuing its remarkable progress in poverty reduction. Per the latest 2016/17 estimates, 24.3 per cent of the population lived in poverty, and 12.9 per cent were in extreme poverty (Table V). This represents a 24.6 percentage point reduction in the upper poverty rate since 2000 and 7.2 percentage points since 2010. Annex 2 presents the estimated poverty rates for all analytical domains.

Importantly, the HIES design is characterised by the following: (i) sampling weights; (ii) sampling of households within clusters or PSUs; and (iii) geographic stratification. These three elements need to be considered to compute adequate statistics using the survey. Using sampling weights (variable POPWGT) is important to calculate correct point estimates (e.g., poverty rate). In addition to the weights, the clustering (PSU variable) and stratification (ZILAID for annual estimates and STRATUM16 for quarterly estimates) of the survey design need to be considered to calculate the correct standard errors. If the analysis ignores the clustering of the survey design, it would produce standard errors that are smaller than they should be (for more details, see Ahmed *et al*. 2017).

TABLE V

NATIONAL POVERTY RATES
HIES 2000-2016/17

Year	Rate	Standard error	95% Confidence interval	
A. Upper poverty				
2000	0.489	0.012	0.464	0.513
2005	0.400	0.011	0.378	0.422
2010	0.315	0.010	0.296	0.334
2016/17	0.243	0.005	0.233	0.254
B. Lower poverty				
2000	0.343	0.012	0.319	0.367
2005	0.251	0.009	0.233	0.270
2010	0.176	0.008	0.160	0.191
2016/17	0.129	0.004	0.122	0.136

Source: Authors' calculations using HIES (various rounds).

TABLE VI

QUARTERLY NATIONAL POVERTY RATES HIES 2016/17

Year	Rate	Standard error	95% Confidence interval	
A. Upper poverty				
Q1 (April-June 2016)	0.225	0.014	0.199	0.252
Q2 (July-September 2016)	0.230	0.012	0.206	0.253
Q3 (October-December 2016)	0.261	0.012	0.238	0.284
Q4 (January-March 2017)	0.271	0.014	0.244	0.298
B. Lower poverty				
Q1 (April-June 2016)	0.1244	0.0092	0.1064	0.1423
Q2 (July-September 2016)	0.1231	0.0092	0.1051	0.1411
Q3 (October-December 2016)	0.1345	0.0085	0.1179	0.1511
Q4 (January-March 2017)	0.1406	0.0103	0.1204	0.1607

Source: Authors' calculations using HIES 2016/17.

IV. POVERTY RATES USING COMPARABLE STRATA AND CORRECTED URBAN CLASSIFICATION

As previously discussed, the substantial increase in the sample size of the HIES 2016/17 required using a different sampling frame to accommodate the larger number of PSUs. The PSUs for all the previous rounds of the HIES were selected from the Integrated Multiple-Purpose Sample (IMPS) – a master sample updated after each *Housing and Population Census*. In the HIES 2016/17, the PSUs come from the list of Enumeration Areas (EAs) used for Bangladesh's 2011 Population and Housing Census.

The use of a different sampling frame affected the comparability of strata across HIES. In this section, we describe the adjustments that need to be made to the HIES 2016/17 microdata to create comparable strata across time. In addition, we include a fix to the urban and rural definition to ensure a comparable and consistent classification of urban areas when using HIES. Finally, we present the poverty rates estimated using the comparable strata and correct urban/rural classification.

Within urban areas the comparability across time was affected, as the concept of SMA was abandoned in the 2011 census. The concept of SMA was replaced by the concept of Rural/Urban/CC (RUC) in the 2011 Census of Population and Housing. The strata used in HIES 2000, 2005, and 2010 were the divisions separated by urban, rural, and SMAs. Instead, the strata used in HIES 2016/17 were the divisions separated by urban, rural, and CCs.

Moreover, the Post-Enumeration Check Survey (PECS) conducted after the completion of the *2011 Household and Population Census* found that there was under-coverage both in urban and rural areas, but that it was more prevalent in

urban areas. BBS thus used a two-step approach to adjust the 2011 census estimates. First, it reclassified urban and rural areas using the concepts of: (i) growth centres, (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 per cent, which is the number that BBS has been using since then to produce official population projections and statistics. These adjustments (reclassification of areas and re-weighting) were also done in the HIES 2016/17 data to ensure a consistent urban share with the corrected 2011 census and with previous HIES rounds. However, 13 out of 2,304 enumeration areas in the HIES microdata were classified as rural when in fact they were urban. This classification error underestimates the urban share of the population. The urban share that is calculated directly from the HIES microdata is 27.3 per cent of the population, which is actually lower than the official share for 2011. The corrected share is 29.1 per cent of the population, which is more consistent with the urbanisation process observed in Bangladesh in the past years.

Annex 4 includes a STATA code that produces a comparable strata variable (replacing STRATUM16) with the previous HIES and corrects the misclassification error in the urban/rural variable. Based on these corrections, new poverty rates were estimated and are presented in Table VII. With these adjustments, the national poverty rate is 0.3 percentage points higher. The urban poverty rate increases from 18.9 to 19.5 per cent and the rural poverty rate also increases from 26.4 to 26.7 per cent. The division-level poverty rates are also presented in Table VII. None of the changes are statistically different from zero.

TABLE VII

POVERTY RATES WITH CORRECT URBAN CLASSIFICATION AND SMA COMPARABLE WITH 2010

	Upper poverty			Lower poverty		
	Official	Fixing urban classification only	Fixing urban classification and SMA comparable with 2010	Official	Fixing urban classification only	Fixing urban classification and SMA comparable with 2010
	(1)	(2)	(3)	(4)	(5)	(6)
National	0.243	0.245	0.246	0.129	0.130	0.130
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)
Rural	0.264	0.267	0.267	0.149	0.150	0.150
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)
Urban	0.189	0.193	0.195	0.076	0.080	0.080
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Poverty by division						
Barishal	0.265	0.264	0.264	0.145	0.144	0.144
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)

(Contd. Table VII)

	Upper poverty			Lower poverty		
	Official	Fixing urban classification only	Fixing urban classification and SMA comparable with 2010	Official	Fixing urban classification only	Fixing urban classification and SMA comparable with 2010
	(1)	(2)	(3)	(4)	(5)	(6)
Chattogram	0.184	0.183	0.186	0.087	0.090	0.090
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Dhaka	0.199	0.205	0.206	0.096	0.099	0.099
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Khulna	0.275	0.277	0.275	0.124	0.121	0.122
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Rajshahi	0.289	0.290	0.290	0.142	0.143	0.143
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Rangpur	0.472	0.473	0.473	0.306	0.306	0.306
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Sylhet	0.162	0.162	0.162	0.115	0.115	0.115
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)

Source: Authors' calculations using HIES 2010 and 2016/17.

V. ROBUSTNESS CHECKS

In this section, we investigate the sensitivity of the poverty rates to the imputation of education expenditures, correction of outliers in unit-values, and deflation within the year.

Correction of zeros and missing in education. Education expenditures were collected in Sections 2 and 9 of the survey. Traditionally, for the computation of the consumption aggregate, education expenditures are added using the information from section 9. In 2016/17, 5.6 per cent of households reported missing or zero education expenditures in Section 9, but had positive values in Section 2. In 2010, this was only true for 1.2 per cent of households. The current estimates for 2016/17 replace zeros and missing values in the consumption module with the information from the education section. This imputation is considered to be important for comparability with 2010.

Outlier adjustment of unit-values. When comparing the distribution of unit values between 2010 and 2016/17, it was found that the 2016/17 data had more extreme values. Table VIII presents the distribution of unit values at the national level for some key items as an example.

TABLE VIII

RIGHT-TAIL DISTRIBUTION OF UNIT VALUES AT THE NATIONAL LEVEL, HIES 2010 AND 2016/17

Item	2010				2016/17			
	mean	p95	p99	max	mean	p95	p99	max
Coarse rice	3	4	4	38	3	4	5	3600
Lentil (musur)	11	12	13	24	13	16	20	1841
Puti/Big Puti/Telapia/Nilotica	10	16	20	48	13	20	30	1400
Hen eggs	633	700	800	7000	877	1000	1050	85000
Beef	24	26	27	42	48	50	80	45000
Potato	1	2	2	14	2	3	4	3250
Liquid milk	4	5	6	12	7	8	12	9000
Sugar	5	6	6	14	8	10	25	10000
Mustard oil	13	20	20	25	16	25	50	10000
Ripe banana	5	9	10	13	17	15	500	10000
Soft drinks	5	8	9	16	14	12	50	6500
Cigarettes	149	325	600	1000	300	600	1100	35300

Note: Authors' calculations using HIES 2010 and 2016/17.

Two approaches to deal with unit values were compared: (i) identification of outliers using their distribution at the stratum level and imputation of unit values using median values from the lowest level possible (household) to the highest (national); (ii) identification of unit values using the distribution at the division level and imputation of median values of the division.

Quarterly inflation adjustment. Another option that was explored was to deflate the consumption aggregate within the year, to express all values to one quarter of the year. The objective of the adjustment was to test the importance of accounting for inflation within the year to calculate the 2016/17 poverty numbers.

Table IX presents the estimated upper and lower poverty rates under different adjustments. Overall, the imputation of education expenditures, outlier corrections, or deflating expenditures within the year do not change the poverty rates in a statistically significant sense. Analysis available by request also shows limited changes to the consumption distribution. Therefore, the preferred methodology was option 3, where education expenditures were imputed, and outliers were corrected using the distribution at the stratum level. This option was considered the most comparable to the 2010 methodology.

TABLE IX

NATIONAL POVERTY RATE 2016/17, DIFFERENT APPROACHES

Options	Mean	SE	95% Confidence interval	
A. Upper poverty				
1. Original	0.251	0.005	0.240	0.262
2. Imputing zeros and missings in education	0.248	0.005	0.237	0.259
3. Imputing zeros and missings in education + Outlier adjustment using stratum	0.243	0.005	0.233	0.254
4. Imputing zeros and missings in education + Outlier adjustment using division	0.240	0.005	0.230	0.251
5. Imputing zeros and missings in education + Outlier adjustment using stratum + quarterly inflation adjustment	0.247	0.005	0.236	0.258
B. Lower poverty				
1. Original	0.135	0.004	0.128	0.143
2. Imputing zeros and missings in education	0.133	0.004	0.125	0.140
3. Imputing zeros and missings in education + Outlier adjustment using stratum	0.129	0.004	0.122	0.136
4. Imputing zeros and missings in education + Outlier adjustment using division	0.127	0.004	0.120	0.134
5. Imputing zeros and missings in education + Outlier adjustment using stratum + quarterly inflation adjustment	0.131	0.004	0.123	0.138

Note: Quarterly inflation adjustment means that the consumption aggregate for Q2, Q3, and Q4 was expressed in prices of Q1.

Source: Authors' calculations using HIES 2016/17.

VI. OFFICIAL HOUSEHOLD INCOME ESTIMATES

Income in Bangladesh was defined as money inflows into the household occurring during the last 12 months. Household income was computed using a set of questions from the HIES and by adding together all the sources of family income described in detail below.

A. Labour income: The total labour income was defined as the total amount earned or received (in-cash or in-kind) for the last 12 months from each activity by household members aged five years and above who were engaged in economic activities and were classified as day-labourer or employees in agricultural and non-agricultural activities. The total labour income included other benefits that salaried workers received over the past 12 months (tips, bonuses, or transport). It was found that 7 per cent of wages for day-labourers and employees were missing. This was not the case in 2010, where only 1 per cent of this information was not reported. In this case, missing daily wages, net remunerations for salaried workers and other benefits were replaced by the median of the stratum and industry at the two-digit level when there were more than 30 observations. If that industry did not have enough observations at the stratum level, then the missing wage

was replaced by the area (urban/rural) median—conditioned to have more than 30 observations. Otherwise, national median per industry was used.

B. Business income: For households owning or running businesses, net revenue over the last 12 months was calculated as the difference between total gross revenue and total expenditures. The latter were estimated by adding up expenditures on wages, rent, raw materials, kerosene, electricity, expenditure on finished goods purchased for reselling, and other operating expenses in the past 12 months. This number was multiplied by the share of the company's profit that was owned by the household. One extreme gross revenue value reported by one household was fixed because there was an additional digit compared to the number reported in the physical questionnaire.

C. Agricultural income: This source of income was defined as total crop production consumed and sold by the household, and was computed by multiplying crops' unit values from total production by quantities sold and consumed. In the presence of missing values for quantities consumed and sold but complete information for total production, the total value of the latter was used. In addition, outliers in unit values were identified, when the value was above 3.5 standard deviations of each crop distribution within the strata. These cases were imputed by the median at the stratum level for each crop if the number of observations exceeded 30. Otherwise, the area (urban/rural) median unit value by crop was used if there were more than 30 observations. Crops without enough observations used the national median. Total livestock and poultry sold, and the total value of livestock products, fish, and forest products sold and consumed in the last 12 months were also included in the agricultural income.

D. Non-labour income: The family non-labour income was the sum of rent from land, rent from other properties, other profits and dividends received as partner or shareholder, interest from banks and other sources, social incomes such as insurances, lotteries, charities, or assistance in cash or kind, and gratuities, separation payments, or retirement benefits, all of them received during the past 12 months.

E. Other sources of income: Other sources of income included the total amount of remittances in cash and in kind sent in the last 12 months by household members who migrated to other districts inside the country or

abroad. Total payments received in cash or in kind in the last 12 months for all the household members currently enrolled in social safety nets programmes, stipends for household members who were currently studying, and self-reported imputed rent were also included.

Table X shows the average income, expenditure, and consumption expenditure. Expenditure was defined as total consumption plus lumpy life-cycle expenditures, income tax, and interest charges. Average income was computed only for positive values. In total, 0.61 per cent of the sample had a negative income, and 0.59 per cent reported zero income. Negative incomes were the result of negative profits for self-employed individuals and might not reflect the permanent income of the household. Zero incomes arose when none in the household earned any income during the last 12 months, or individuals earned non-monetary income such as charities, transfers, or social assistance, but there was a misreport, and this information was not recorded (Socio-Economic Database for Latin America and the Caribbean and World Bank 2014).

TABLE X

MONTHLY HOUSEHOLD NOMINAL INCOME, EXPENDITURE, AND CONSUMPTION, 2016

Residence	Average Monthly (Taka)		
	Income	Expenditure	Consumption expenditure
National	15,945	15,715	15,420
Rural	13,353	14,156	13,868
Urban	22,565	19,697	19,383

Note: Figures are not deflated spatially, as presented in the HIES 2016 preliminary report.

REFERENCES

Ahmed, F., D. Roy, M. Yanez-Pagans, and N. Yoshida. 2017. "Design of a Multi-Stage Stratified Sample for Poverty and Welfare Monitoring with Multiple Objectives: A Bangladesh Case Study." World Bank Research Working Paper. Washington, DC: World Bank.

Alamgir, M. 1974. "Some Analysis of Distribution of Income, Consumption, Saving and Poverty in Bangladesh." *Bangladesh Development Studies,* 2: 737-818.

Jolliffe, D., I. Sharif, L. Gimenez, and F. Ahmed. 2013. "Bangladesh - Poverty Assessment: Assessing a Decade of Progress in Reducing Poverty, 2000-2010." *Bangladesh Development Series*, paper no. 31. Washington, DC: World Bank.

National Institute of Population Research and Training (NIPORT), Mitra and Associates, and ICF International. 2016. *Bangladesh Demographic and Health Survey 2014.* Dhaka, Bangladesh, and Rockville, Maryland, USA: NIPORT, Mitra and Associates, and ICF International.

Ravallion, M. and B. Sen. 1996. "When Method Matters: Monitoring Poverty in Bangladesh." *Economic Development and Cultural Change*, 44: 761-792.

Socio-Economic Database for Latin America and the Caribbean and World Bank. 2014. *A Guide to SEDLAC Socio-Economic Database for Latin America and the Caribbean*. Methodological document.

ANNEX 1

Table A1.1: **HIES 2016/17 QUARTERLY POVERTY LINES**

stratum16	Lower poverty lines				Upper poverty lines			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Barishal Rural	1770	1802	1827	1829	2047	2085	2113	2116
Barishal Urban	1983	1977	2041	2020	2742	2733	2822	2793
Chattogram Rural	1974	2010	2087	2056	2373	2415	2508	2471
Chattogram Urban	2044	2153	2202	2193	2495	2629	2688	2677
Chattogram City Corp.	2039	2104	2172	2105	2587	2670	2756	2670
Dhaka Rural	1793	1837	1898	1882	2103	2154	2226	2208
Dhaka Urban	1894	1949	1928	1991	2584	2659	2631	2717
Dhaka City Corp.	1973	2013	2043	2032	2860	2919	2962	2946
Khulna Rural	1621	1663	1757	1703	1952	2003	2115	2051
Khulna Urban	1788	1796	1861	1836	2380	2391	2478	2444
Khulna City Corp.	1919	1913	1952	1982	2332	2325	2373	2409
Rajshahi Rural	1592	1677	1776	1740	1915	2018	2137	2094
Rajshahi Urban	1799	1834	1929	1903	2174	2216	2330	2299
Rajshahi City Corp.	1659	1767	1825	1850	2111	2248	2321	2354
Sylhet Rural	1706	1785	1826	1842	1804	1887	1931	1948
Sylhet Urban	1837	1833	1954	1952	2226	2221	2367	2365

ANNEX 2: HIES 2016/17 POVERTY ESTIMATES

Table A2.1: **National Upper Poverty Rates by Area, 2016/17**

Area	Mean	Standard error	95% Confidence interval	
Rural	0.26	0.01	0.25	0.28
Urban	0.19	0.01	0.16	0.21

Table A2.2: **National Lower Poverty Rates by Area, 2016/17**

Area	Mean	Standard error	95% Confidence interval	
Rural	0.15	0.00	0.14	0.16
Urban	0.08	0.01	0.06	0.09

Table A2.3: **National Upper Poverty Rates by Division, 2016/17**

Division name	Mean	Standard error	95% Confidence interval	
Barishal	0.26	0.02	0.23	0.30
Chattogram	0.18	0.01	0.16	0.21
Dhaka	0.16	0.01	0.13	0.19
Khulna	0.27	0.01	0.25	0.30
Mymensingh	0.33	0.02	0.29	0.37
Rajshahi	0.29	0.02	0.26	0.32
Rangpur	0.47	0.01	0.45	0.50
Sylhet	0.16	0.02	0.13	0.20

Table A2.4: **National Lower Poverty Rates by Division, 2016/17**

Division name	Mean	Standard error	95% Confidence interval	
Barisal	0.14	0.01	0.12	0.17
Chittagong	0.09	0.01	0.07	0.10
Dhaka	0.07	0.01	0.06	0.09
Khulna	0.12	0.01	0.11	0.14
Mymensingh	0.18	0.01	0.15	0.20
Rajshahi	0.14	0.01	0.12	0.16
Rangpur	0.31	0.01	0.28	0.33
Sylhet	0.11	0.01	0.09	0.14

Table A2.5: **National Upper Poverty Rates by District, 2016/17**

District name	Mean	Standard error	95% Confidence interval	
Bagerhat	0.31	0.04	0.23	0.40
Bandarban	0.63	0.08	0.48	0.78
Barguna	0.26	0.03	0.19	0.32
Barisal	0.27	0.03	0.21	0.34
Bhola	0.15	0.03	0.10	0.21
Bogura	0.27	0.04	0.20	0.34
Brahmanbaria	0.10	0.03	0.05	0.16
Chandpur	0.29	0.04	0.21	0.38
Chittagong	0.14	0.03	0.08	0.20
Chuadanga	0.32	0.03	0.26	0.37
Comilla	0.14	0.02	0.10	0.17
Cox's Bazar	0.17	0.04	0.09	0.25
Dhaka	0.10	0.04	0.03	0.17
Dinajpur	0.64	0.03	0.58	0.71
Faridpur	0.08	0.02	0.04	0.12
Feni	0.08	0.02	0.05	0.12
Gaibandha	0.47	0.04	0.40	0.54
Gazipur	0.07	0.01	0.04	0.10
Gopalganj	0.30	0.03	0.23	0.36
Habiganj	0.13	0.03	0.08	0.19
Joypurhat	0.21	0.03	0.16	0.27
Jamalpur	0.53	0.03	0.46	0.59
Jashore	0.27	0.03	0.21	0.33
Jhalokathi	0.22	0.02	0.17	0.26
Jhenaidah	0.27	0.04	0.18	0.35
Khagrachhari	0.53	0.08	0.38	0.68
Khulna	0.31	0.05	0.22	0.40
Kishoreganj	0.54	0.04	0.45	0.62
Kurigram	0.71	0.03	0.64	0.77
Kushtia	0.18	0.03	0.12	0.23
Lakshmipur	0.33	0.04	0.25	0.40
Lalmonirhat	0.42	0.05	0.33	0.51

(Contd. Table A2.5)

District name	Mean	Standard error	95% Confidence interval	
Madaripur	0.04	0.01	0.02	0.06
Magura	0.57	0.05	0.47	0.66
Manikganj	0.31	0.04	0.24	0.38
Meherpur	0.32	0.04	0.25	0.39
Maulvibazar	0.11	0.03	0.06	0.16
Munshiganj	0.03	0.01	0.01	0.05
Mymensingh	0.22	0.04	0.15	0.29
Naogaon	0.32	0.03	0.26	0.38
Narail	0.17	0.03	0.11	0.22
Narayanganj	0.03	0.01	0.01	0.05
Narsingdi	0.10	0.03	0.05	0.16
Natore	0.24	0.03	0.17	0.30
Chapai Nawabganj	0.40	0.03	0.34	0.46
Netrakona	0.34	0.04	0.27	0.41
Nilphamari	0.32	0.03	0.27	0.38
Noakhali	0.23	0.04	0.15	0.32
Pabna	0.33	0.03	0.27	0.39
Panchagarh	0.26	0.05	0.17	0.36
Patuakhali	0.37	0.05	0.27	0.47
Pirojpur	0.32	0.03	0.26	0.39
Rajshahi	0.20	0.07	0.07	0.33
Rajbari	0.34	0.03	0.28	0.40
Rangamati	0.29	0.05	0.20	0.37
Rangpur	0.44	0.04	0.37	0.51
Shariatpur	0.16	0.03	0.11	0.21
Satkhira	0.19	0.03	0.12	0.25
Sirajganj	0.30	0.04	0.23	0.38
Sherpur	0.41	0.04	0.33	0.50
Sunamganj	0.26	0.05	0.17	0.35
Sylhet	0.13	0.03	0.08	0.18
Tangail	0.19	0.03	0.13	0.25
Thakurgaon	0.23	0.04	0.17	0.30

Table A2.6: **National Lower Poverty Rates by District, 2016/17**

District name	Mean	Standard error	95% Confidence interval	
Bagerhat	0.14	0.03	0.08	0.20
Bandarban	0.50	0.08	0.35	0.66
Barguna	0.12	0.03	0.07	0.17
Barisal	0.14	0.03	0.09	0.19
Bhola	0.09	0.02	0.04	0.13
Bogura	0.14	0.02	0.09	0.18
Brahmanbaria	0.05	0.02	0.02	0.08
Chandpur	0.15	0.03	0.09	0.22
Chittagong	0.04	0.02	0.00	0.07
Chuadanga	0.12	0.01	0.10	0.15
Comilla	0.05	0.01	0.03	0.07
Cox's Bazar	0.08	0.03	0.01	0.14
Dhaka	0.02	0.01	-0.01	0.04
Dinajpur	0.45	0.03	0.39	0.52
Faridpur	0.03	0.02	0.00	0.07
Feni	0.03	0.01	0.01	0.06
Gaibandha	0.29	0.03	0.23	0.35
Gazipur	0.02	0.01	0.00	0.04
Gopalganj	0.15	0.03	0.10	0.21
Habiganj	0.10	0.03	0.05	0.15
Joypurhat	0.10	0.02	0.06	0.13
Jamalpur	0.35	0.03	0.29	0.42
Jashore	0.09	0.01	0.06	0.12
Jhalokathi	0.10	0.02	0.06	0.14
Jhenaidah	0.13	0.03	0.07	0.19
Khagrachhari	0.33	0.06	0.21	0.45
Khulna	0.14	0.03	0.08	0.19
Kishoreganj	0.34	0.05	0.25	0.44
Kurigram	0.54	0.04	0.46	0.62
Kushtia	0.07	0.02	0.04	0.10
Lakshmipur	0.20	0.03	0.14	0.27
Lalmonirhat	0.23	0.04	0.16	0.30
Madaripur	0.01	0.00	0.00	0.02
Magura	0.38	0.05	0.28	0.47
Manikganj	0.16	0.03	0.11	0.21
Meherpur	0.12	0.02	0.09	0.16
Maulvibazar	0.07	0.02	0.03	0.11
Munshiganj	0.01	0.01	0.00	0.03
Mymensingh	0.10	0.02	0.05	0.14
Naogaon	0.18	0.03	0.12	0.24
Narail	0.06	0.02	0.02	0.09

(Contd. Table A2.6)

District name	Mean	Standard error	95% Confidence interval	
Narayanganj	0.00	0.00	0.00	0.00
Narsingdi	0.05	0.02	0.00	0.09
Natore	0.13	0.02	0.08	0.17
Chapai Nawabganj	0.24	0.03	0.19	0.29
Netrakona	0.16	0.02	0.11	0.20
Nilphamari	0.14	0.02	0.11	0.18
Noakhali	0.13	0.03	0.07	0.19
Pabna	0.17	0.02	0.12	0.21
Panchagarh	0.14	0.04	0.07	0.21
Patuakhali	0.24	0.04	0.17	0.32
Pirojpur	0.18	0.03	0.12	0.23
Rajshahi	0.07	0.05	-0.02	0.16
Rajbari	0.16	0.03	0.11	0.21
Rangamati	0.11	0.03	0.06	0.16
Rangpur	0.27	0.03	0.22	0.32
Shariatpur	0.05	0.02	0.02	0.08
Satkhira	0.09	0.02	0.05	0.13
Sirajganj	0.12	0.02	0.08	0.16
Sherpur	0.24	0.03	0.18	0.31
Sunamganj	0.19	0.04	0.12	0.27
Sylhet	0.09	0.02	0.05	0.13
Tangail	0.09	0.02	0.04	0.13
Thakurgaon	0.15	0.03	0.10	0.21

ANNEX 3. ASSESSING CONSISTENCY OF HOUSEHOLD SIZE ESTIMATES

The average household size obtained from the HIES 2016/17 was 4.06 members. This implies a significant reduction in the average household size compared to the previous HIES 2010, which is not explained by differences in the definition of households (Table 3.1).

Table A3.1: **Average household size in HIES**

HIES	Mean	Standard error	95% Confidence interval	
2000	5.18	0.04	5.10	5.26
2005	4.85	0.03	4.78	4.91
2010	4.50	0.03	4.44	4.55
2016/17	4.06	0.02	4.03	4.09

Source: HIES 2000, 2005, 2010, 2016/17.

Other recent nationally representative surveys, like the Multiple Indicator Cluster Survey (MICS 2012/13) and Demographic and Health Survey (BDHS 2014) which have in principle consistent definitions of households, show a larger average household size – 4.57 and 4.69 members. However, more recent national representative surveys collected by BBS show consistent large reductions in household size. For example, the first quarter of the new Quarterly Labour Force Survey (QLFS 2015/16), collected between July and September 2015, shows an average household size of 4.2. Similarly, HIES estimates of the percentage of single-member households seems aligned with the most recent QLFS (Table A3.2.)

Table A3.2: **Household Size Based on Different Nationally Representative Surveys**

	HIES 2010	Population Census 2011	MICS 2012/13	LFS 2013	DHS 2014	QLFS 2015	HIES 2016/17
Average household size	4.50	4.45	4.57	4.30	4.69	4.20	4.06
Single-member households (%)	2.4	3.4	1.9	-	1.5	3.3	2.8

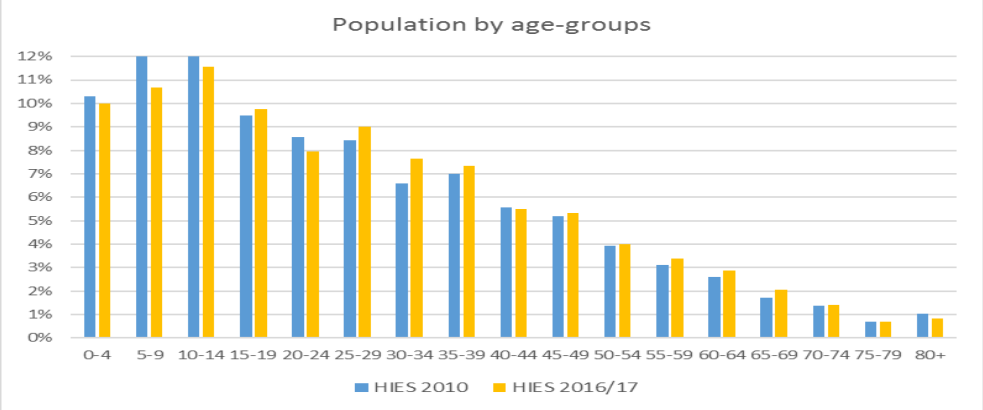
To assess the consistency of the average household size estimates based on the HIES 2016/17, we compared projections starting from a baseline using HIES 2000. Table A3.3 compares two types of projection (linear and compound) with observed estimates from HIES and the Census. The results suggest that the reduction in household size is consistent with an expected declining trend.

Table A3.1: **Projections of Household Size**

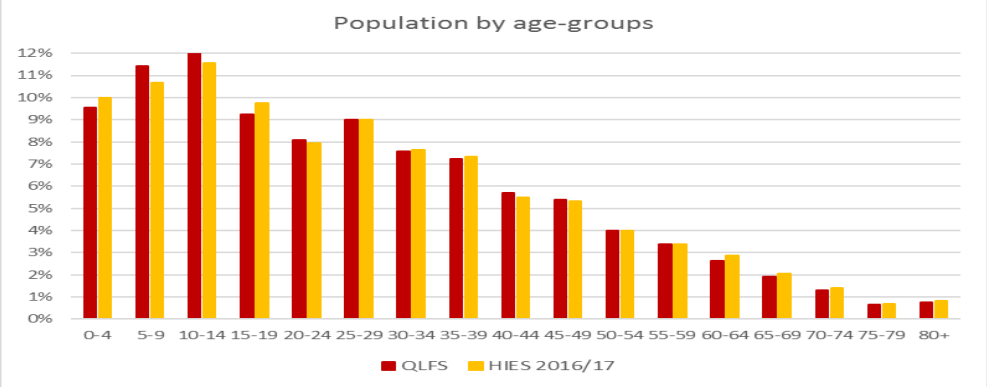
Name of survey or census	Direct estimation	Projections	
		Compound	Linear
HIES 2000	5.18	5.18	5.18
HIES 2005	4.85	4.83	4.85
HIES 2010	4.50	4.5	4.51
Population Census 2011	4.45	4.44	4.44
HIES 2016/17	4.06	4.12	4.07

In addition, we compared the population pyramids based on the HIES 2016/17 with the ones produced using the HIES 2010 data, the official BBS population projections (BBS 2015), and the QLFS. The different population pyramids estimated are shown in Figure A3.1, and none of them seem to suggest any strange pattern or important differences.

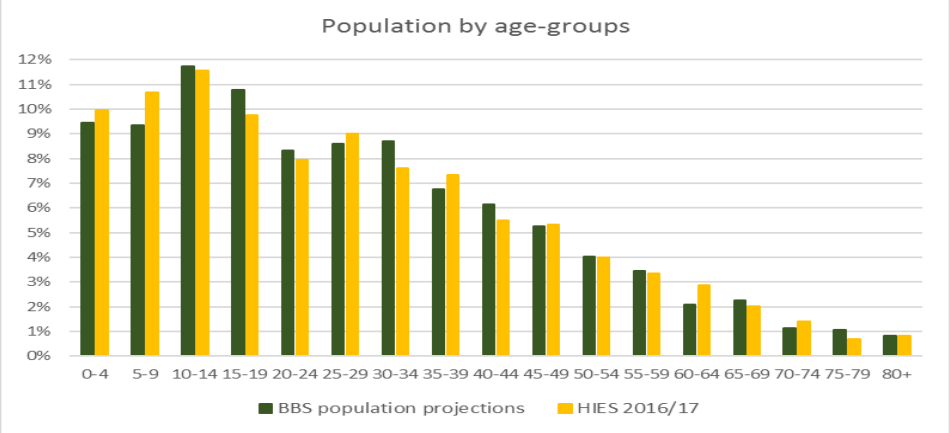
Figure A3.1: **Population Pyramids**
Panel A: **HIES 2010 versus HIES 2016/17**



Panel B: **QLFS versus HIES 2016/17**



Panel C: **BBS population projections versus HIES 2016/17**



Note: QLFS estimates reported are based on July-September 2015. BBS population projections are based on the official publication disseminated in 2015.

Consequently, there does not seem to be any reason to suspect that the average household size estimated based on the HIES 2016/17 round is much lower or inconsistent with what other official national representative surveys are suggesting.

ANNEX 4. STATA CODE TO PRODUCE COMPARABLE STRATA AND CORRECT THE MISCLASSIFICATION ERROR IN THE URBAN/RURAL VARIABLE

```
/*==================================================
1: Fix the Urban Rural misclassification
==================================================*/

use "$hies_raw/HH_SEC_A_Q1Q2Q3Q4",clear

gen ruc_new=ruc
*reclassify 259 households from rural to urban
#delimit
replace ruc_new=2 if inlist(psu
,222
,343
,641
,642
,705
,706
,928
,929
,1239
,1253
,1717
,1781
,1789
)
;
#d cr

compare ruc_new ruc

gen stratum16_new=stratum16

*Reclassify  40  households  from  Stratum  Chattogram  rural  to
Chattogram Urban
replace stratum16_new=4 if inlist(psu,222,343)

*reclassify  120 households from Dhaka Rural to Urban stratum
replace stratum16_new=7 if inlist(psu,641,642,705,706,928,929)

*reclassify  40 households from Khulna Rural to Urban stratum
replace stratum16_new=10 if inlist(psu,1239,1253)

*reclassify  60 households from Rajshahi Rural to Urban stratum
replace stratum16_new=13 if inlist(psu,1717,1781,1789)

compare stratum16_new stratum16
```

```
keep hhold ruc_new stratum16_new psu

lab var stratum16_new "Stratum 16 urban area fix"
lab var ruc_new "Rural Urban urban area fix"

#d;
la de stratum16
1      "Barishal Rural"
2      "Barishal Urban"
3      "Chattogram Rural"
4      "Chattogram Urban"
5      "Chattogram CC"
6      "Dhaka Rural"
7      "Dhaka Urban"
8      "Dhaka CC"
9      "Khulna Rural"
10     "Khulna Urban"
11     "Khulna CC"
12     "Rajshahi Rural"
13     "Rajshahi Urban"
14     "Rajshahi CC"
15     "Sylhet Rural"
16     "Sylhet Urban"
,modify
;

la de ruc
1      "Rural"
2      "Urban"
3   "City Corporation"
;
#d cr


*rename variable
rename (stratum16_new  ruc_new) (stratum16 ruc)

*label values
la val stratum16 stratum16
la val ruc ruc


/*==================================================
 2: Generate stratum 16 comparable across time
==================================================*/

*stratum comparable from stratum 16
gen stratum16_comparable=stratum16

*gen ruc comparable
gen ruc_comparable=ruc
```

```
*from Chattogram Urban to Chittagon CC (now becoming SMA) 20
households
replace stratum16_comparable=5 if inlist(psu,343)
replace ruc_comparable=3 if inlist(psu,343)

*from Khulna Urban to Khulna CC (now becoming SMA) 40 households
replace stratum16_comparable=11 if inlist(psu,1239,1253)
replace ruc_comparable=3 if inlist(psu,1239,1253)

*from Dhaka Urban to Dhaka CC (now becoming SMA) 120 households
#delimit
replace stratum16_comparable=8 if inlist(psu
,641
,642
,705
,706
,928
,929
)
;
replace ruc_comparable=3 if inlist(psu
,641
,642
,705
,706
,928
,929
)
;
#d cr


*Var labels
#d ;
la de stratum16_comparable
1      "Barishal Rural"
2      "Barishal Urban"
3      "Chattogram Rural"
4      "Chattogram Urban"
5      "Chattogram SMA"
6      "Dhaka Rural"
7      "Dhaka Urban"
8      "Dhaka SMA"
9      "Khulna Rural"
10     "Khulna Urban"
11     "Khulna SMA"
12     "Rajshahi Rural"
13     "Rajshahi Urban"
14     "Rajshahi SMA"
15     "Sylhet Rural"
```

```
16     "Sylhet Urban"
,modify
;

la de ruc_comparable
1      "Rural"
2      "Urban"
3   "SMA"
;
#d cr


la val stratum16_comparable stratum16_comparable
la val ruc_comparable ruc_comparable

lab var stratum16_comparable "Stratum 16 Comparable acrros time with
urban area fix"
lab var ruc_comparable "1 Rural 2 Urban 3 SMA Comparable acrros time
       with urban area fix"
```