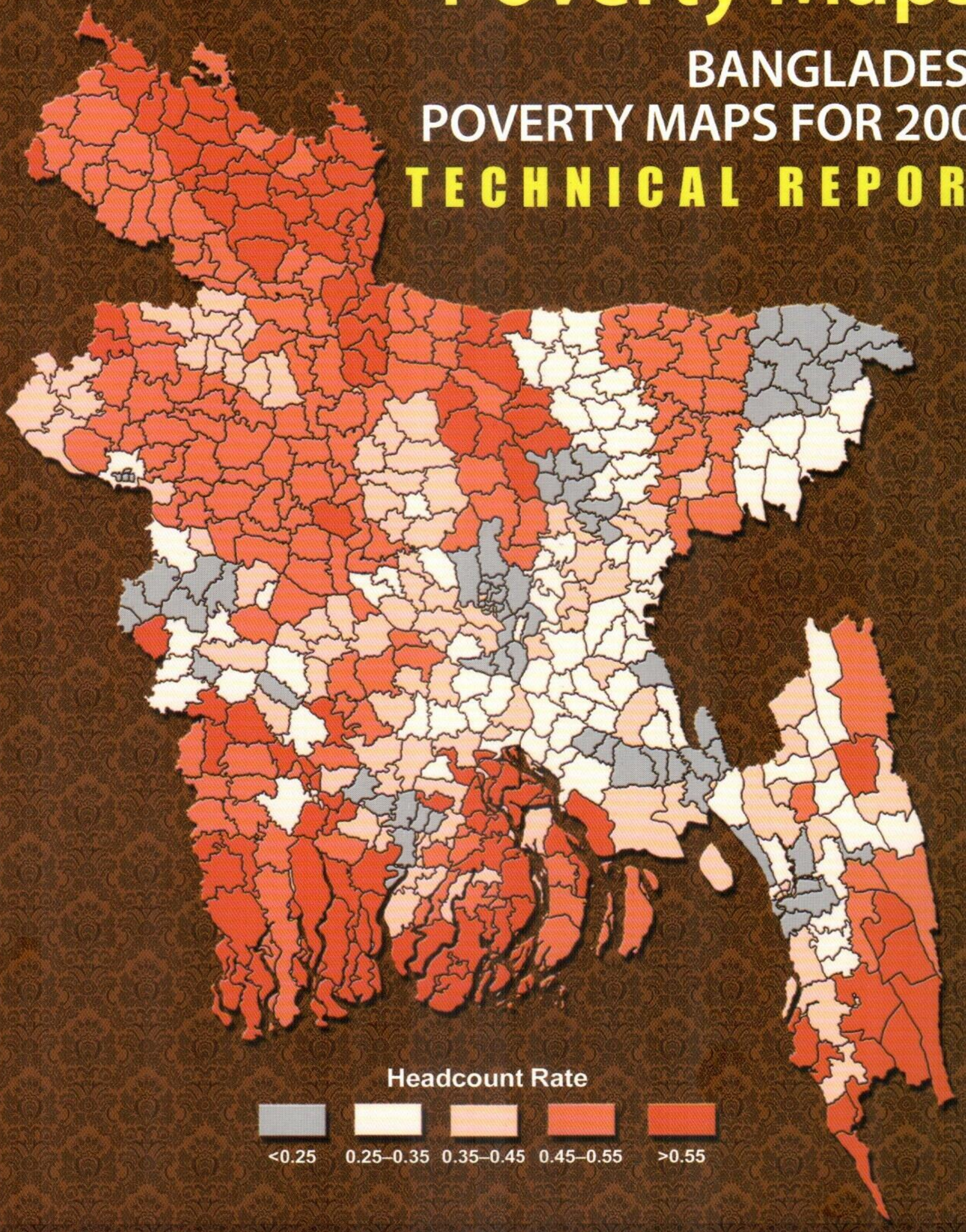
Updating Poverty Maps:

BANGLADESH POVERTY MAPS FOR 2005

TECHNICAL REPORT



THE WORLD BANK

WFP

Updating Poverty Maps:

BANGLADESH POVERTY MAPS FOR 2005

TECHNICAL REPORT

THE WORLD BANK

WFP

Disclaimer

Photo Courtesy

The World Bank

Designed and Printed by Macro Graphics Pvt. Ltd. www.macrographics.com

Table of Contents

List of Tables

List of Figures

ACKNOWLEDGEMENTS

This technical report summarizes the results and experience of the Bangladesh poverty map update that has been carried out by the Bangladesh Bureau of Statistics (BBS), the World Bank, and the United Nations World Food Programme (WFP).

We gratefully acknowledge the leadership and support of Mr. AYM Ekramul Hoque (the former Director General, BBS). Mr. Hoque made resources available for the team to expedite this exercise. We also acknowledge technical inputs and arrangements of Mr. Md. Shamsul Alam (the focal point of this poverty mapping exercise, BBS), Mr. Zahidul Hoque Sarder (Project Director, HIES, BBS) and Mr. Md. Abdul Latif (ASO, HIES, BBS). They contributed to the entire exercise by providing their in-depth knowledge on a wide variety of data available in the BBS and also made it easier for the team to access to the data. They also organized various meetings and a training event, such as poverty mapping training (June 2008), the technical committee meeting (June 2008), and the steering committee meeting (February 2009).

We would also like to thank WFP partner staff including Ms. Nusha Yamina Choudhury, Mr. John McHarris, and Mr. Mahabubul Alam. Ms. Choudhury and Mr. Alam produced many maps of natural disasters and educational attainment and led the perception survey to validate the results of poverty maps. Both Ms. Choudhury and Mr. McHarris provided their experience of the previous poverty mapping exercise and provided valuable comments on the preliminary results and the final maps.

The core task team in the World Bank consisted of Nobuo Yoshida (Task Team Leader), Faizuddin Ahmed, and Ericka Rascon. Ericka Rascon made a substantial contribution to building consumption models. Faizuddin Ahmed has been a center of this exercise and made significant contributions to producing poverty estimates at the Upazila level. Nobuo Yoshida has supervised this exercise while providing the latest international experience and technical innovations to the team.

The following persons have also made significant contributions to this exercise. Ambar Narayan started this exercise as a co-Task Team Leader and continued to provide useful advice and suggestions to the team. Tara Vishwanath provided technical advice as well as guidance on implementation of this exercise, especially coordination with stakeholders. Peter Lanjouw supported the team by providing solutions for some technical issues and informing the team of the state-of-art knowledge on the methodology and the latest international experience. Aphichoke Kotikula helped the team use the economic census data to improve the precision of poverty maps further. Ananya Basu provided various suggestions and advice to the team. Siobhan Murray converted the results of poverty mapping to maps and created various maps of other socio-economic indicators. Tahrat Shahid contributed to the completion of this report. Diepak Elmer and Mehrin Mahbub contributed to organizing a dissemination workshop and preparing a brochure – a brief summary of this report. Mildred Gonsalvez provided editorial support and helped organizing various training and workshops. Cheku Dorji also helped finalizing this report.

We would also like to thank members of both the technical committee and the steering committee for their useful comments and suggestions. The technical committee members carefully reviewed the methodology as well as the preliminary results in June 2008.

Their comments, which were incorporated in the final maps, improved the quality of poverty maps. The steering committee meeting was held in February 2009. After careful reviews on the final poverty maps, the members endorsed the findings and cleared wider dissemination.

We greatly acknowledge Department for International Development for their financial support and useful comments. The generous support enabled us to provide technical assistance for capacity building at BBS and facilitated the dissemination process.

Finally we would like to thank Md. Shahjahan Ali Mollah (Director General of BBS), Xian Zhu (the former country director of Bangladesh of World Bank), Ellen Goldstein (the current country director of World Bank), and John Aylieff (Country Representative of WFP) for their continued support and guidance.

1 INTRODUCTION

Bangladesh has experienced considerable poverty reduction, especially since 2000. Poverty incidence, which was as high as 57 percent at the beginning of the 1990s, declined to 49 percent in 2000. This trend accelerated subsequently, reducing the poverty headcount rate to 40 percent in 2005.

However, growing inequality among regions is a concern. There is evidence to suggest that the eastern region has increasingly benefited from integration with growth centers, namely Dhaka and Chittagong, in contrast to the more isolated regions in the west and southwest. The two largest river systems, the Brahmaputra/Jamuna and the Ganges/Padma, crisscross the country and appear to act as natural boundaries by imposing strong connectivity/trade barriers. Given the spatial inequality in growth and poverty reduction, policy interventions and foreign aid are likely to be far more effective if resources can be allocated and distributed based on local level poverty data.

Poverty maps and poverty estimates at even sub-district levels, are not new in Bangladesh. Several poverty maps were produced using Household Income and Expenditure Survey (HIES) 2000 data. However, there has been a growing demand for new poverty maps using the latest information such as HIES 2005 data.[1]

In response to this demand, the BBS, the World Bank, and WFP have updated maps of poverty estimates at up to the *upazila* (sub-district) level for 2005. Poverty mapping is an exercise to estimate poverty incidence at a level where a typical household income and expenditure survey cannot produce statistically reliable poverty estimates due to high sampling errors. In Bangladesh, official poverty rates are not produced below *division* level where sampling errors of HIES data become non-negligible. Various poverty mapping methodologies were devised to overcome increasing imprecision of poverty estimates as they are disaggregated.

A poverty mapping methodology used for this exercise is the Small Area Estimation (SAE) method developed by Elbers, et al. (2003). This methodology is one of the most commonly used poverty mapping methodologies around the world and has been widely tested and validated. The SAE method used to produce the new poverty maps of Bangladesh is based on two primary data sets; the HIES 2005 survey and the Census from 2001. The method takes advantage of the strengths of both sources: in the case of the HIES data its strength is associated with the fact that direct measures of poverty (i.e., income and expenditure data) are available, whereas in the case of the Census data its strength is associated with its size, meaning that data were collected from all households in the country as opposed to 'sampled' from a primary sampling unit.

The Bangladesh poverty map update exercise faced some technical challenges. For example, the interval between Population Census and HIES is relatively long. In Bangladesh, the latest Population Census was fielded in 2001 and the latest HIES was in 2005. The four years interval could cause a substantial bias in poverty estimates if appropriate treatments were not undertaken. In addition, a recent study by Tarrozi and Deaton (2008) showed the importance of incorporating regional

1 The BBS and WFP with technical support from Massey University, New Zealand produced poverty maps for 2001 using the 5 percent sample of Population Census 2001 and the HIES 2000. IRRI also produced a rural poverty map using HIES 2000.

variations in consumption patterns into SAE methods.

A novelty of the Bangladesh poverty map update is that it did not only attempt remedies for the above technical challenges but it also involved validation exercises to check whether potential bias was successfully mitigated. These challenges are not new, and solutions to resolving such concerns and new ways of validating poverty mapping results have recently been proposed (e.g., Elbers et al. 2008). This exercise made use of these proposed solutions and conducted new validation exercises showing some evidence that biases due to the above problems were minimal.

Another noteworthy aspect of this exercise is the establishment of strong country ownership of the maps. The government took an initiative to scrutinize results by organizing a technical committee meeting in June 2008 and a steering committee meeting in February 2009. The technical committee reviewed the quality of preliminary results and the recommendations were reflected in the final version of maps. After carefully reviewing the results of the final maps, the steering committee endorsed the results of poverty maps as robust and reliable, and cleared wider dissemination of the results. The World Bank's efforts for capacity building facilitated this country ownership, providing capacity building at the BBS by organizing a training session and sharing user-friendly software, PovMap2, with the BBS and other stakeholders in Bangladesh.

The objective of this report is to describe, in detail, how this Bangladesh poverty map update was conducted. For those readers whom are not interested in technical detail, a more general summary level brochure, entitled 'Updating Poverty Maps of Bangladesh', is recommended.[2]

The structure of this report herein is as follows. Section II discusses the SAE method and data used, includes technical challenges, and explains how the method was executed. Section III illustrates the results, while Section IV shows the results and interpretations of the validation exercises. Section V gives a brief conclusion.

2 It can be found and downloaded at: http://www.bbs.gov.bd/dataindex/povertymb.pdf

2 Modeling Exercises

The Elbers et al. (2002, 2003) small-area estimation (SAE) procedure offers a powerful approach to produce statistically reliable poverty estimates for small areas. Historically, poverty has been measured on the basis of sample survey consumption data in which household per capita expenditures are compared against a poverty line set by the government. Under this approach the sampling error of poverty estimates rise rapidly as the target area gets smaller. This precludes analysts from estimating poverty at the local level. In the present SAE method also shares a similar tendency, but the increases in the standard errors of poverty estimates are far slower than in the traditional method. As a result, reasonably precise poverty estimates can be obtained at the district, and even at the sub-district level.

However, implementation of the SAE method is fairly complex and, without careful implementation, resulting estimates may be unreliable. A number of studies have been undertaken that document these risks and that describe how to best apply the method. In this exercise, an exhaustive range of tests and checks have been applied along each step of the process. This section describes the methodology, data needs, and implementation process (including various tests and checks) of the Bangladesh Poverty Mapping.

HIES 2005 was collected by the BBS, and includes 10,800 households and 16 strata. Most variables are representative at the division level. The survey collected detailed information on consumption and income, and the data contains rich information on employment, ownership of assets, housing condition, and access to services such as education and health. The large set of variables helps precise imputation of household consumption into the census.

Along with the Population Census 2001 and HIES 2005, the Population Sample Census (PSC) 2004 data are also used in this exercise. The PSC 2004 data includes a wealth of information on household demographics, employment, educational attainments, health outcomes, asset ownership, and migration conditions. The PSC 2004 includes 150,000 households and most of the variables are representative at the *zila* level. As described in Section IV, this dataset was used to validate the findings from the poverty mapping exercise.

A. Methodology

The selection of poverty mapping methodology is critical; numerous methods are available and have been documented by Bigman and Deichmann (2000). An SAE method developed by Elbers et al. (2003) (henceforth referred to as ELL) has gained wide popularity amongst development practitioners around the world.

This Bangladesh poverty map update adopted the SAE method developed by ELL. It imputes consumption levels into census households based on a consumption model estimated from the household survey. In order for this to be possible, the consumption model must include explanatory variables (household and individual characteristics) that are available in both the census and the survey. By applying the estimated coefficients to the "common" variables from the census data, consumption expenditures of census households are imputed. Poverty and inequality statistics for small areas are then calculated with the imputed consumption of census households.

One advantage of this method is that it does not only estimate poverty incidence but also estimates standard errors of poverty estimates. Since poverty estimates are computed based on imputed consumption, they cannot escape imputation errors, which are their standard errors. ELL analyzed the properties of such imputation errors in detail and derived a procedure to compute standard errors of poverty estimates. Please see Box 1 for greater detail on this method.

B. Main Data Sources

The SAE method generally makes use of household survey and population census data. The Bangladesh poverty map update is no exception, using the unit record Population Census 2001 data and HIES 2005 data. The census data was collected by the BBS, and covered roughly 30 million households. A wide range of household information was collected including religion, educational attainments, labor activities, residential information, and employment and housing conditions. As is the practice in all countries, the Bangladesh Census did not include household consumption and income levels, but its wide coverage of household characteristics is an advantage for imputing household consumption precisely.[3]

HIES 2005 was collected by the BBS, and includes 10,800 households and 16 strata. Most variables are representative at the division level. The survey collected detailed information on consumption and income, and the data contains rich information on employment, ownership of assets, housing condition, and access to services such as education and health. The large set of variables helps precise imputation of household consumption into the census.

Along with the Population Census 2001 and HIES 2005, the Population Sample Census (PSC) 2004 data are also used in this exercise. The PSC 2004 data includes a wealth of information on household demographics, employment, educational attainments, health outcomes, asset ownership, and migration conditions. The PSC 2004 includes 150,000 households and most of the variables are representative at the *zila* level. As described in Section IV, this dataset was used to validate the findings from the poverty mapping exercise.

3 The poverty map of 2001, which was produced by BBS and WFP, used only a 5 percent sample of the census data since the full census data were not available then.

C. Technical Challenges

The ELL poverty mapping methodology continues to evolve in response to ongoing scrutiny from researchers. To this end a variety of documents and manuals are available on the World Bank website to inform development practitioners of the latest developments and methodological improvements in the SAE method. These improvements are also reflected in the updated versions of the PovMap2 software produced by the World Bank to assist with application of the procedure.

In the context of this Bangladesh poverty map update, two major technical challenges were apparent: (i) Long interval between Population Census 2001 and HIES 2005; (ii) Tarrozi and Deaton (2008) critique.

(i) Long Interval between Population Census 2001 and HIES 2005

Using ELL's method, the Bangladesh poverty mapping update derives a consumption model in HIES 2005 by regressing household expenditure on a set of proxies from household and individual characteristics, which are also available in Population Census 2001. The model is then used to predict household expenditure for each census household.

This approach works well if Census 2001 reflects the situation of 2005 properly. This assumption is, however, doubtful given the fact that Bangladesh experienced substantial economic growth between 2001 and 2005. As a result, consumption patterns might have changed dramatically between 2001 and 2005. Also, population distribution might have changed substantially due to active migration. Both changes in consumption pattern and population distribution can cause sizable biases and standard errors in poverty estimates derived from ELL's method.

BOX 1: The small area estimation method developed by ELL (2003)

The method proposed by ELL has two stages. In the first part, a model of log per capita consumption expenditure ($\ln y_{ch}$) is estimated in the survey data:

$$\ln y_{ch} = X_{ch}'\beta + Z'\gamma + u_{ch}$$

where X_{ch}' is the vector of explanatory variables for household *h* in cluster *c*, β is the vector of regression coefficients, Z' is the vector of location specific variables, γ is the vector of coefficients, and u_{ch} is the regression disturbances due to the discrepancy between the predicted household consumption and the actual value. This disturbance term is decomposed into two independent components: $u_{ch} = \eta_c + \varepsilon_{ch}$ where a cluster-specific effect, η_c and a household-specific effect, ε_{ch}. This error structure allows for both a location effect – common to all households in the same area—and heteroskedasticity in the household-specific errors. The location variables can be any level – *Zila, Upazila*, Union, *Mauza*, and Village – and can be drawn from any data sources that include all locations in the country. All parameters regarding the regression coefficients (β, γ) and distributions of the disturbance terms are estimated by Feasible Generalized Least Square (FGLS). In the second part of the analysis, poverty estimates and their standard errors are computed. There are two sources of errors involved in the estimation process: errors in the estimated regression coefficients ($\hat{\beta}$, $\hat{\gamma}$) and the disturbance terms, both of which affect poverty estimates and the level of their accuracy. ELL propose a way to properly calculate poverty estimates as well as their standard errors while taking into account these sources of bias. A simulated value of expenditure for each census household is calculated with predicted log expenditure $X_{ch}'\hat{\beta} + Z'\hat{\gamma}$ and random draws from the estimated distributions of the disturbance terms, η_c and ε_{ch}. These simulations are repeated 100 times. For any given location (such as a *zila* or an *upazila*), the mean across the 100 simulations of a poverty statistic provides a point estimate of the statistic, and the standard deviation provides an estimate of the standard error.

(ii) Tarrozi and Deaton (2008) Critique

In a recent contribution, Tarrozi and Deaton (2008) highlighted a number of concerns with the ELL methodology. Notably, they show that, under certain circumstances, the ELL method can result in an overly optimistic assessment of the statistical precision of the poverty map estimates. The present India Poverty Mapping Pilot has paid special attention to this concern and has undertaken a number of robustness checks to gauge its applicability.

The specific concerns raised by Tarrozi and Deaton (2008) can be summarized as follows. First, differences in consumption patterns can bias both poverty estimates and the standard errors. The ELL method estimates a consumption model that is assumed to apply to all households within each domain. The implicit assumption is that the relationship between household expenditures and its correlates is the same for all households within the domain, and that all remaining differences are due not to structural factors, but are attributable to errors. This is not a minor assumption and is explicitly acknowledged as such in ELL (2003).

Second, Tarrozi and Deaton (2008) caution that the misspecification in the error structure can lead to overstating the precision of poverty estimates. PovMap2, the software used for poverty mapping, in its current configuration can incorporate only two layers of errors (or residuals): at the levels of the household and at the level of some unit of aggregation above the household. In the case of this Bangladesh poverty mapping update, in addition to household level, errors at *mauza* level were incorporated in a consumption model. This does not mean, however, that there is no correlation in errors at the level of *zila* or *upazila*. Tarrozi and Deaton (2008) show that under some conditions, ignoring the *zila* or *upazila* level correlation can cause a large bias in standard errors of poverty estimates. An obvious solution for this issue is to introduce multiple layers of errors during the consumption modeling. However, this is not a practical solution for practitioners since PovMap2 currently allows only two layers of errors, as mentioned above.

Alternative remedies to resolve this issue were explored in the Bangladesh Poverty Mapping Pilot. These are suggestive, but are not able to entirely remove the potential concern. As a result, a set of additional validation exercises were undertaken to buttress the poverty map results on the basis of indirect empirical evidence (see Section IV).

D. Construction of the Bangladesh Poverty Maps of 2005

The poverty mapping procedure comprises two main components: selecting sound consumption models and selecting the level of disaggregation. This sub-section describes this process in detail. Final models are listed in Table A-3 of Annex 1.

As can be seen below, careful execution of poverty mapping is critical despite the convenience and user-friendliness of PovMap2 – new software developed by the World Bank's research department. PovMap2 facilitates this process, providing various statistics to help us undertake the above selections properly. Nevertheless, it is worth noting that the software cannot solve all problems and technical challenges, and thus users need to check every step carefully.

Model Selection

(a) The number of consumption models

The Bangladesh poverty mapping prepared 16 different consumption models, each corresponding to a stratum defined for the HIES 2005. As mentioned earlier, failure to capture regional differences in consumption patterns could bias poverty estimates produced with the ELL method. Regional differences in consumption patterns can often be substantial. For example, the educational attainment of household heads might be a good predictor of household wealth in urban areas, whereas it might not be as important in rural areas where the agricultural sector dominates.

Despite potential heterogeneity across areas, increasing the number of consumption models does not necessarily improve the statistical performance of poverty mapping. As the number of models rises, the sample size in the HIES 2005 data for each model declines, lowering the accuracy and stability of the consumption model.

In order to balance between capturing regional heterogeneity and maintaining adequate sample sizes it was decided to create a consumption model for each HIES 2005 stratum, resulting in 16 consumption models. This choice seems appealing since the sampling frame of the HIES 2005 data is stratified at the stratum level.

(b) Explanatory power of consumption models

Both R-square and Adjusted R-square provide information on how well a consumption model can predict the actual consumption expenditure of each census household. More specifically, R-square is a statistic that indicates how well predicted expenditure from a consumption model fits actual household expenditure. The higher the R-square, the

TABLE 1: R-square (R2) and adjusted R-square (adjR2)

Stratum	Name	R2	adjR2
1	Barisal (Rural)	0.40	0.39
2	Barisal (Muni.)	0.54	0.52
3	Chittagong (Rural)	0.46	0.45
4	Chittagong (Muni.)	0.52	0.50
5	Chittagong (SMA)	0.58	0.57
6	Dhaka (Rural)	0.40	0.39
7	Dhaka (Muni.)	0.44	0.43
8	Dhaka (SMA)	0.49	0.48
9	Khulna (Rural)	0.37	0.36
10	Khulna (Muni.)	0.59	0.58
11	Khulna (SMA)	0.51	0.49
12	Rajshahi (Rural)	0.33	0.33
13	Rajshahi (Muni.)	0.46	0.45
14	Rajshahi (SMA)	0.60	0.57
15	Sylhet (Rural)	0.38	0.36
16	Sylhet (Muni.)	0.68	0.66

Source: World Bank staff estimation using HIES 2005 data with Population Census 2001, Economic Census, and Natural Disaster data.

better predicted expenditure fits actual household expenditure. Adjusted R-square is a modification of R-square that adjusts for the number of terms in a model. R-square always increases when a new variable is added to a model, but Adjusted R-square increases only if the new variable improves the model more than would be expected by chance.

In the Bangladesh poverty mapping update, both R-square and Adjusted R-square are in general high. Eleven out of 16 models record an Adjusted R-square of over 40 percent and only one model (Rajshahi Rural) records an Adjusted R-square of below 35 percent (see Table 1).

(c) Share of variance of residuals at the mauza level

The consumption model cannot capture all the variation in household expenditures and the unexplained variation is accounted for by residuals (or simply errors).

The consumption model cannot explain all variations in household expenditure and the unexplained variations will go to residuals (or simply errors). These have two layers in the present analysis – household and cluster ("*Mauza*" in rural areas and "*Ward*" for urban areas). The cluster effect is included since consumption expenditures can be affected by region specific factors that are common across households, some of which may be observable while others not. The cluster effect is included since consumption expenditures can be affected by region specific factors that are common across households, some of which may be observable while others not.

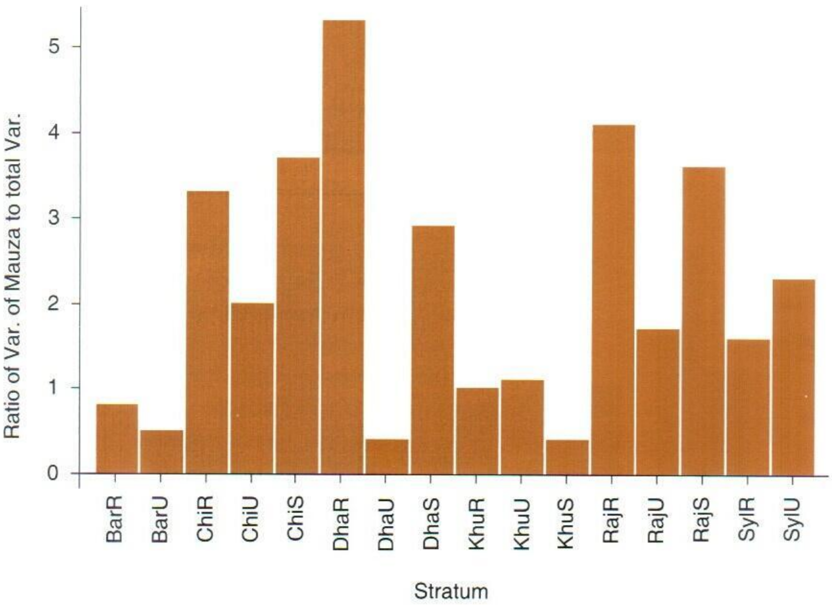
Since residual location effects such as cluster effects can reduce the precision of poverty and inequality estimates, ELL (2002, 2003) recommend applying great effort to capturing variation in consumption by observables as far as possible. A rule of thumb is to reduce the share of the variance of the cluster effect to the total variance of residuals to 10 percent or lower. International experience suggests that in rural areas achievement of this goal often remains elusive (see Mistiaen, et al. 2002).

One strategy for reducing the share of the variance of the cluster effect is to include location specific variables in the regression models. Such location specific variables can be constructed by aggregating data from the Population Census and can also be drawn from the village and town directory. They can be constructed not only at the cluster level but also at other administrative unit levels. Expending effort along these lines has been found to be of great importance also in addressing the concerns raised by Tarozzi and Deaton (2008) regarding the precision of poverty map estimates.

In the case of the Bangladesh Poverty Mapping Pilot, *upazila* level errors are not explicitly controlled for. However, as was shown by Elbers et al. (2008) for the case of Brazil, adding location specific variables at the cluster level helps reduce not only village level errors but also errors at a higher level (in this case the *upazila* level).[4]

The strategy outlined above has been quite successful in the Bangladesh Poverty Mapping. For all regions, the variance of cluster (mauza or ward) level errors constitutes less than 6 percent of the total variance (see Figure 1). In general, urban areas record a lower contribution of the town level errors.

FIGURE 1: Contribution of *mauza* level residuals



Source: World Bank staff estimation.

4 "Brazil within Brazil: Testing the Poverty Map Methodology in Minas Gerais," Policy Research Working Paper World Bank, WPS4513.

(d) The impact of the errors at the upazila and union levels

An ideal solution for the aforementioned problem of potential high errors in consumption models at the *upazila* and *union* levels is to introduce multi-layers of cluster effects to the consumption model. However, this is practically difficult since the poverty mapping software allows only one layer of cluster effects.

TABLE 2: Impact of switching the cluster from *mauza* to *upazila* on standard errors of the *upazila* level poverty estimates (%)

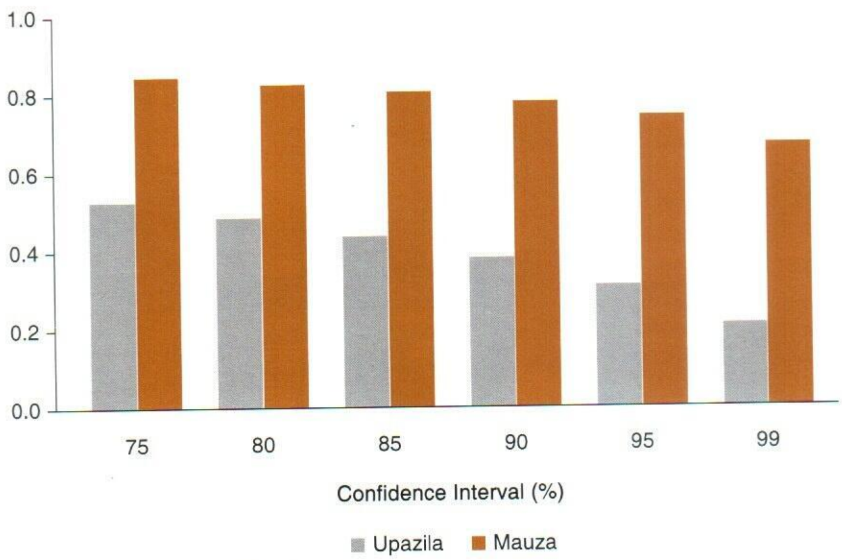
Percentile	1%	Median	99%
Mauza	0.6	1.7	5.7
Upazila	1.4	6.6	10.7

Source: World Bank staff estimation.

Note: Rows "Mauza" and "Upazila" correspond to the standard errors if a cluster effect is set at *mauza* and *upazila*, respectively.

Instead, Elbers et al. (2008) propose two tests for the level of risk associated with underestimation of standard errors of poverty estimates. One way to test for this is to switch the level of cluster from *mauza* to *upazila* or *union* and compare the proportion of statistically distinguishable rankings among *upazilas*. Switching a cluster from a smaller unit to a larger unit tends to increase standard errors of poverty estimates and thus reduces the proportion of statistically distinguishable rankings. We thus expect that the standard errors of poverty estimates will rise as the cluster level shifts from *mauza* to *union*, and from *union* to *upazila*.

FIGURE 2: Proportion of statistically significant rankings



Source: World Bank staff estimation.

In reality, a consumption model tends to have errors at *mauza* level; ignoring *mauza* level errors likely exaggerates the size of standard errors, producing a more pessimistic outcome than the reality, while including only *mauza* level errors likely understates the size of standard errors and is too optimistic about the precision of poverty estimates. The true level of standard errors of poverty estimates thus lies somewhere in between.

Standard errors of poverty estimates in fact rise significantly after shifting a cluster effect from *mauza* to *upazila* (see Table 2). If the cluster effect is set at the *upazila* level, the median standard error of the *upazila* level poverty estimate rises from 1.7 percent to 6.6 percent.

Although the results demonstrate a significant deterioration in the precision of poverty estimates, they are encouraging for the following reasons. First, even the 99th percentile of the *upazila* level estimate is not far off from stratum level estimates when poverty mapping is not applied. For example, the poverty estimate for Rajshahi metropolitan areas (stratum number 14) has a standard error of 9.6 if poverty mapping is not used. Second, as mentioned above, the results here are exaggerating the standard error, since only the *upazila* level cluster is introduced. The true standard error must be lower than if the cluster is set at the *upazila* level, but might be higher than if the cluster is set at the *mauza* level.

Reflecting the increase in standard errors, the proportion of statistically significant rankings declines considerably after switching the cluster effect from *mauza* to *upazila*. Around half of the rankings of upazila poverty rates are statistically distinguishable if 75 percent confidence intervals are adopted, even after switching the cluster effect from *mauza* to *upazila* (Figure 2).

The second approach proposed by Elbers et al. (2008) is to carry out a simple multi-layer maximum likelihood model. This approach allows more than two layers of errors, but it is still limited in that it does not introduce as complex a heteroskedasticity model as PovMap2 and applies a different optimization method from PovMap2 (maximum likelihood). Despite these differences, it is still useful to see the relevance of *upazila* and *union* level errors using this approach.

The contributions of *upazila* and *union* level errors are measured by the ratio of variances of these errors to the total variance and presented for each stratum separately in Annex (Table A-1). Note that there are some strata that do not have any number because the maximum likelihood estimation does not converge. Non-convergence seems to occur more frequently as complexity of error structure increases.

In most strata, the contribution of both *upazila* and *union* level errors (or if not available, the *upazila* level error) is limited to no more than 5 percent. Most errors are concentrated in the *mauza* and the household levels, of which PovMap2 can take explicit account. Preferred results are obtained likely because many area specific characteristics are included in the consumption models.

(e) Reducing incidence of trimming

A further set of important model selection criteria is associated with the handling of outliers in the simulated household expenditures of census households. The ELL method simulates household expenditure for all census households by randomly drawing parameters (including both regression coefficients and residuals) from their corresponding distributions as estimated in the survey-based consumption model. One issue with this method is that random drawing can potentially pick extreme values, albeit with low probability. Simulated household expenditures can thus include a few outlier values. PovMap2 allows for the elimination of such outliers by dropping them before estimating poverty and inequality indicators. Such an adjustment, which is often called 'trimming,' is needed since a few outliers can produce huge biases, especially in inequality statistics. However, trimming is more of a practical solution than one derived from rigorous statistical theory. In this sense, it would always be preferable if a consumption model could be specified from which a need for trimming did not arise.

TABLE 3: Incidence of trimming at various levels

	Share of trimmed simulated expenditures (%)			
Percentile	**Stratum**	**Zila**	**Upazila**	**Union**
Median	0.7	0.2	0.2	0.1
95%	2.0	1.8	1.8	1.9
Max	2.0	5.1	13.6	52.2

Source: World Bank staff estimation.

Table 3 summarizes the incidence of trimming at four different administrative unit levels if the final consumption models were adopted. For each level, it ranks all corresponding administrative units by incidence of trimming (the share of trimmed simulated expenditures), and shows the median, the 95th percentile, and the maximum number. Table 3 ensures that all strata and *zilas* involve very low incidence of trimming. At the *upazila* and *union* levels, the incidence of trimming is still low except for the largest five percent. However, the maximum number at the *union* level is as high as 52.2 percent – over half of simulated expenditures were dropped before estimating poverty headcount rates. This is one reason why we believe the Bangladesh poverty estimations of 2005 can be disaggregated up to the *upazila* level, but not below. This point will be revisited later.

(f) Mitigating the bias due to a long interval between census 2001 and HIES 2005

To mitigate issues arising from the long interval between Population Census 2001 and HIES 2005, only variables whose means did not change much at the stratum level among HIES 2000, Population Census 2001, and HIES 2005, are selected for the consumption modeling.[5] For example, household size did

5 More specifically, we estimated the 95 percent confidence interval of all variables for all HIES 2000, Census 2001 and HIES 2005 and check whether they are overlapped. Due to sampling errors, a difference in means from two or three datasets does not necessarily mean the true means are also different. To incorporate the effects of sampling errors, the 95 percent confidence intervals are calculated. Even if the means are different, if the 95 percent confidence intervals are overlapped, the difference is not statistically significant.

not change much in most divisions and strata. Although a reasonable number of variables satisfy this condition, it certainly limits model fitness, that is, the accuracy of predicting the true level of household expenditure.

Another way to mitigate this issue is inclusion of location specific variables. Household and individual variables are problematic since the source of data differs between consumption modeling and poverty simulations. As described above, household and individual characteristics are drawn from HIES 2005 at the consumption modeling stage, while these characteristics are drawn from Population Census 2001 at the poverty simulation stage. On the other hand, location specific variables from the same source can be used at both stages as long as they are available for all administrative units. One example is the construction of census means of household size at the *union* level. This variable can be merged with HIES 2005 data and included in a consumption model. Also, since Census 2001 data includes this variable, it can also be used to predict the household expenditure of census households and thus simulate poverty estimates. The inclusion of location specific variables also helps to mitigate the problems raised by Tarrozi and Deaton (2008), discussed in further detail below.

(g) The level of disaggregation

As noted earlier, ELL's method produces margins of error in poverty estimates, which can be used to practitioners in find the appropriate level of disaggregation of poverty estimates. Although most statistics of this type are associated with certain margins of error, results of poverty maps are frequently reported without providing any information about such errors. PovMap2 provides both poverty estimates and their standard errors.

As revealed in Table 4, standard errors are reasonably small at the stratum, *zila*, and *upazila* levels. For example, the largest standard error among all *upazilas* is just 5.7 percentage points. In addition, the performance at the *union* level is relatively good, 6.5 percentage points at the 95[th] percentile; however, the *union* level maximum reaches nearly 30 percentage points. This means that, given the fact that the 95 percent confidence interval lies in the range of +/- two standard errors, the true rate of poverty incidence of this *union* can be anywhere between 0 and 100 percent with 95 percent of probability. The relative standard error – a measure dividing the standard error by the corresponding mean – also suggests a similar conclusion: poverty maps can be disaggregated to the *upazila* level at a reasonable level of statistical accuracy.

Simulation results also support this decision. At up to the upazila level, very few administrative units record a high incidence of trimming. Only at the upazila level does the maximum figure reach 13.6 percent. At the union level, one union lost more than half of simulated expenditures because they are unusually low or high. Such a high incidence of trimming certainly reduces the reliability of poverty estimate of the union.

TABLE 4: Assessment of simulation results at various levels

	Standard Errors of Poverty Estimates (%)				Relative Standard Errors (%)			
Percentile	**Stratum**	**Zila**	**Upazila**	**Union**	**Stratum**	**Zila**	**Upazila**	**Union**
Median	0.8	1.3	1.7	3.3	2.1	3.4	4.6	9.3
95%	3.2	2.9	4.0	6.5	13.7	12.0	13.6	24.2
Max	3.2	5.4	5.7	29.4	13.7	25.6	63.0	124.8

Source: World Bank staff estimation.

Note: All numbers are calculated using the upper poverty lines of 2005.

3 Illustrations of the Bangladesh Poverty Maps of 2005

This section illustrates the results of the Bangladesh Poverty Maps of 2005. The presentation of these results is important, as careful illustration of results will help readers draw useful information from the maps.

A. Poverty Map vs. Extreme Poverty Map

First, we compare two poverty maps called "Poverty Map" and "Extreme Poverty Map" based on a different set of poverty lines (see Figure 3). The BBS produced official poverty headcount rates based on two sets of poverty lines, the upper poverty and lower poverty lines.

Both poverty lines, which are defined for each of 16 strata, consist of food and nonfood poverty lines while they adopt the same levels of food poverty lines, they adopt different levels of nonfood poverty lines. The upper poverty lines select higher allowances for nonfood consumption than the lower poverty lines. Such differences in nonfood allowances emerge since there is no clear consensus on what level of nonfood consumption is needed to sustain a minimum standard of living.[6] In Bangladesh, poverty defined by lower poverty lines is often called "Extreme Poverty" among development practitioners, while poverty defined by upper poverty lines is simply called "Poverty."

Both "Poverty" and "Extreme Poverty" are concentrated in the North-West, and the coastal areas. However, both maps suggest that areas between Dhaka city and Chittagong city are in general better-off than other areas. Poverty rates are generally lower in the Poverty Map than in the Extreme Poverty Map, since "Extreme Poverty" is defined using lower poverty lines than "Poverty."

It is nevertheless noteworthy that part of northwest and the coastal areas exhibits similarly high poverty rates in both maps. This suggests that severe deprivation is concentrated in these areas.

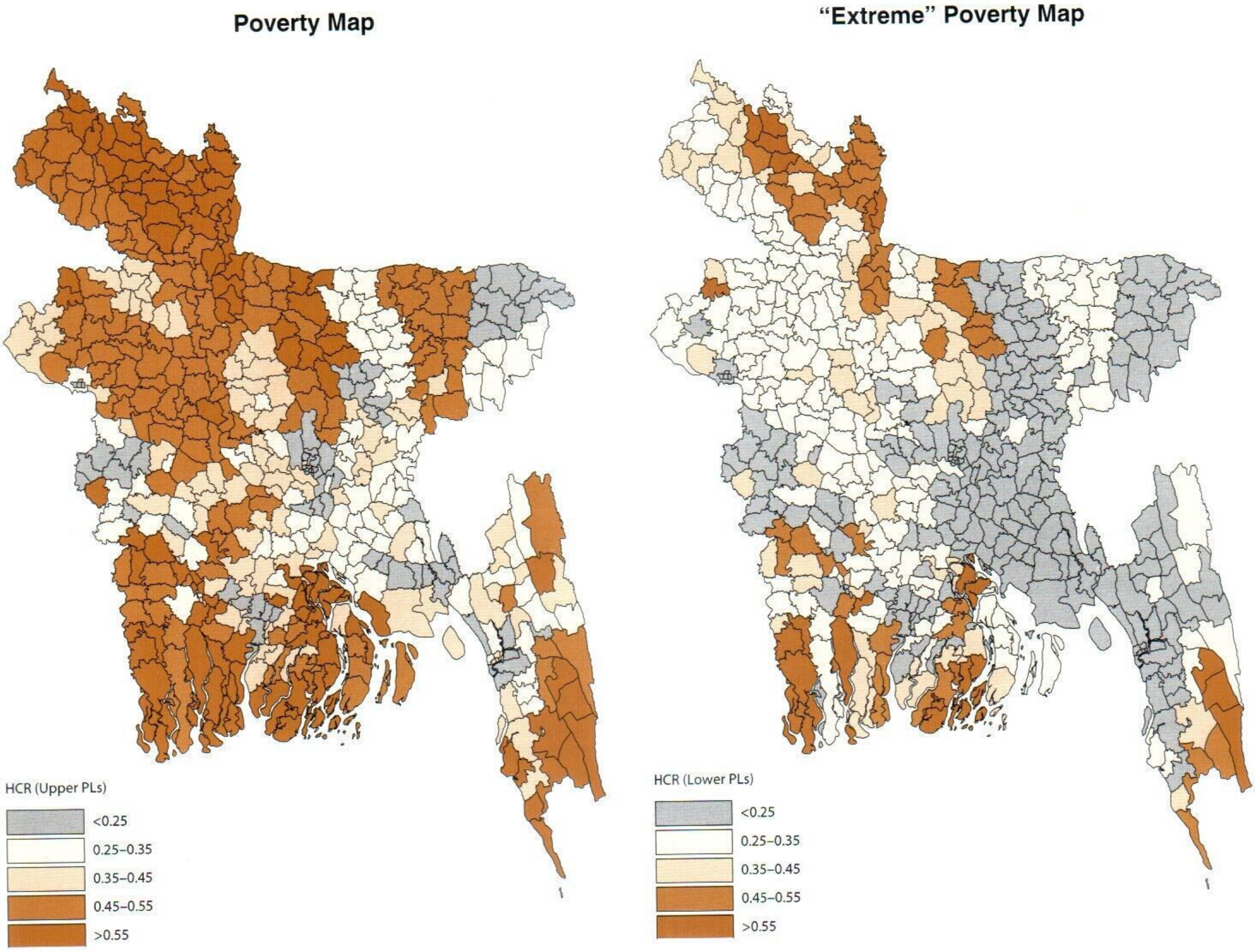
Both the Poverty Map and Extreme Poverty Map are useful but for different purposes. One possible use of the Extreme Poverty Map is to prioritize areas with high incidence of "Extreme Poverty" in terms of resource allocations if resources are limited. Also since areas with higher prevalence of extreme poverty are areas where food insecurity is likely to be more severe, those planning interventions that have food security as their primary objective might wish to prioritize their focus on the Extreme Poverty map as opposed to the Poverty map.

B. Power of Disaggregation

Next, we show the power of disaggregation of poverty estimation by comparing the Division level poverty map with the *upazila* level poverty map (see Figure 4). Note that all maps include poverty headcount rates based on the upper official poverty lines. As the World Bank (2008) argues, the east-west division is clear in the Division map, although this is not as clear in the *upazila* map. According to the *upazila* map, coastal areas appear to be poor irrespective of whether they are located in the west or in the east. Also, there are pockets of severe deprivation in the west, even near the main growth poles of Dhaka and Chittagong, while there are pockets of affluence in the east. Nevertheless, all maps confirm that areas between Dhaka and Chittagong are in general

6 See Ravallion and Sen (1996) for more details.

FIGURE 3: Bangladesh poverty maps of 2005



Source: World Bank Staff estimation.

Note: The Poverty Map was produced using the upper poverty lines while the "Extreme" Poverty Map was produced using the lower poverty lines.

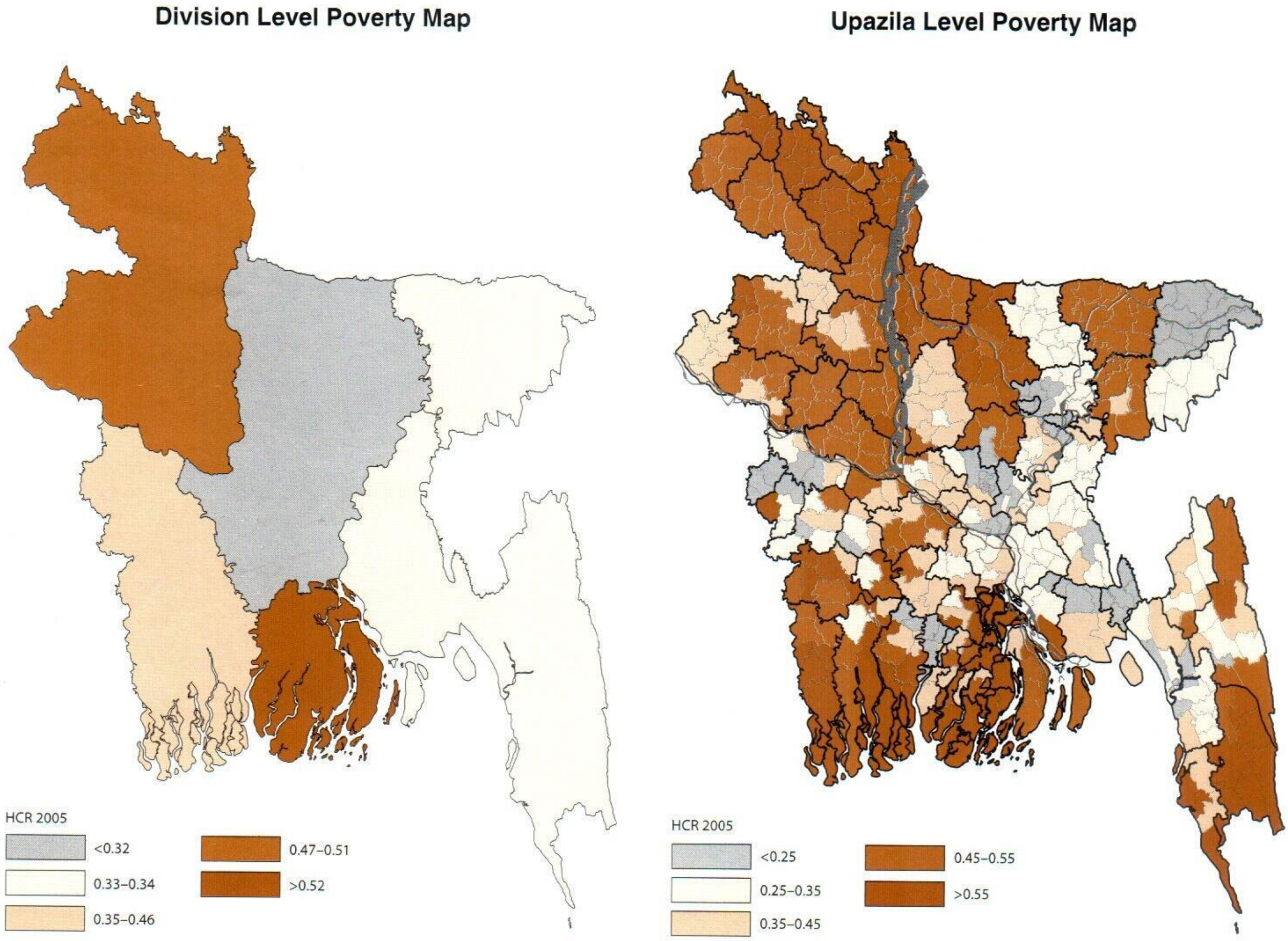
better off. Further investigation of these areas would be useful to find effective growth strategies with local conditions in mind.

C. Poor Population vs. Poverty Headcount Rates

It is interesting to compare a map of poverty headcount rates and a map of poor populations. A poverty headcount rate, a percentage of poor population in a given area, often receives more attention from governments, civil societies, and development partners than an absolute size of poor population. However, if a policy goal is to eradicate poverty, identifying areas with high concentration of poor population is equally or even more important than identifying areas with high poverty rates. This is particularly so since areas with high concentration of poor population often exist in areas with low poverty rates.

The maps in Figure 5 below, illustrate the important difference between poverty rates (headcount rates), and poor population sizes. The brown circle on the left map shows Dhaka's poverty headcount rate as relatively low, as indicated by the lighter shades of brown. By contrast, the brown circle on the map on the right shows a greater number of

FIGURE 4: Poverty maps (based on upper poverty lines) at different levels of spatial disaggregation



Source: World Bank Staff estimation.

Note: All poverty maps are produced using the upper poverty lines.

dots within the same area, indicating that the absolute size of Dhaka's poor population is relatively large.

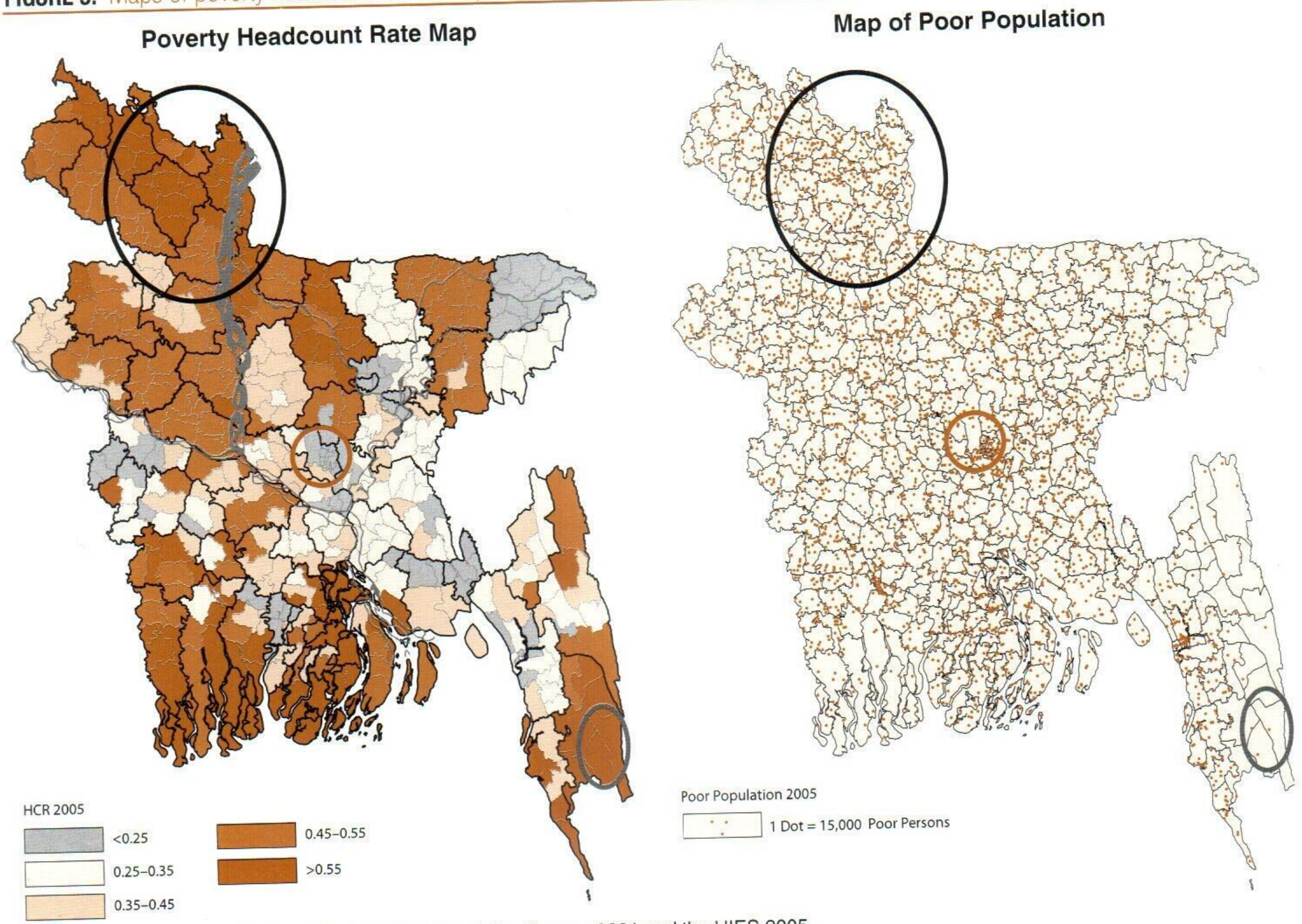
Other areas within Bangladesh have both relatively high poverty rates and relatively large poor population sizes. The 'Monga' (seasonal hunger) areas in the northwest of Bangladesh are one such example, and this area is highlighted with black colored circles on both maps below. The dark brown shades within the black circle on the left map indicate high rates of poverty within Monga areas, while the relatively large number of dots within the same area on the right map indicate that the absolute size of the poor population living within Monga areas is relatively large.

Lastly, combinations of high poverty rates and relatively small poor population sizes are also possible within the same given area, as demonstrated by the grey circles in the extreme Southeast of the country on both maps below. This area, is part of the 'Chittagong Hill Tracts/CHT' region, where poverty rates can be quite high (as indicated by the dark brown shades on the map on the left) despite the fact that the absolute size of the poor population is relatively small as indicated by the few dots within the grey circle on the map on the right.

Poverty and Inequality

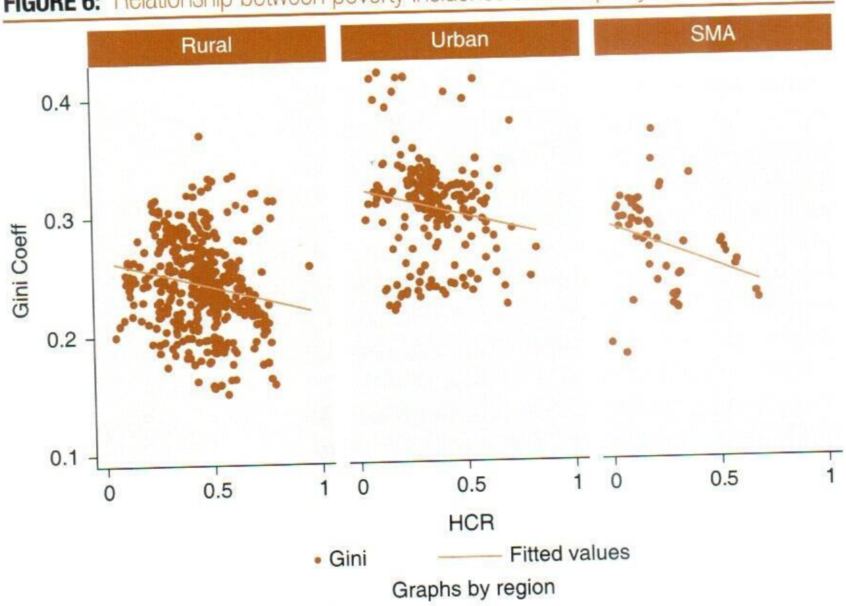
The Bangladesh poverty map update has produced not only poverty headcount rates

FIGURE 5: Maps of poverty headcount rates and poor populations at the *upazila* level for 2005



Source: World Bank staff estimation using the Population Census 2001 and the HIES 2005.

FIGURE 6: Relationship between poverty Incidence and inequality at the *upazila* level



Source: World Bank staff estimation.

but also various inequality measures such as Gini coefficients at up to the *upazila* level. Both poverty headcount rates and Gini coefficients are estimated for Statistical Metropolitan Area (SMA), urban areas, and rural areas separately. Although it is easy to aggregate poverty headcount rates across regions, doing the same for Gini coefficients is far more challenging. As a result, we compare poverty incidence with inequality for each region separately. This type of exercise is interesting because the relationship between poverty and inequality could differ substantially across regions.

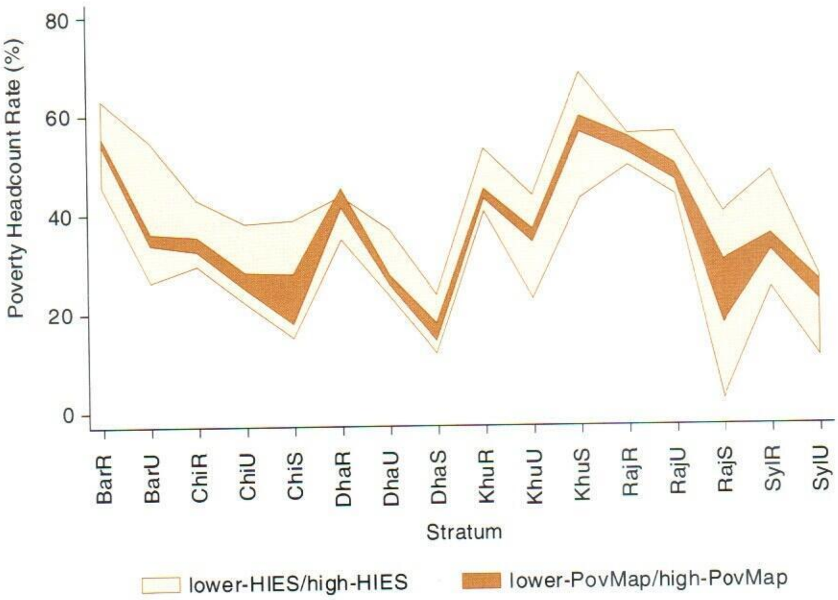
Figure 6 summarizes these results. Each dot represents an *upazila's* combination of a poverty headcount rate and a Gini coefficient, while the line represents a linear projection. The figure illustrates the trend of higher poverty headcount rates corresponding with lower inequality. There is thus a tendency that many people are equally poor in areas with high poverty rates. It is nevertheless also true that there is large heterogeneity within a region and the observations based on the linear projection should be viewed with caveats in mind. For example, some SMA *upazila* exhibit among the lowest inequalities and poverty rates in the country.

This comparison across regions suggests that urban areas tend to have higher inequality than the other two types of regions. In terms of a linear projection, the SMA exhibits the steepest relationship between poverty headcount rates and Gini coefficients.

4 VALIDATION EXERCISES

The Bangladesh Poverty Mapping Update incorporated a number of validation exercises including: (i) confirming consistency in the stratum level poverty estimates between direct estimation from HIES 2005 and the Small Area Estimation (SAE) method; (ii) comparing the poverty map based on Census 2001 with that of PSC 2004; (iii) comparing a perceptions-based poverty map with the updated Poverty map; and (iv) checking for consistency between the poverty map and other regional characteristics.

FIGURE 7: Comparison between the direct estimation from HIES 2005 and the Small Area Estimation (SAE) method



Source: World Bank staff estimation.

A. Consistency in the Stratum Level Poverty Estimates between Direct Estimation from HIES 2005 Data and the Small Area Estimation Method

One way to check the reliability of estimates from the ELL method is to compare them with the corresponding numbers estimated directly from the HIES 2005 data.

Key variables in the HIES 2005 data are stratified at the stratum level, The ELL method can obviously generate estimates at the stratum level as well. Presumably, if underlying assumptions of within-region homogeneity and of relative stability between 2001 and 2005 do not hold, there would be little reason to expect estimates based on the ELL method to be close to those from the HIES data directly. Conversely, if the ELL method produces a good predictor of true poverty incidence, it should be consistent with that estimated from HIES 2005 data.

Consistency checks are applied using the 95 percent confidence intervals of both estimates. Both poverty estimates are statistics rather than true levels, and their 95 confidence intervals reflect the margins of errors of the poverty estimates. These two estimates can be considered as consistent if the 95 percent confidence intervals are overlapping.

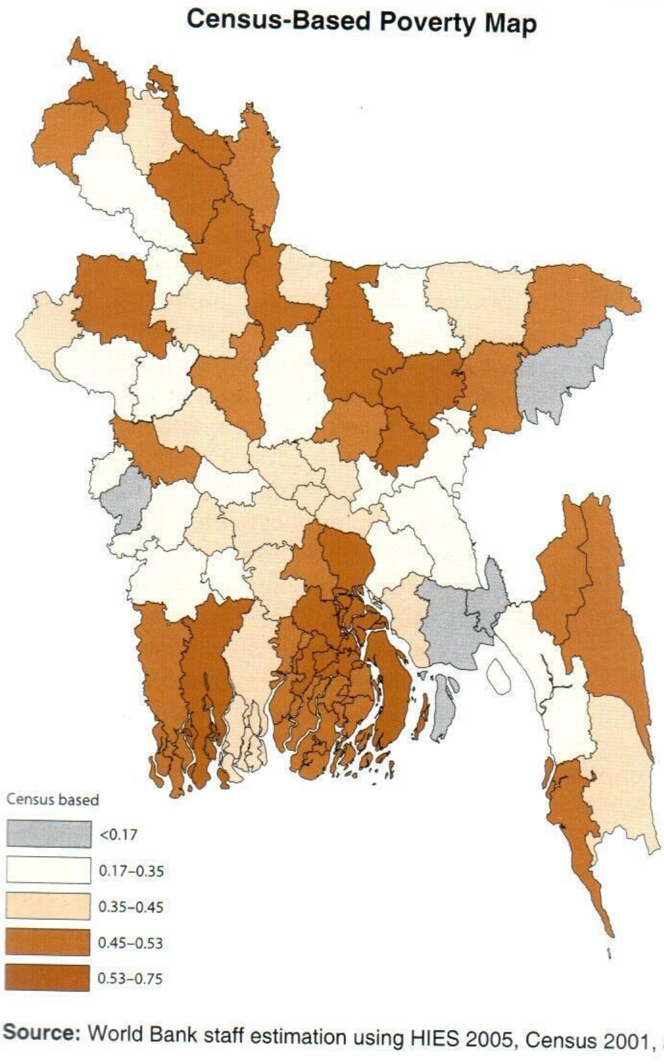
As Figure 7 shows, both estimates are overlapping in all strata except for rural Dhaka but the mismatch is minute. Another interesting observation is that poverty estimates from the ELL method appear to contain much smaller margins of error than the estimates of the direct estimation from the HIES 2005. This reflects the fact that estimates directly from the HIES survey are based on far fewer data points than are those based on the population census.

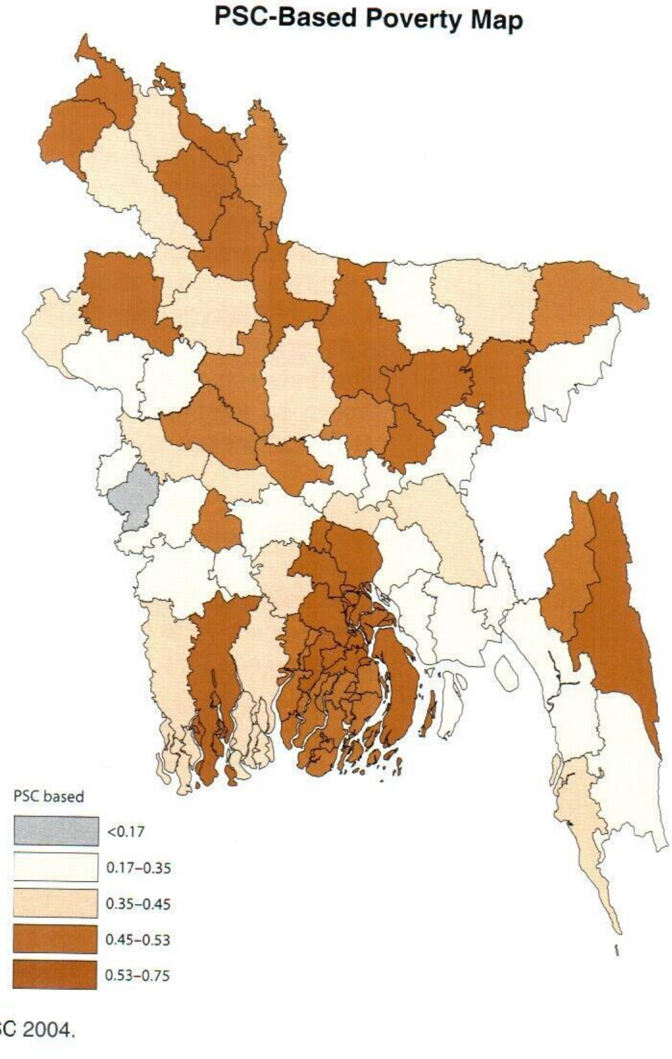
B. Creation of Poverty Maps Using HIES 2005 and Population Sample Census (PSC) 2004

Poverty maps based on PSC 2004 (instead of Census 2001) are produced to see whether the long interval between Census 2001 and HIES 2005 cause biases in poverty estimates. Since PSC 2004 was conducted just one year before HIES 2005, the poverty maps based on PSC 2004 are unlikely to be vulnerable to the potential bias caused by the long interval between Census 2001 and HIES 2005. If the poverty maps based on Census 2001 are very similar to those of PSC 2004, the potential bias is likely to be low. Nevertheless, since PSC 2004 is representative at the *zila* (district) level, the comparison is therefore conducted only at the *zila* level.

The poverty maps based on PSC 2004 were also used to make assessments of the impact of population movements. Since PSC 2004

FIGURE 8: Population sample census 2004





Source: World Bank staff estimation using HIES 2005, Census 2001, and PSC 2004.

reflects population in 2004, if the poverty maps based on the Census 2001 are similar to those of PSC 2004, the poverty impact of population movements is likely to be limited.

Also, the PSC 2004 data helps us measure the impact of migration on poverty estimates. If migrants are significantly poorer or better off than the populations that they are migrating into, the poverty estimates using the Census 2001 data would either under or over estimate poverty rates. Using migration data available in PSC 2004, we can assess how much difference can be made by migration between 2001 and 2004.

Figure 8 depicts that the *zila* level poverty map based on Census 2001 is similar to that of PSC 2004. In fact, both the correlation coefficient and the rank correlation are nearly 90 percent. This similarity between the two poverty maps suggests that the long interval between Census 2001 and HIES 2005 is unlikely to cause a large bias in the updated poverty estimates.

The PSC 2004 data also suggest that the impact of migration on poverty estimates is likely to be small. According to PSC 2004, recent migrants constitute less than 5 percent of total population in all but three *zilas* - Dhaka (11.4%), Narayanganj (9.2%), and Gazipur (8.4%). Even in these three zilas, poverty rates between recent migrants and others are similar; so the impact of recent migration on poverty incidence is limited.

In conclusion, the validation work described above confirmed that *neither* the time lag between the Census 2001 data and the HIES 2005 data, *nor* migration activities during the same period, were likely to have a big impact on updated poverty estimates and data quality.

C. Comparison between the Perception Map and the Poverty Map

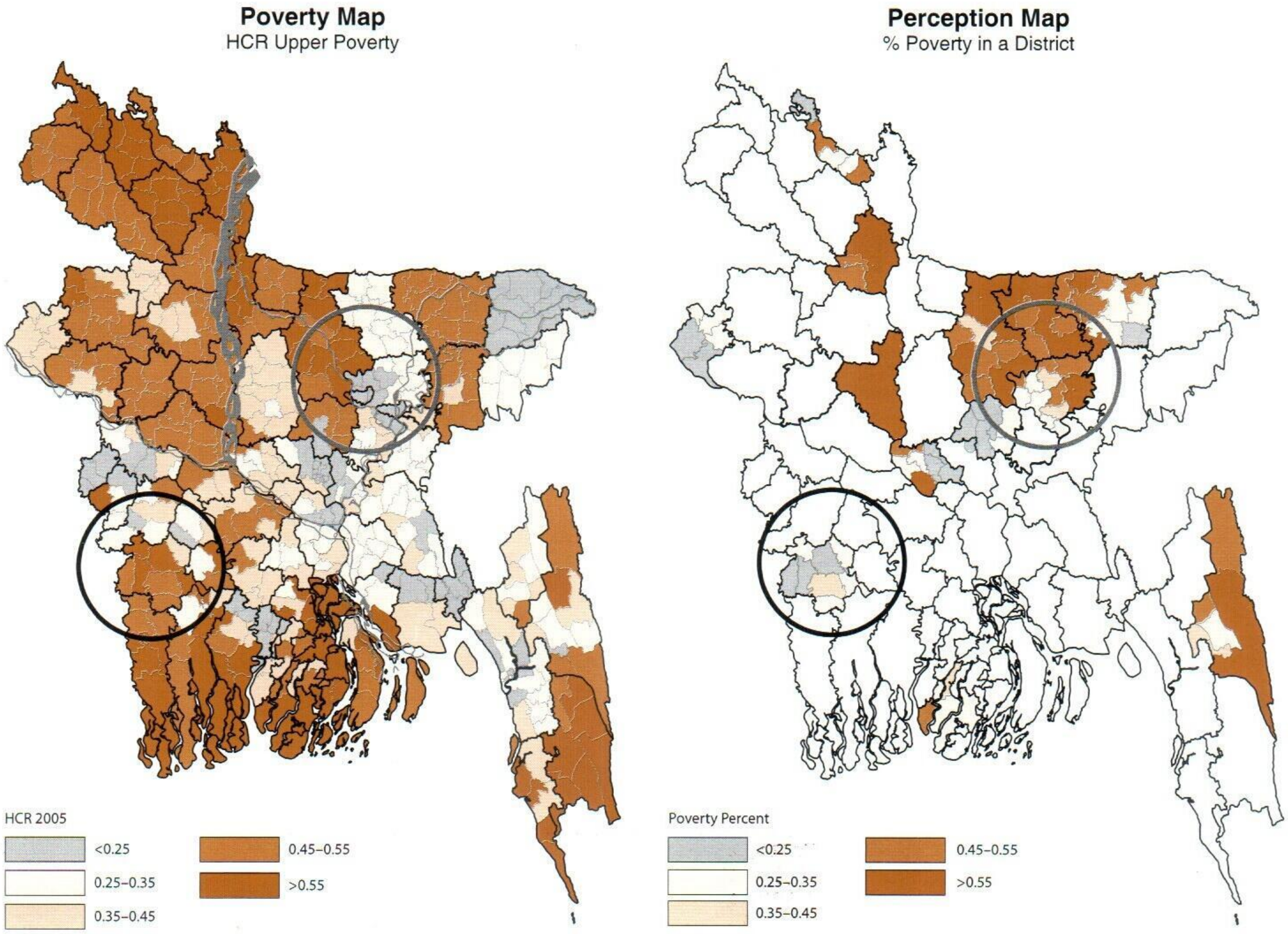
Estimates of poverty prevalence based on perceptions often differ significantly from estimates based on more statistical or objective methods, such as the SAE based method described earlier. Nevertheless, perception can play an important role on resource allocation decisions. Also, perceptions sometimes help to identify data problems in poverty mapping. As part of the validation and cross checking work associated with the Poverty Mapping Update exercise, WFP conducted a perceptions survey for selected areas during 2008, to see how perception differs from SAE predicted poverty estimates.[7] A preliminary version of the updated poverty map, and consultations with key informants including BBS staff, World Bank staff, WFP staff, and technical committee members, were used to focus and prioritize areas where perceptions data was collected. The results of this perception survey were used to refine consumption models before finalizing the poverty maps of 2005. Nevertheless, the perceptions survey suggests that even after refinements of the consumption models, there still exists some areas where perception differ significantly from the poverty incidence estimated by the SAE method. Both Jessore and Netrakona districts, highlighted with circles in Figure 9 correspond to two such areas.

Before discussing the divergence between the perception and poverty maps, it is worth noting that there are some differences in the definition of poverty. First, in the perceptions survey, poverty means inability to meet basic needs in terms of: (1) food and clothing consumption, (2) housing conditions, and (3) access to clean water, health services, and schools. On the other hand, the poverty map is produced based on consumption poverty, which BBS also adopted in producing the official poverty estimates of 2005. Although the consumption poverty is often highly correlated with the multiple dimensions of basic needs, there are also some exceptional cases. Second, the perceptions survey focused on "chronic" poverty while the poverty map represents a snapshot of poverty in 2005. Again, in theory, if households have full access to credit markets, the level of consumption should be constant over time ("consumption smoothing") but, in reality, it is quite volatile.

Considering these differences, the divergence between the perception estimates and SAE estimates are highlighted in Figure 9. It shows that Jessore district in the southwest of the

7 See Hassan and Hassan (2008) for further details.

FIGURE 9: Comparison between the poverty map and the perceptions map



Source: The perception map was created by the perception survey conducted by the WFP in 2008. The details in the methodology are available in Hassan and Hassan (2008).

country (circled in black) is a good example of a case where poverty estimates based on perceptions (see map on the right) differed significantly from estimates based on the SAE approach used for the updated Poverty map (see map on the left). It may be noted that Jessore has been relatively less affected by natural disasters in recent years, and this could contribute to economic growth and/or the perception of lower poverty as reflected in the map. Netrakona district, on the other hand (circled in grey), was badly affected by the very large and relatively recent floods of 2007, and this could have contributed to either significant economic losses or to the perception of worsening poverty amongst key informants. Even though the perceptions survey clearly asked interviewees to provide their perception of poverty in 2005, this perception is likely to be affected by the aftermath of the 2007 flood.

In conclusion, there are significant evidence of differences between perceptions and results from the SAE method. Although, a further careful assessment is necessary before accepting the differences, they seem to be attributed to the aforementioned differences in concept of 'poverty' and also 'timing'. However, acknowledging the differences is critical when using poverty maps based on the SAE method. In areas where perceptions differ substantially from the statistical results, it is likely to face more resistance against poverty ranking based on the SAE based poverty maps.

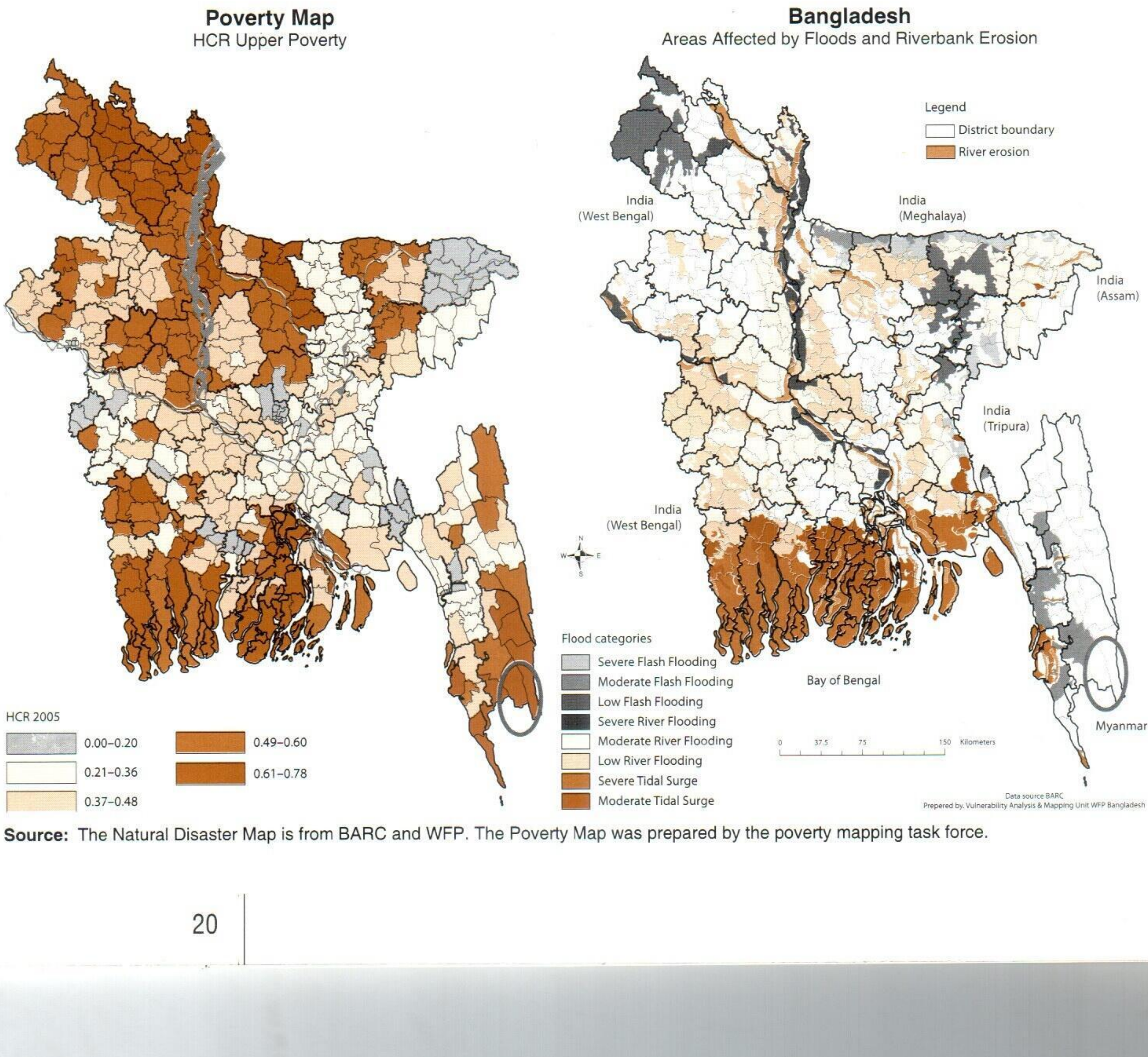
D. Consistency between the Poverty Map and Other Regional Characteristics

Another way to evaluate the reliability of poverty maps is to compare them with other geographic or regional characteristics that are likely correlated with poverty incidence. For example, natural disasters like floods, droughts, and cyclones are known to affect people's livelihoods in Bangladesh. Therefore, such disasters are important to be related with the poverty maps. Below we check the reliability of the poverty maps by comparing them with maps of poverty related variables such as natural disasters as well as to access to infrastructure and educational attainment of household heads.

Figure 10 compares the poverty map (based on upper poverty lines) with a hazards map related to floods, cyclones, and other natural disasters. The map on the right suggests that many areas with severe river floods are indeed very poor. Furthermore, most areas affected by cyclones and severe tidal surges (colored dark brown, in the south), also suffer from abject poverty. However, it is also clear that proneness to natural disasters is not a sole determinant of regional poverty. For example, Bandarban District (circled in grey) records a very high poverty headcount rate but it is not very prone to natural disasters such as floods and cyclones.

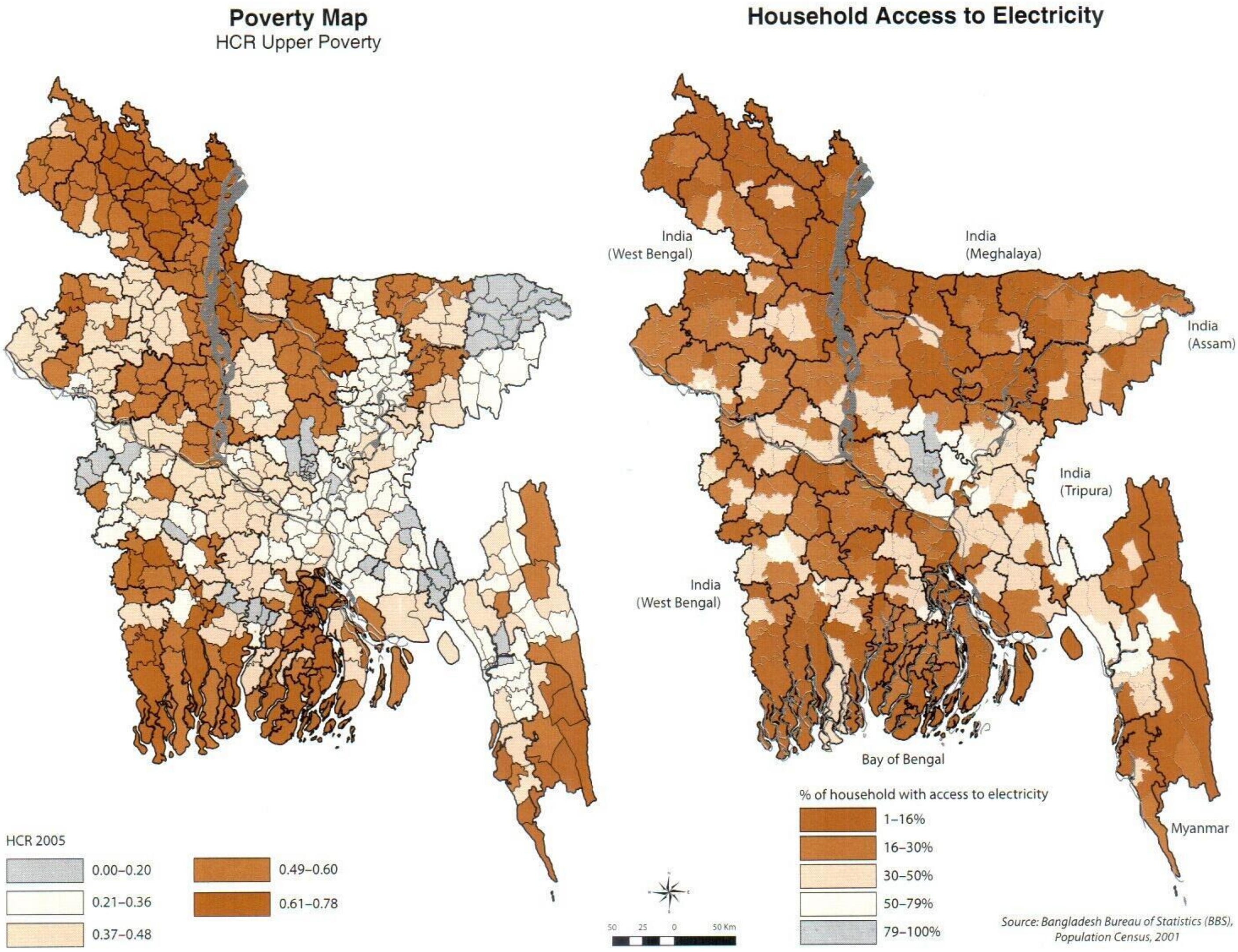
Figure 11 indicates that poverty appears to be closely associated with access to electricity. Darker brown areas on the map correspond with lower access to electricity,

FIGURE 10: The poverty map and the natural disaster map



Source: The Natural Disaster Map is from BARC and WFP. The Poverty Map was prepared by the poverty mapping task force.

FIGURE 11: The poverty map with access to electricity



Source: The map of household access to electricity is prepared by WFP from Population Census 2001.

while areas in lighter brown indicate greater accessibility.

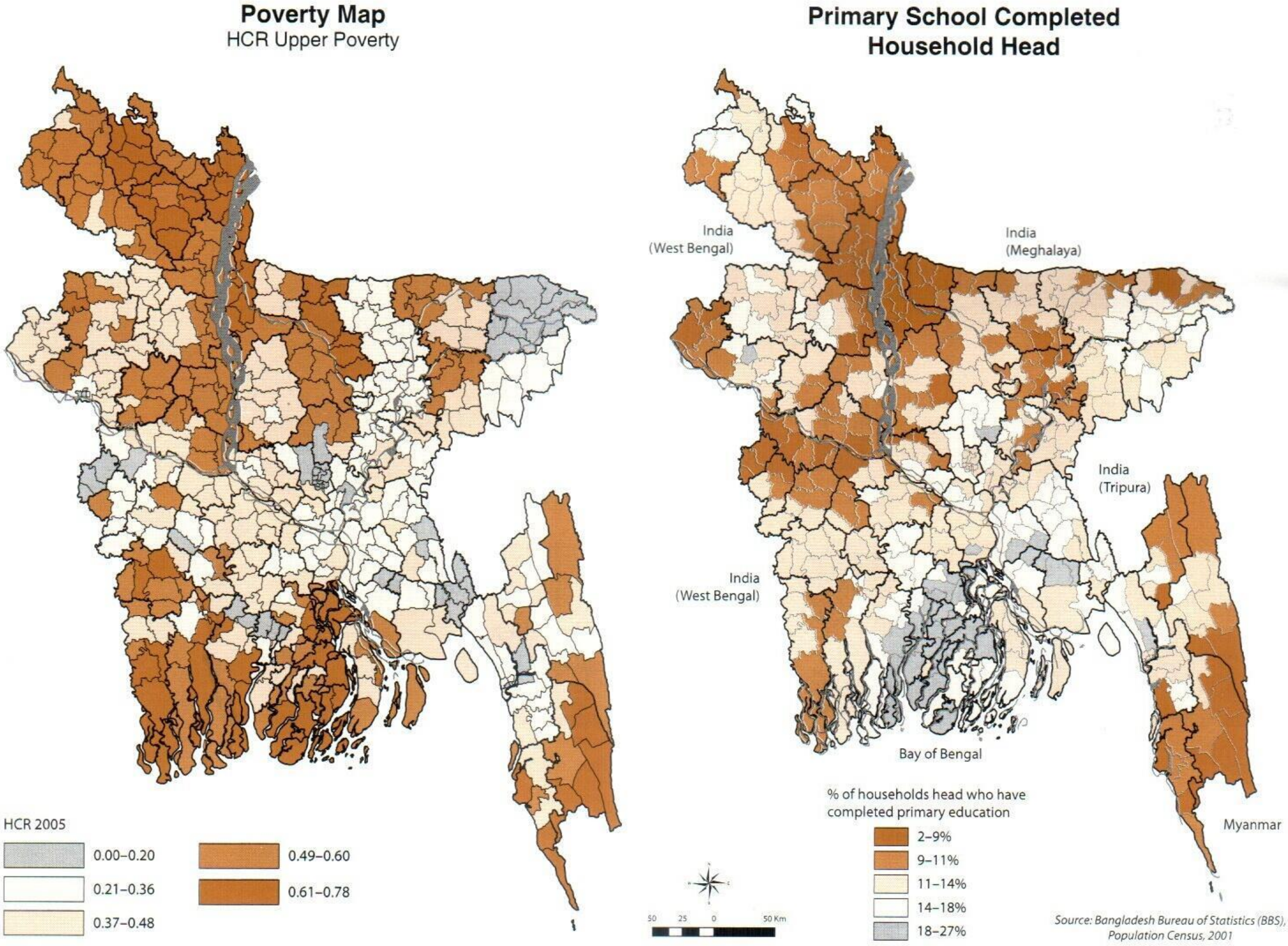
It is observed that northwestern Monga areas and southern coastal areas whose poverty rates are high also suffer from limited access to electricity. A similar situation of generally high poverty rates, and low electricity access rates, is seen in sections of the Chittagong Hill Tribal region and in the southeast. In contrast, the region between and around Bangladesh's two largest cities, Dhaka and Chittagong, is characterized both by low poverty rates and by high rates of access to electricity.

However, access to electricity cannot explain all variations in poverty. For example, Sylhet district, located in the northeast, is among the richest districts in the country although some areas in the district has limited access to electricity.

Poverty and educational attainment also appear to be spatially correlated (see Figure 12). In general, areas with a higher proportion of household heads with completed primary education, poverty rates are lower. For example, the completion rate of primary education is relatively high in Dhaka city and the surrounding areas, where poverty rates are among the lowest areas in Bangladesh. The completion rate of primary education is also low in northwest Monga areas, where the poverty rates are among the highest in the country. On the contrary, there are some areas where almost inverse correlations are exhibited. A number of southern coastal districts have relatively higher proportion of household heads with completed primary education, while these same areas have very high poverty rates.

In conclusion, the poverty map generally exhibits signs of spatial correlations with key regional characteristics as expected. However, these observations also suggest that no single regional characteristic can completely explain variations in poverty.

FIGURE 12: The poverty map and educational attainment of household heads



Source: The map of the share of household head completing primary school education is prepared by WFP from Population Census 2001.

5 Conclusions and Next Steps

This report summarizes the findings from the Bangladesh poverty map update conducted by the BBS, the World Bank, and WFP. In Bangladesh, poverty maps are not new. At least two sets of poverty maps were produced using HIES 2000, Census 2001, and other auxiliary data. Updating the poverty maps is however, new in that (i) it uses the latest available household survey data, HIES 2005, (ii) it incorporates many methodological improvements since the previous poverty maps, and (iii) it carries out a wide variety of validation exercises to ensure the quality of results.

Among all, the main technical challenge in updating the maps is the long interval between Census 2001 and HIES 2005. Consumption patterns and population distribution might have changed over four years, which could have created a large bias in both poverty estimates and their standard errors. Consequently, to minimize bias, a number of adjustments were made, including election of time-invariant household and individual characteristics and the use of location specific variables.

Careful validation exercises are conducted and the results confirm that the bias due to the long interval is limited. Among these validations, producing another set of poverty maps using HIES 2005 and PSC 2004 is a particularly noteworthy exercise. The result of the validation exercise is encouraging since the poverty maps based on HIES 2005 and PSC 2004, which have only one year interval, are approximately the same as those based on HIES 2005 and Census 2001.

Poverty mapping is a powerful instrument to see variations in poverty with great accuracy. However, it should be pointed out that the usefulness of this type of exercise becomes limited when it is not undertaken regularly. The success of this exercise suggests that, even between inter-census years, high quality poverty maps can be produced by focusing on time-invariant variables and using many location specific variables. Capacity building is also important in ensuring the sustainability of this exercise. The World Bank has provided a training workshop at the BBS in June 2008. This type of training should be repeated to facilitate the building of capacity at the BBS and other relevant government agencies.

Annex 1: Additional Tables

TABLE A-1: Shares of variance of errors in each layer (%)

Stratum	Three layers model					Two layers model				One Layer model		
	Up	Un	Mz	HH	All	Up	Mz	HH	All	Mz	HH	All
Barisal (Rural)	Convergence not achieved					Convergence not achieved				Convergence not achieved		
Barisal (Muni.)						0	3	97	100	3	97	100
Chittagong (Rural)						2	2	95	100	5	95	100
Chittagong (Muni.)						3	3	94	100	5	95	100
Chittagong (SMA)						3	3	94	100	5	95	100
Dhaka (Rural)						Convergence not achieved				Convergence not achieved		
Dhaka (Muni.)	0	1	1	98	100	0	2	98	100	2	98	100
Dhaka (SMA)	0	2	2	95	100	0	5	95	100	5	95	100
Khulna (Rural)	0	2	2	96	100	0	4	96	100	4	96	100
Khulna (Muni.)	2	3	3	92	100	0	4	96	100	4	96	100
Khulna (SMA)	1	1	1	96	100	2	1	97	100	0	100	100
Rajshahi (Rural)	2	2	2	95	100	2	3	95	100	5	95	100
Rajshahi (Muni.)	2	2	2	94	100	2	2	96	100	3	97	100
Rajshahi (SMA)	4	1	1	94	100	4	2	94	100	6	94	100
Sylhet (Rural)	1	1	1	97	100	1	1	97	100	3	97	100
Sylhet (Muni.)	0	3	3	94	100	0	6	94	100	3	50	100

Source: The World Bank staff estimation using STATA's program XTMIXED.

Note: Household heterogenity is not included.

TABLE A-2: List of time invariant variables

	Barisal_Pur	Barisal_Urb	Chittagong_ Rur	Chittagong_ Urb	Chittagong_ SMA	Dhaka_Rur	Dhaka_Urb	Dhaka_SMA
	Str1	Str2	Str3	Str4	Str5	Str6	Str7	Str8
1	chid0yrp	chid0yrp	chid0yrp	chid0yrp	chid0yrp	chid0yrp	chid0yrp	Chid0yrp
2	dhd_single	n60plusp	n60plusp	chid1_4p	chid1_4p	dhd_maried	chid1_4p	dhdwid_div
3	dhd_maried	dhd_maried	dhd_maried	chid5_14p	n60plusp	dhdnmuslim	n60plusp	dhdothswrk
4	dhdwid_div	dhdwid_div	dhdwid_div	dhd_maried	dhdwid_div	dhd_wrk	dhd_maried	dsemipucca
5	dhd_lit	dhd_lit	dhd_lit	dhdnmuslim	dhdnmuslim	dhdothswrk	dhdwid_div	dpucca
6	dhdothswrk	dhd_wrk	dhdothswrk	dhd_lit	dhd_wrk	dpucca	dhd_lit	downed_hh
7	dtap_water	dhdothswrk	dpucca	dhd_wrk	dhdothswrk	downed_hh	dhd_wrk	drentfree
8	dsemipucca	dsemipucca	downed_hh	dhdothswrk	downed_hh	drentfree	dhdothswrk	djsec_edu
9	djsec_edu	drentfree	drentfree	drented_hh	drented_hh	dhsec_edu	dtubewater	dhsec_edu
10	dhsec_edu	dsec_edu	dhsec_edu	drentfree	djsec_edu	dgra_edu	downed_hh	dvoc_edu
11	dvoc_edu	dhsec_edu	dvoc_edu	djsec_edu	dsec_edu	dpgra_edu	drented_hh	dgra_edu
12	dgra_edu	dgra_edu	dgra_edu	dsec_edu			drentfree	dpgra_edu
13		dpgra_edu	dpgra_dud	dhsec_edu			dno_edu	
14		chma_femp		dpgra_edu			djsec_edu	
15				chma_femp			dhsec_edu	
16							dgra_edu	
17							dpgra_edu	
18								
19								
20								
21								

TABLE A-2: List of time invariant variables

	Khulna_Rur	Khulna_Urb	Khulna_ SMA	Rajshahi_ Rur	Rajshahi_ Urb	Rajshahi_ SM	Sylhet_Rur	Sylhet_Urb
	Str9	Str10	Str11	Str12	Str13	Str14	Str15	Str16
1	chld0yrp	chld0yrp	chld0yrp	dhd_single	chld0yrp	chld0yrp	chld0yrp	chd0yrp
2	dhd_single	chld1_4p	child1_4P	dhdnmuslim	dhd_single	n60plusp	n60plusp	child1_4p
3	dhd_maried	chld5_14p	n60plusp	dhd_lit	dhd_maried	dhd_maried	dhd_single	n60plusp
4	dhdnmuslim	n60plusp	dhd_single	djsec_edu	dhdwid_div	dhdwid_div	dhd_maried	dhd_single
5	dhd_lit	dhd_maried	dhd_maried	dsec_edu	dhd_lit	dhd_lit	dhdwid_div	dhd_maried
6	dhd_wrk	dhdwid_div	dhdwid_div	dhsec_edu	delectric	dhd_wrk	dhd_lit	dhdwid_div
7	dhdothswrk	dhdnmuslim	dhd_lit	*chld0yrp**	dnolatrine	dtap_water	dhdothswrk	dhdothswrk
8	downed_hh	dhd_lit	dhd_wrk	*chld5_14p**	djsec_edu	dtubewater	dsemipucca	dsemipucca
9	djsec_edu	dhd_wrk	dhdothswrk	*n60plusp**	dsec_edu	dpucca	downed_hh	djsec_edu
10	dsec_edu	dhdothswrk	dtap_water	*tmem**	dgra_edu	dslatrine	drentfree	dsec_edu
11	dhsec_edu	dtap_water	dtubewater	*malep**	dpgra_edu	downed_hh	dno_edu	dhsec_edu
12	dgra_edu	dtubewater	dsemipucca			drented_hh	djsec_edu	
13	dpgra_edu	dpondwater	dpucca			dno_edu	dhsec_edu	
14		dsemipucca	delectric			dpri_edu	dgra_edu	
15		downed_hh	dslatrine			djsec_edu		
16		dsec_edu	downed_hh			dsec_edu		
17		dhsec_edu	drented_hh			dhsec_edu		
18		dvoc_edu	dsec_edu			dvoc_edu		
19		dgra_edu	dhsec_edu			dpgra_edu		
20		dpgra_edu	dgra_edu					
21		chma_femp						

** Selection of variables based on HIES 2005 and Census 2001 comparability.*

TABLE A-2: Final models with the OLS coefficients

Variables	Description of the variables	Estimated coefficient
	Barisal Rural [STRATUM 1]	
intercept	Constant used in the model	11.28
CHLD1_4P_MEAN_U	Average proportion of child aged 1-4 yr at upazila level	19.83
CHLD1_4P_MEAN_UN	Average proportion of child aged 1-4 yr at union level	-19.85
DGRA_EDU_1	Graduate head in the household	0.38
DHDNMUSLIM_MEAN_U	Mean of non-muslim head at upazila level	1.60
DHD_LIT_1	Literate head in the household	0.29
DHD_MARIED_1	Married head in the household	-0.16
DHSEC_EDU_1	Head with higher secondary education in the household	0.25
DT_06	Dummy for district Barisal	-0.60
N15_59YRP_MEAN_UN	Average proportion of persons aged 15-59 yr at union level	-10.35
N60PLUSP_MEAN_U	Average proportion of persons aged 60+ yr at upazila level	13.77
TMEM2	Household size squared	0.00
_DT$DHD_LIT_780	District=78 and head not literate	-0.39
_DT$DSEMIPUCCA_420	District=42 and not a semipucca house	-0.40
	Barisal Urban [STRATUM 2]	
intercept	Constant used in the model	6.62
DGRA_EDU_1	Graduate head in the household	0.50
DHDWID_DIV_1	Head of the household is widowed or divorced	0.44
DHSEC_EDU_1	Head of the household with higher secondary education	0.47
DPGRA_EDU_1	Head of the household with post graduate education	0.71
DSEC_EDU_1	Head of the household with secondary education	0.37
TPE06_SRVC	Total persons engaged in the service sector at upazila level, 2006	0.00
_DHD_WRK$DHD_LIT_01	Literate head of the household not working	0.28
_DHD_WRK$DHD_LIT_11	Literate head of the household and working	0.30
_DT$DHDOTHSWRK_090	District=09 and both head & others in the household working	0.58
_DT$DSEC_EDU_420	District=42 and head with not secondary education	-0.59
_DT$DSEMIPUCCA_090	District=09 and house is not semipucca	-0.34
	Chittagong Rural [STRATUM 3]	
intercept	Constant used in the model	7.39
CHLD0YRP	Proportion of children aged 0 yr in the household	-0.61
DGRA_EDU_1	Graduate head in the household	0.41
DHDOTHSWRK_1	Both head & others working in the household	0.05
DHDWID_DIV_1	Head of the household is widowed or divorced	0.09
DHD_LIT_1	Head of the household is literate	0.20
DHSEC_EDU_1	Head of the household with higher secondary education	0.28
DNOLATRINE_MEAN_UN	Proportion of household with no latrine at union level	-0.43
DNO_EDU_MEAN_UN	Proportion of household with head not working at union level	-0.39
DPONDWATER_MEAN_U	Proportion of household using pond water for drinking at union level	-1.24

Variables	Description of the variables	Estimated coefficient
DPUCCA_1	Pucca house	0.54
DRENTFREE_1	House is rent free	-0.18
N60PLUSP	Proportion of elderly people (60+) in the household	0.33
TMEM2	Household size squared	-0.01
TMEM3	Household size cubed	0.00
TPE03_CONS	Total persons engaged in the construction sector in 2003	0.00
_DT#015$N60PLUSP	District=15 & Proportion of Elderly people in the household	0.45
_DT#030$N60PLUSP	District=30 & Proportion of Elderly people in the household	-1.07
_DT$DPGRA_EDU_130	District=13 and head with not post graduate education	0.12
_DT$DPUCCA_300	District=30 and not a pucca house	0.27
Chittagong Urban [STRATUM 4]		
intercept	Constant used in the model	7.39
CHLD0YRP	Proportion of the children aged 0 yr in the household	-0.59
CHLD1_4P	Proportion of the children aged 1-4 yr in the household	-0.91
CHLD5_14P	Proportion of the children aged 5-14 yr in the household	-0.77
DHDNMUSLIM_1	Non muslim head in the household	-0.15
DHDNMUSLIM_MEAN_UN	Mean of non-muslim head at union level	0.41
DHD_LIT_1	Literate head in the household	0.44
DJSEC_EDU_1	Head with junior secondary education in the household	-0.19
DRENTFREE_1	Household with rent free house	-0.29
DT_30	Dummy for district=30	0.93
DT_46	Dummy for district=46	-0.34
TMEM2	Household size squared	0.00
_DT$DHSEC_EDU_030	Dummy for district=03 and head with not higher secondary education	-0.32
_DT$DHSEC_EDU_130	Dummy for district=13 and head with not higher secondary education	-0.35
_DT$DHSEC_EDU_300	Dummy for district=30 and head with not higher secondary education	-0.82
_DT$DPGRA_EDU_190	Dummy for district=19 and head with not post graduate education	-0.18
Chittagong SMA [STRATUM 5]		
intercept	Constant used in the model	7.38
CHLD1_4P	Proportion of children aged 1-4 yr	-1.14
DHDNMUSLIM_1	Head of the household is a non-muslim	-0.23
TPE06_CONS	Total persons engaged in the construction sector for 2006	0.01
Dhaka Rural [STRATUM 6]		
intercept	Constant used in the model	6.59
CHLD0YRP	Proportion of children aged 0 yr	-1.11
CHLD1_4P	Proportion of children aged 1-4 yr	-1.00
CHLD5_14P	Proportion of children aged 5-14 yr	-0.57
DAGR_WKER_MEAN_UN	Mean of head working in Agri. Sector at union level	0.49
DBUSS_WKER_MEAN_M	Mean of head working in Bussiness Sector at mauza level	0.88

Variables	Description of the variables	Estimated coefficient
DBUSS_WKER_MEAN_U	Mean of head working in Bussiness Sector at upazila level	1.21
DELECTRIC_1	Household with access to electricity	0.32
DGRA_EDU_1	Household with graduate head	0.31
DHD_LIT_1	Household with literate head	0.24
DHSEC_EDU_1	Head with higher secondary education	0.37
DSEC_EDU_1	Head with secondary education	0.11
DT_48	Dummy for district=48	0.26
DT_72	Dummy for district=72	0.20
TMEM2	Household size squared	0.00
_DT$DHSEC_EDU_330	District=33 and Head's education not higher secondary	-0.23
_DT$DHSEC_EDU_390	District=39 and Head's education not higher secondary	-0.15
_DT$DHSEC_EDU_541	District=54 and Head's education is higher secondary	-0.95
_DT$DHSEC_EDU_610	District=61 and Head's education not higher secondary	-0.14
_DT$DHSEC_EDU_860	District=86 and Head's education not higher secondary	0.19
Dhaka Rural [STRATUM 7]		
intercept	Constant used in the model	7.66
CHLD0YRP	Proportion of children aged 0 yr	-0.72
CHLD1_4P	Proportion of children aged 1-4 yr	-0.71
DGRA_EDU_1	Household with graduate head	0.39
DHD_WRK_1	Household with working head	-0.17
DHSEC_EDU_1	Head with higher secondary education	0.24
DNO_EDU_1	Head with no education	-0.45
DPGRA_EDU_1	Household with post graduate head	0.43
DRENTFREE_1	Rent free household	-0.22
DT_26	Dummy for district=26	0.33
DT_48	Dummy for district=48	0.52
DT_54	Dummy for district=54	0.17
DT_59	Dummy for district=59	0.71
DT_86	Dummy for district=86	0.33
DTUBEWATER_1	Household with Tubewell	-0.20
N60PLUSP	Proportion of Elderly people in the household	0.30
TMEM2	Household size squared	0.00
TPE06_SALE	Total persons engaged in wholesale & retail sector	0.00
Dhaka SMA [STRATUM 8]		
intercept	Constant used in the model	8.58
CHLD1_4P_MEAN_U	Average proportion of children aged 1-4 yr at upazila level	-9.06
CHLD1_4P_MEAN_UN	Average proportion of children aged 1-4 yr at union level	-6.92
DELECTRIC_MEAN_U	Average proportion of household with access to electricity at upazila level	0.64
DELECTRIC_MEAN_UN	Average proportion of household with access to electricity at union level	-0.70
DOWNED_HH_1	Proportion of household owned a house	0.26
DPGRA_EDU_1	Proportion of household with post graduate head	0.17
DPUCCA_1	Proportion of household with pucca house	0.63

Variables	Description of the variables	Estimated coefficient
DSEMIPUCCA_1	Proportion of household with semi pucca house	0.26
_TMEM2$DGRA_EDU#0	Interaction of household size squared and head not graduate	0.00
	Khulna Rural [STRATUM 9]	
intercept	Constant term used in the model	6.65
CHLD0YRP	Proportion of children aged 0 yr in the household	-0.71
DAGR_WKER_MEAN_U	Upazila level census mean of head working in Agriculture sector	1.41
DBUSS_WKER_MEAN_UN	Union level census mean of head working in Bussiness sector	1.18
DELECTRIC_MEAN_M	Mauza level census mean of household with access to electricity	0.59
DGRA_EDU_1	Household with graduate head	0.43
DHDNMUSLIM_MEAN_M	Mauza level census mean of household with head non-muslim	-0.28
DHDNMUSLIM_MEAN_U	Upazila level census mean of household with head non-muslim	0.78
DHD_LIT_1	Head literate in the household	0.18
DHSEC_EDU_1	Head with higher secondary education in the household	0.24
DPGRA_EDU_1	Head with post graduate education in the household	0.30
DSEC_EDU_1	Head with secondary education in the household	0.12
DT_87	Dummy for district=87	-0.29
N60PLUSP_MEAN_M	Mauza level census mean of elderly people in the household	-5.57
N60PLUSP_MEAN_U	Upazila level census mean of elderly people in the household	-24.84
N60PLUSP_MEAN_UN	Union level census mean of elderly people in the household	16.27
SEV_DROU	Proportion of total area subject to severe drought	-0.49
TPE06_CONS	Total persons engqazed in construction sector at upazila level	0.00
_DT$DGRA_EDU_010	Dummy for district=01 & head not graduate	0.25
_DT$DGRA_EDU_410	Dummy for district=41 & head not graduate	-0.35
_DT$DGRA_EDU_551	Dummy for district=55 & head graduate	0.64
_DT$DHSEC_EDU_011	Dummy for district=01 & head with higher secondary education	1.46
_DT$DHSEC_EDU_440	Dummy for district=44 & head not higher secondary education	-0.15
_DT$DOWNED_HH_501	Dummy for district=50 & owned a house	0.19
_DT$DOWNED_HH_571	Dummy for district=57 & owned a house	0.31
	Khulna Urban [STRATUM 10]	
intercept	Constant used in the model	6.86
ANYCHAR_1	Dummy for char area	0.72
CHLD1_4P	Proportion of children aged 0-4 yr in the household	-0.78
CHLD5_14P	Proportion of children aged 5-14 yr in the household	-0.58
DBUSS_WKER_MEAN_M	Mauza level census mean of head working in bussiness sector	1.05
DELECTRIC_MEAN_U	Upazila level census mean of household with access to electricity	-0.39

Variables	Description of the variables	Estimated coefficient
DGRA_EDU_1	Household with graduate head	0.58
DHDOTHSWRK_1	Household with both head & others working	-0.19
DHD_LIT_1	Household with literate head	0.22
DHSEC_EDU_1	Household with head completed higher secondary education	0.66
DOWNED_HH_1	Household owned a house	0.18
DPGRA_EDU_1	Household with post graduate head\	0.75
DSEC_EDU_1	Household with head completed secondary education	0.24
DT_01	Dummy for district=01	-0.39
DT_57	Dummy for district=57	0.20
DVOC_EDU_1	Head with vocational education	0.99
N60PLUSP	Proportion of household with elderly people	-0.65
SEV_DROUGH	Proportion of total area subject to severe drought	-7.53
Khulna SMA [STRATUM 11]		
intercept	Constant used in the model	6.63
CHLD1_4P	Proportion of children aged 1-4 yr in the household	-0.55
DGRA_EDU_1	Household with graduate head	1.17
DSEC_EDU_1	Household with head completed secondary education	0.27
_DHD_MARIED$DHD_LIT_01	Dummy for head not married but literate	0.40
_DPUCCA$DELECTRIC_11	Dummy for pucca house with electricity\	0.50
_TMEM2$DGRA_EDU#1	Interaction of household size squared & head graduate	-0.03
Rajshahi Rural [STRATUM 12]		
intercept	Constant used in the model	6.58
CHLD0YRP	Proportion of children aged 0 yr in the household	-1.31
CHLD1_4P	Proportion of children aged 1-4 yr in the household	-1.02
CHLD5_14P	Proportion of children aged 5-14 yr in the household	-0.59
DHDNMUSLIM_1	Household with head non-muslim	-0.08
DHD_LIT_1	Household with head literate	0.21
DHSEC_EDU_1	Household with head completed higher secondary education	0.40
DSEC_EDU_1	Household with head completed secondary education	0.18
DTUBEWATER_MEAN_M	Mauza level census mean of household with access to tube-wel	0.30
DT_49	Dummy for district=49	-0.13
DT_70	Dummy for district=70	0.14
DT_73	Dummy for district=73	-0.13
DT_85	Dummy for district=85	-0.10
Rajshahi Urban [STRATUM 13]		
intercept	Constant used in the model	6.64
ANYCHAR_1	Upazila having char area	0.19
DELECTRIC_1	Household with access to electricity	0.35
DGRA_EDU_1	Household with head graduate	0.48
DHD_LIT_1	Household with literate graduate	0.22
DHD_MARIED_1	Household with head married	-0.12

Variables	Description of the variables	Estimated coefficient
DNOLATRINE_1	Household with no latrine	-0.15
DPGRA_EDU_1	Household with head post graduate	0.48
DSEC_EDU_1	Household with head completed secondary education	0.13
DT_52	Dummy for district=52	-0.14
DT_69	Dummy for district=69	-0.24
TPE06_MANU	Total person engaged in the mfg. sector at upazila level	0.00
_DT#049$TMEM2	Interacton of district=49 & household size squared	-0.01
Rajshahi SMA [STRATUM 14]		
intercept	Constant used in the model	6.79
CHLD0YRP	Proportion of the children aged 0 yr in the household	-1.32
DHD_LIT_1	Household with literate head	0.61
DPGRA_EDU_1	Household with post graduate head	0.98
DPUCCA_1	Household with pucca house	0.40
N60PLUSP	Proportion of the children elderly people in the household	0.53
_TMEM2$DHD_LIT#1	Interaction of household size squared & head literate	-0.01
Sylhet Rural [STRATUM 15]		
intercept	constant used in the model	6.88
DHD_LIT_1	Household with head literate	0.23
DHSEC_EDU_1	Household with head completed higher secondary education	0.90
DSEC_EDU_1	Household with head completed secondary education	0.22
DT_58	Dummy for district=58	0.20
DT_91	Dummy for district=91	0.45
N60PLUSP	Proportion of elderly people in the household	0.47
TMEM2	Household size squared	0.00
DGRA EDU$DSEMIPUCCA_00	Dummy for not graduate & not semipucca house	-0.18
_DT$DSEMIPUCCA_901	Dummy for district=90 & semipucca house	1.00
Sylhet Urban [STRATUM 16]		
intercept	Constant used in the model	7.68
CHLD0YRP	Proportion of the children aged 0 yr in the household	-1.98
CHLD1_4P	Proportion of the children aged 1-4 yr in the household	-0.64
DAGR_WKER_MEAN_U	Upazila level census mean of head working in Agriculture Sector	-1.85
DPUCCA_1	Household is a pucca house	0.54
DSEMIPUCCA_1	Household is a semi-pucca house	0.53
TMEM2	Household size squared	0.00
_DT$DOWNAGLND_580	Dummy for district=58 & not owning agri. Land	-0.45
_DT$DOWNAGLND_901	Dummy for district=90 & owning agri. Land	0.45
_DT$DPGRA_EDU_910	Dummy for district=91 & head not post graduate	-0.26
_DT$DPUCCA_911	Dummy for district=91 & pucca house	0.66

Annex 2: Descriptions of Main Datasets

Population Census 2001

The population census was conducted by Bangladesh Bureau of Statistics during the period January 23 to January 27, 2001. The enumeration of the census was aimed to take place at a point of time, called the census night (midnight of 22nd January, 2001). Data collected through the population census are cross-classified and analyzed to meet the demand of users and policy makers.

For the convenience of the field operation and to ensure full coverage, the whole country was divided into 262,000 Enumeration Areas (EA's). Each Enumeration Area was formed taking around 100 households. In order to facilitate identification of all EA's within administrative areas, unique geocodes were assigned to each EA, and a map of each EA was prepared and supplied to each enumerator. This helped avoiding omissions and duplications to a great extent in the field work.

The questionnaire used for the population census 2001 contained 28 basic questions. 16 questions were related to housing and household characteristics and 12 questions for individual members. In field operation mainly de-facto method was followed for enumeration. To ensure speedy and timely processing of census data, the questionnaire for the census was designed in OCR/OMR format.

Population Sample Census (PSC) 2004

The Population Sample Census survey was aimed to supplement the information collected in the Population Census 2001. More specifically it collected socio-economic and demographic status of population in the country that the Population Census 2001 did not collect. For this purpose, it is desirable to conduct the PSC immediately after the Population Census; however, for some reasons, the PSC was delayed and conducted only during the period January 8 to January 25, 2004.

Questionnaire of PSC 2004

The PSC 2004 includes the following information:

- Household & housing status
- Household structural composition, relationship, size and pattern
- Obtain data on demographic characteristics by gender and other gender issues
- Population health & health services, water and sanitation
- Income, occupation and other socio-economic activities
- Ownership of assets etc.

Sampling of PSC 2004

The sampling design of the PSC 2004 adopts a two stage stratified cluster design. The whole country was first divided into three basic strata viz.

1. SMA (statistical metropolitan area)
2. Municipal Areas
3. Rural Areas.

SMA stratum was further divided into 4 sub-strata as follows:

i. Dhaka SMA
ii. Chittagong SMA
iii. Khulna SMA
iv. Rajshahi SMA.

Municipal stratum was divided into 58 sub-strata having one for each zila excepting for Dhaka, Gazipur, Narayanganj, Chittagong, Khulna and Rajshahi zilas where their municipalities were merged with the respective SMA's.

Similarly Rural stratum was divided into 64 sub_strata having one for each zila. In total, there were 126 (4+58+64) substrata in the population sample census, 2004. The Enumeration Areas (EA's) of population census 2001, a cluster of around 100 households, were treated as Primary Sampling Unit (PSU). A total of 6000 PSU's (EA's) were allocated to three basic strata. The proportional allocation method was adopted in allocating the PSU's to each basic stratum. Allocation of PSU's to substrata was also done using the proportional allocation method. Thus zila level estimates with urban rural breakdowns are aimed to be representative.

At the first stage, PSU's were selected systematically with a random start within each substratum. Prior to enumeration, a complete household listing operation was done in the selected PSU's. At the second stage, 25 households were selected randomly from each selected PSU. Thus the sample size for the PSC 2004 amounts to 150,000 households.

References

Bigman, D. and U. Deichmann. (2000). 'Spatial indicators of access and fairness for the location of public facilities', in Geographical Targeting for Poverty Alleviation. Methodology and Applications, edited by D. Bigman and H. Fofack, World Bank Regional and Sectoral Studies, Washington DC.

Elbers, C., J.O. Lanjouw, and P. Lanjouw (2003). "Micro-level Estimation of Poverty and Inequality," *Econometrica*, 71(1):355-364.

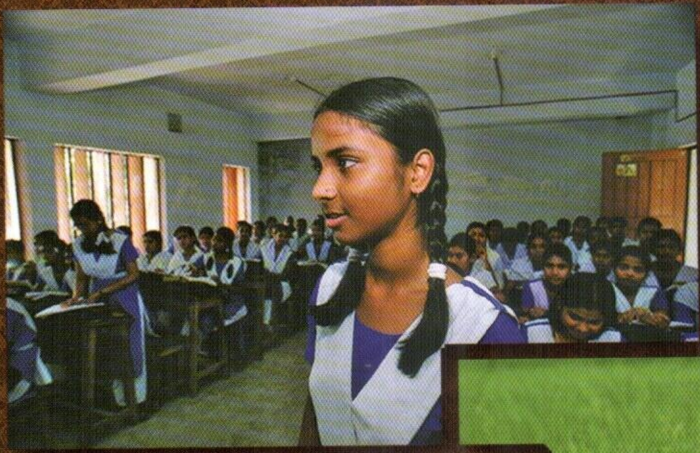
Elbers, C., P. Lanjouw, and P. G. Leite (2008). "Brazil within Brazil: Testing the poverty map methodology in Minas Gerais," *Policy Research Working Paper* No. 4513, The World Bank.

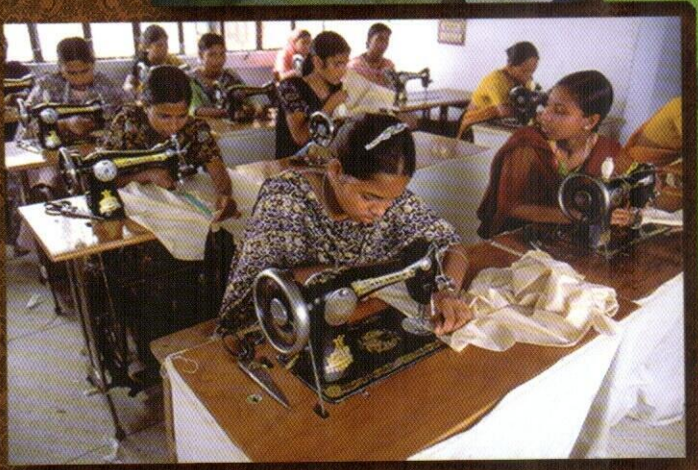
Hassan, M. and M. Hassan (2008). "Field Validation of Poverty Estimates, 2008. Pair wise comparison: A Perception Survey on Relative Prevalence of Poverty," mimeo. Data Analysis and Technical Assistance, Limited.

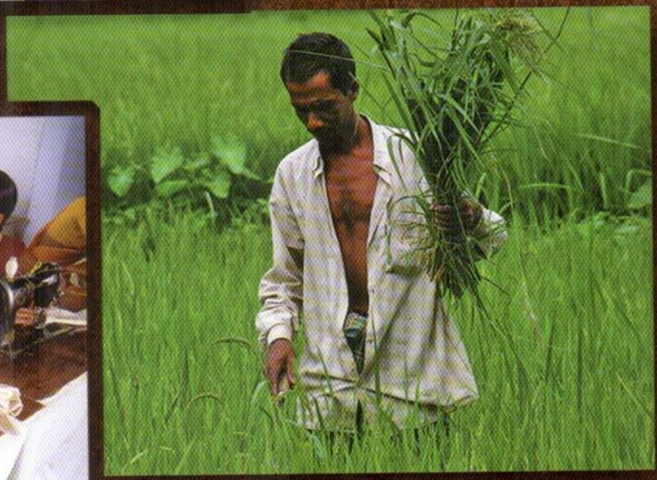
Ravallion, M., and B. Sen (1996). "When method matters: monitoring poverty in Bangladesh," *Economic Development and Cultural Change*, 44(4): 761-92.

Tarrozi, A. and A. Deaton (2008). "Using Census and Survey Data to Estimate Poverty and Inequality for Small Areas," forthcoming in *Review of Economics and Statistics* and available in http://www.princeton.edu/~deaton/downloads/20080301SmallAreas_FINAL.pdf.

World Bank (2008). *Poverty Assessment for Bangladesh – Creating Opportunities and Bridging the East-West Divide*, World Bank, Washington DC.







Bangladesh Bureau of Statistics
Statistics Division
Ministry of Planning
Government of the People's Republic of Bangladesh

The World Bank
1818 H Street N.W.
Washington D.C.
20433 USA

World Food Programme
Via C.G. Viola 68
Parco dei Medici
00148 Rome - Italy