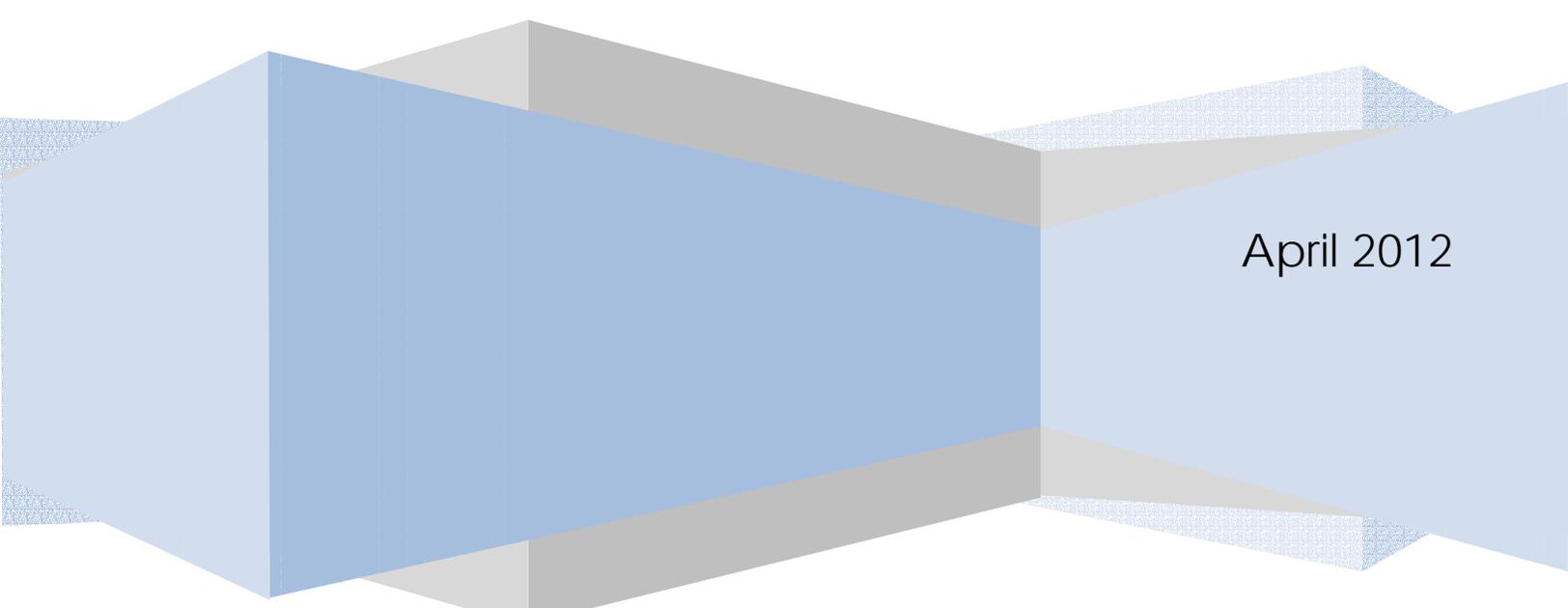# Report of the Post Enumeration Check (PEC) of the Population and Housing Census, 2011

Bangladesh Institute of Development Studies (BIDS)
E-17, Agargaon, Sher-e-Bangla Nagar, Dhaka 1207

April 2012

Contents

FOREWORD

The Bangladesh Bureau of Statistics (BBS), the national data collecting and disseminating agency, is responsible for conducting the decennial population censuses. Since the first population census in 1974, Post Enumeration Check (PEC) has been conducted after every census by BBS. The Population and Housing Census 2011 was conducted by BBS during 15-19 March 2011. Later on, it was decided that the PEC would be conducted by an independent organization and the responsibility of conducting the PEC was given to the Bangladesh Institute of Development Studies (BIDS). It may be mentioned here that this is the first time that the PEC has been planned and implemented by an organization other than BBS since the history of censuses in Bangladesh.

Accordingly, BIDS conducted the PEC in April 2011 to provide information on census coverage and magnitude of content errors. In order to achieve the PEC objective of providing quantitative information on census accuracy, a one-stage stratified cluster design was used in selecting the population for interview. The PEC involved: (i) re-enumerating designated sample segments or Enumeration Areas (EAs), and (ii) re-enumerating the persons in a sample of previously enumerated households. To minimize non-sampling errors, adequate emphasis was placed on extensive training of enumerators/supervisors, proper supervision of fieldwork and use of an effective matching operation.

On behalf of BIDS, I would like to extend my deep appreciation to BBS for making available the necessary financial and technical resources for undertaking the PEC. I would also like to thank the BBS staff who extended all cooperation to BIDS in carrying out this exercise. I have no doubt that if it were not possible to foster the strong partnership between BIDS and BBS, then the PEC would not have been properly organized and implemented.

I extend my sincere appreciation to the Secretary, Statistics and Informatics Division and the Director General of BBS as well as the Project Director of the Population and Housing Census 2011 for their whole hearted cooperation and valuable inputs that made the PEC exercise a success. I would also like to thank the management of BBS, the focal person of BBS, and all others who in one way or another participated in the planning and implementation of the PEC, especially in data processing/matching operations, field supervision and data collection, and above all the individual respondents.

The results indicate a high coverage rate of the Population and Housing Census 2011. This gives confidence in the census data. We hope that the results of PEC will be useful to the government, data analysts and other users of the census data.

Finally, we would like to express our deep gratitude to the Hon'ble Minister, Ministry of Planning, Government of the People's Republic of Bangladesh for his constant guidance and encouragement in conducting the PEC.

Mustafa K. Mujeri
Director General

EXECUTIVE SUMMARY

The Bangladesh Institute of Development Studies (BIDS) conducted the Post Enumeration Check (PEC) in April 2011 to evaluate the coverage and content errors of the Population and Housing Census 2011. Two major domains were selected, namely urban and rural areas. The urban domain was stratified into three different categories: Urban/Municipality, Other Urban, and City Corporation. A total of 280 enumeration areas (EAs) were selected as primary sampling units (PSUs): 140 Rural, 60 Urban, 20 Other Urban, and 60 City Corporation.

Highly experienced enumerators of BBS who worked during the census were retrained to collect data on selected variables. For absolute independence, it was ensured that the enumerators worked in different areas from those covered during the main census.

A matching exercise was undertaken after data collection, which was aimed at investigating whether the PEC persons/households were enumerated during the census. Unmatched records were then reconciled in the field with the purpose of identifying erroneous inclusions. Out of a total of 280 EAs, a sub-sample of 61 EAs --23 rural and 38 urban -- were selected for reconciliation visits. These visits gave an opportunity to identify erroneous census enumerations and the resolution of doubtful/unmatched cases in order to arrive at a realistic match status for each census and PEC element.

Evaluation of Coverage Errors

In evaluating the coverage, a dual system of estimation was used. The PEC results showed that the 2011 Census national coverage rate was around 96 percent giving an omission rate of about 4 percent. At the national level, the PEC has shown a net omission of 41.6 persons for every 1,000 persons enumerated. The figures compare favorably with data from previous PECs. The PEC indicates that the coverage in rural areas was higher than in urban areas. The net omission was 54.3 persons (per 1,000 persons enumerated) in the Municipality compared with 39.8 persons (per 1,000 persons) in the rural area.

The analysis of regional coverage rate indicates moderate differences. The rural region has the highest coverage rate of 96.02 percent. However, urban areas showed a somewhat lower coverage compared with the rural areas. There is no significant difference in coverage between the three urban regions; Municipality area had the lowest coverage rate of 94.56 percent, followed by other urban areas (95.58 percent) and City Corporation (95.94 percent). The net error rate was 3.80 percent for the rural region; while the urban net error rate ranged within 3.86 percent and 5.26 percent.

Evaluation of Content Errors

In order to measure the correctness of responses between the census and the PEC, the rate of agreement, net difference rate and index of inconsistency are used. Sex has the highest rate of agreement of 98 percent and lowest aggregate index of inconsistency of 4 percent. In contrast, age has the lowest rate of agreement of 67 percent. This is because while sex as a characteristic of individuals is easy to report accurately, age depends on the person having accurate idea about the exact age. The rates of agreement of other characteristics are as follows: marital status 97 percent, source of drinking water 89.9 percent, literacy 87.4 percent, and type of toilet facility 69.5 percent.

Implications of the Results and Lessons Learned

The PEC findings are supposed to give credibility to the Population Census 2011 results and would guide the users to better interpret the census results. However, best practices of conducting PEC could not be strictly adhered to in all respects during the process. For instance, due to time and resource constraints the same organization (BBS) that did the census also provided the personnel for the fieldwork, data matching, data analysis and field reconciliation for the PEC. This arrangement partially compromised the independence of PEC from the census, a cardinal assumption behind conducting the PEC.

If the results of PEC are to be credible, the independence of PEC from census should be observed at all stages of PEC implementation; such as fieldwork, data matching, data management and field reconciliation. The lesson here is that independence of PEC from census is vital and should be maintained at all costs. This can be achieved if the PEC unit in BBS is manned by the personnel who do not participate in census activities. Alternatively, an organization outside of BBS can be contracted to plan and implement the PEC, from the beginning to the end. Adequate funds for PEC should be allocated in advance to ensure smooth running of the activities of the survey.

 The lesson learned is that if the budget and manpower for different stages of PEC is not planned, secured and strictly followed, it would cause delays in the implementation of different stages of PEC. In order to successfully implement the PEC by an independent body outside BBS, the outside organization should be well equipped to face the challenges in planning the exercise-- how to deal with the problems of fieldwork after a census, should have trained manpower to implement the matching procedures including data collection and field reconciliation, with adequate manpower and logistic facilities to carry out data processing, tabulation and analysis for write up.

Chapter 1

Background and Objectives

1.1 Background and Objectives

Errors are inevitable in a large data collection exercise such as a census. Errors in census can arise from many sources, such as flawed data collection and processing procedures. For evaluating census results, a number of methods are available which include demographic analysis, record checks and comparison of census data with results of existing household surveys, and post enumeration surveys. The Population and Housing Census (PHC) of Bangladesh was conducted during 15-19 March 2011. For assessing the census outcomes, the decision was to conduct the Post Enumeration Check (PEC) by an independent organization soon after the completion of census enumeration.

It is widely recognized that the independence between the PEC and the PHC is a prime requisite for ensuring the reliability of the dual system of estimation. The validity of the PEC estimates hinges critically on the assumption of independence between the PHC and PEC. In view of the above, the technical responsibility of conducting the PEC was given to the Bangladesh Institute of Development Studies (BIDS), the premier multi-disciplinary autonomous public research organization of Bangladesh.

In order to make the process of conducting the PEC transparent, all census enumeration booklets were transferred from the zonal offices to the central office of the Bangladesh Bureau of Statistics (BBS) at Dhaka prior to the initiation of the PEC operation. Every effort was made by BIDS to maintain the required independence between the two operations (i.e. PHC and PEC) by conducting necessary operational procedures separately. For instance, different enumerators and supervisors were selected and deployed for the PEC. More importantly, the PEC implementation maintained operational independence from the PHC at every stage such as enumeration, data processing, and administering the PEC. Accordingly, all personnel assigned to the PEC had no operational responsibility associated with the main Census.

1.2 What is PEC?

According to the United Nations Principles and Recommendations for Population and Housing Census Rev. 2 (UN 2008), a PEC is a complete re-enumeration of a representative sample of a census population followed by matching each individual enumerated in the PEC with information obtained from the census enumeration. The results of the comparison are mainly used to measure coverage and content error.

Coverage error refers to either an under count or over count of units owing to omissions of persons/ housing units or duplication/erroneous inclusion, respectively. *Content error* pertains to the error in the characteristics that are reported for the persons or housing units that are enumerated. Both types of error can affect the distribution of the population with respect to their characteristics.

In practice, three types of coverage error may exist:

- Omissions;
- Duplications; and
- Erroneous inclusions.

Omissions tend to "undercount" the exact population which results mainly from the failure of census enumerators to record the names of all persons who stayed in the household during the census night. This includes both the omission of persons in the households, and the omission of entire households. Erroneous enumerations of persons, on the other hand, include those who were not residing in the household on the census night or inclusion of persons who died before the census night. Duplications include persons counted more than once or listing names of fictitious persons. Duplications and erroneous inclusions tend to 'over-count' the true population.

These errors can also occur in the recorded characteristics with respect to enumerated persons such as age, family relations, marital status, and other features. While coverage error refers to housing units and people missed in the census or those erroneously included, content errors, on the other hand, evaluate the response quality of selected questions in a census, which are also a basis for evaluating reliability of some characteristics reported in the census (such as age, sex, education and occupation of household members).

1.3 Objectives of the PEC

The goal of PEC exercise is to establish accurately who lived in a particular housing unit on the census night and the next step is to match the results from

the PEC with those of the census questionnaires. Thus, the primary objective of the PEC is to determine sources and magnitude of coverage error and content error.

The specific objectives are to measure: (i) Under-coverage and over-coverage; and (ii) Levels of agreement for responses to questions on selected characteristics, such as sex, age, marital status, education of household members, sources of drinking water, and types of toilet facilities. The main purpose of the PEC for the 2011 PHC was to estimate the 'undercount' and 'over-count' of population on the basis of which estimate of true population would be obtained.

1.4 Scope of the PEC

The intention of the PHC 2011 was to cover every person present in Bangladesh on the census night (except foreign diplomats and their families). However, due to difficulties in enumeration and matching it was not possible to include individuals in prisons, hospitals and other institutions in the PEC. Homeless people were also beyond the scope of the PEC although, as the process was applied at the final stage, they would have received the same adjustment for undercount as people who lived in residential dwellings in urban and rural areas. Thus, for practical reasons, the coverage of the PEC was limited to persons living in households in residential dwellings.

Chapter 2

Methodology of the PEC

2.1 Issues in Methodology

In general, PEC involves: (i) re-enumerating designated sample segments of the enumeration areas (EAs), and (ii) re-enumerating the persons in a sample of previously enumerated households.

2.2 Sample Size

The decision on an appropriate sample size has many facets, such as availability of resources; required precision; type and size of domains; variability of the attributes being measured; and expected level of non-response. For determining the sample size, the most important task was to prepare a sample design so that unbiased estimates of **coverage error** and **content error** can be obtained for selected variables by sex, age-group and rural-urban residence. The following methodology was adopted.

- The total number of PSUs was chosen at 282 for the PEC in 2011. It may be mentioned that the number was 255 for the 1991 PEC. For inter-census comparability, the number of PSUs for PEC of 2011 was 280.

- The sample had two-way stratification design. The primary stratification was based on the grouping of EAs by Rural, Other Urban (THQ), Municipality, and City Corporations (CCs).

- The second stratification was based on geographic locations within each primary stratum. Within each sampling scheme, each area of the country was represented in proportion to population size. No sampling of dwellings was undertaken within the EAs. Thus, the PEC enumerators had to visit every dwelling in each selected EAs. As a result, the sample was self-weighted in each stratum.

- The urban-rural breakdown of EAs for the PEC of 2011 is given in Table 2.1.

Table 2.1: Rural-Urban Breakdown of Selected EAs

Area	No. of EAs covered
Rural Areas	140
Other Urban (THQ)	20
Municipality	60
City Corporations	60

The urban areas covering all three types were over-sampled for several reasons. First, even though the population of urban areas constitute only one-third of the total population, it was decided to obtain separate estimates for the three categories of urban population, namely municipality/urban, other urban, and City Corporation. Second, and more important was the fact that during the earlier PECs it has been found that the extent of under enumeration is higher in urban than in rural areas.

The sample size of 280 EAs (i.e. about 33,600 households) is large enough to provide estimates of census coverage at the national (including rural-urban) level with reasonably low sampling errors. It needs to be emphasized here that the PEC can only generate reliable and accurate results if the sample is well-designed; its implementation is efficiently managed; the matching operation is meticulously done; and the data analysis and estimations are correctly executed.

To minimize non-sampling errors, adequate emphasis was placed on extensive training of enumerators/supervisors, supervision of fieldwork and developing an effective matching operation.

2.3 PEC Implementation

The PEC survey was conducted after the completion of the census enumeration, from 10-14 April 2011. Within each selected sample EAs, every household was covered using better quality and senior census fieldwork staff under the direct supervision of BIDS professionals. The interviewer posed a series of questions about the household regarding all persons present on the night of 9-10 April 2011 (PEC night) as well as on the census night (14-15 March 2011). In addition to obtaining basic demographic information about all household members, the PEC questionnaire also included questions regarding the presence of persons in the household on the census/PEC night ('Was the person present on both the census and PEC nights, or only on census night, or present only on PEC night?').

Subsequently, in a more protracted undertaking, matching of the PEC questionnaires with the corresponding census questionnaires was done to determine whether the people included in the PEC questionnaire were enumerated in the census at the same address. The final calculation of undercount incorporated the results of the matching exercise, plus some intricate imputations for different kinds of non-match, providing a more reliable *final* undercount.

The PEC was designed to provide an independent check of census coverage. As such, it was important that the survey was conducted as independently of the census as possible. Even though BIDS was assigned responsibility for the PEC, it was necessary for some aspects of the PEC to utilize the infrastructure developed for the census. However, measures were taken to ensure that the processes were as independent as possible. Thus, while the enumeration area boundaries defined were those prepared for the census, the listing of dwellings within these areas was redone for the PEC. In addition, while most of the interviewers for the PEC had previously worked as supervisors on the census, they were allocated to areas different from those where they had worked in the census.

2.4 Sample Design

As already mentioned, a sample of 280 enumerator areas (EAs) was drawn for the PEC--140 rural and 140 urban. The sample was stratified by rural-urban differentiation. The EAs were classified as rural. urban/municipalities, other urban (Thana headquarters), and City Corporations. The sample was drawn from a list compiled from administrative records created prior to the census enumeration. It may be mentioned here that the number and boundaries of EAs changed to some extent during census enumeration; some areas were still being demarcated as enumeration began, while other EAs were split or combined during enumeration.

2.5 Questionnaire Design

The PEC questionnaire was brief. Every person who spent the previous night in the household was included in the PEC questionnaire. In addition, persons who were not present on the previous night but who spent the census night in the household were also included. The questions covered the following issues: Basic demographic information for each person present (age, sex, marital status, literacy). A reproduction of the questionnaire is included in Appendix A. There was also information on people who moved since the census night. Information was also obtained on in-and-out movers.

In order to identify possible *overcount*, the last question in the PEC questionnaire obtained a complete listing of all people who should be included in the census questionnaire. This was done by asking about people who were absent at the time of the

PEC but present on census night. People are required to be enumerated in a census questionnaire at the address where they spent the census night of 14-15 March 2011. Thus, for household members who had not moved since the census, anyone present on the census questionnaire but not on the PEC questionnaire should, in theory, not have been entered into the census questionnaire at that household, as they were not present on census night. These people may have been over-counted if they were also entered in the census questionnaire at the address where they actually spent the census night. However, examination of the PEC showed that it was not so easy to draw this conclusion as there were various causes of discrepancies between the PHC and PEC listings. For example, as the census was conducted over a period of 5 (five) days, many people were enumerated at a location other than their census night addresses. However, it was possible to distinguish the people absent at the time of the PEC but present on the census night from those present at the time of the PEC, due to the way in which the information was captured.

2.6 Information on Coverage and Content Error

The questionnaire collected information on persons who resided at a sample address on census/PEC nights along with relevant information to ensure that persons should have been enumerated at the current address. The PEC questionnaire was designed in such a way that the status of each person with regard to their presence on the census/PEC night can be ascertained.

The basic set of questions, in line with the census questionnaire, included:

(i)   What are the names of all persons living in the household on the PEC night?
(ii)   What is the relationship to the head of household or reference person?
(iii)   What is each person's age and sex?
(iv)   What are the names and relationships of other people who lived in the household on census night but not on the PEC night?
(v)   Was the person present on both the census and PEC nights, or was present only on the census/ PEC night?

For measuring coverage error due to omissions, duplications and erroneous inclusion, the PEC questionnaire had the provision for the classification of each listed person in a particular household as non mover; out mover, in mover, or out of scope. Based on this, the status of each person in the household was determined, as shown in Table 2.2.

Table 2.2: PEC Enumeration Status of Household Members

	Status
1	Non mover
2	Out mover
3	In mover
4	Out of scope (Born after)

The above population parameters have been defined as follows:

(a) *Non mover:* a person who resided in a particular household as of the census date and still living at the PEC time (i.e. present on both census and PEC nights);

(b) *Out movers*: refer to persons who lived in the household on the census date but did not live in the household on the PEC date;

(c) *In mover*: a person who lived in the household on PEC date but did not live in the household on the census date. Such a person arrived in the household after the Census night;

(d) *Out of scope*: a person who does not belong to the target population of the Census date, for example, a child born after the census date and/or in a de facto census person who lived outside the country on census date but lived in the household on the PEC date.

The PEC questionnaire was designed in such a way that it was possible to thoroughly probe during the interview, to make sure that the *correct enumeration status* of each person in the household, was recorded. For movers, when in doubt, the respondent was asked for the exact date of arrival or departure of a household member, in order to determine if the person was an in mover or out mover respectively.

In order to estimate the Response Variance, the following are some of the socio demographic variables which were included in the census questionnaire and also repeated in the PEC questionnaire for matching content error. The listed variables below are relatively easy to measure, which are considered as important demographic and social variables, and worth measuring the response error, if any. They included:

(i)   Age
(ii) Sex
(iii) Marital status
(iv) Education level
(v)   Source of drinking water
(vi) Type of toilet facility

2.7 Phases of the PEC Process

The PEC procedure was divided into three phases. **The first phase** was the 'PEC-A Field Survey' covering 280 EAs. The field work for PEC-A began on 10 April 2011 after all census enumeration booklets have been transferred from the Zonal Offices and put in storage in the *Parishankhyan Bhaban* at the BBS Office in Dhaka. **The second phase** was the 'Headquarters Matching Operation'. A set of matching rules was developed for use by the persons engaged in matching operation. The purpose of the Headquarters Matching Operation was to classify each person listed in the Census and PEC-A according to whether he/she has been correctly enumerated in the census (a matched person) or tentatively missed in the census (a PEC person unmatched with the Census). This required a person by person match between PEC-A and the Census. **The third phase** consisted of 'PEC- B Field Follow-Up Operation'. The PEC-B had two purposes. Firstly, to estimate the number of erroneously enumerated persons included in the Census and secondly, to verify that persons classified as tentatively missed in the Census (non-matches) actually were missed.

2.8 Fieldwork for PEC

The PEC was designed to provide an independent check of the Census count and, accordingly, the team responsible for the PEC was not directly involved in the Census. In practice, PEC enumeration procedures were largely similar to those used in the Census. However, measures were taken to ensure that the PEC was conducted as independently of the Census as possible and was effective in identifying persons missed in the Census.

The PEC field workers were recruited from the BBS staff who had prior experience in the Census work. Although one may argue that this might have compromised the independence of PEC to a certain extent, this was adopted in view of certain advantages of the decision for the PEC. The recruitment procedures were designed in such a way that the staff/interviewers selected were drawn from those known to have performed well as supervisors in the Census. Using experienced interviewers familiar with the processes helped to ensure that PEC enumeration was as complete as possible. At the same time, independence between the Census and PEC was ensured by assigning the interviewers to areas other than where they had worked during the Census.

The entire PEC operation was performed under the overall direction of the Director General of BIDS. Around 350 staffs were involved in the PEC enumeration including 280 interviewers (one per EA), 51 field supervisors, and 14 coordinators (two per Division-one from BIDS and one from BBS). All PEC staff received two days of training. As the staffs were familiar with the Census questionnaire and the PEC questionnaire

was comparatively short and simple, the emphasis of the training was on the concepts and procedures specifically related to the PEC.

The Census enumeration consisted of two main phases. The first was demarcation, when the country was divided into enumerator areas (EAs) consisting of a sufficient number of dwellings (120 households on average) to form a reasonable workload for a Census enumerator. For each EA, an enumerator's summary book was prepared which contained description of the boundaries of the EA, a map or aerial photograph, and a listing of all visiting points within the area. The second phase was the actual census enumeration. During this phase, enumerators collected relevant information from all households in each EA.

The PEC consisted of two similar phases. Firstly, the selected EAs were identified and copies of the maps and boundary descriptions were obtained from enumerators' summary books prepared for the Census. Although it was a feature of the design to use the same boundaries as prepared for the Census, the visiting points were re-listed such that any points not listed during the Census could be included in the PEC. However, the PEC could not identify any households missed in the Census when they were not included within the boundaries of an EA originally demarcated for the Census.

Secondly, each household in the selected EAs was enumerated which was critical for identifying any household missed in the Census. This required the field workers to work diligently to ensure that every dwelling in the EA was enumerated and none were double-counted. As mentioned earlier, the PEC enumeration took place during 10-14 April 2011.

2.9 Methodology for Evaluating Coverage Error

As earlier stated, no matter how carefully a census is planned and executed; coverage and content errors are inevitable. Two types of error are of special significance in the analytical framework, namely, under coverage and erroneous enumeration. Whereas under coverage results from erroneously omitted/excluded persons, housing units and/or households in a census; erroneous enumeration includes duplicate or multiple enumerations such as enumerations that should have not occurred because housing units or persons do not exist or enumerations were wrongly assigned according to geographic or demographic subgroups.

The following are some of the other types of errors normally occur in census enumeration:

(i) The census may contain duplicate or multiple enumerations;

(ii) The census may have people or housing units designated to the wrong geographic area and hence would not match the PEC interview;

(iii) People may be included with less than perfect enumeration in the census so that there would be insufficient information for matching with the PEC interview;

(iv) The census can erroneously enumerate someone who should have been enumerated elsewhere or the enumerator could include fictitious persons.

2.10 First Phase Data Capture: Matching Census and PEC Data

Matching Process

The basic process of matching involves comparing names, addresses and demographic characteristics between Census and PEC results. In general, it is an operation whereby households/housing units and persons enumerated during Census and PEC operations are compared for similarities. A two-way matching process is normally used to identify Census omissions and erroneous inclusions.

Once the PEC questionnaires were returned from the field areas, the next stage required PEC questionnaires to be matched against the corresponding Census questionnaires to check the completeness of the Census enumeration. A set of matching rules was developed for persons engaged in the matching operation and they were also given two-days training before matching exercise was started.

The entire matching operation proved to be a challenging exercise. In the 280 EAs, there were around 34,000 households and 140,000 persons in the PEC and a match status had to be determined for all of them. It took a team of usually 30, but up to 60 during peak periods, people around three months to complete the process.

In the run-up to Census 2011, the census questionnaire asked for only 'first name or initials to make it easy to complete the questionnaire'. The lack of detailed information regarding names made matching to the same address difficult at times. Other obstacles faced included the lack of detailed addresses in many parts of the country and the logistical difficulties encountered in locating census questionnaires at the BBS processing centers. It took unusually lengthy time for some of the PEC EAs to find out the corresponding census EAs (about 60 EAs), which delayed the timely completion of the matching exercise.

Matching involved several stages: matching of EAs, matching of households and matching of individuals. Matching was actually done in two phases. For each of the EAs, all households and persons were independently matched by two matchers- 'A' matcher and 'B' matcher. During the first phase, strict matching rules were followed resulting in obvious 'matches' and 'possible matches'. Both the matchers independently

matched the PEC and Census households and persons. Their main task was simply recording whether each household and person was found, were completed-detailed information being recorded on a 'matching sheet', a reproduction of which is included in Appendix B. This matching sheet had two main purposes:

- To record whether the household was matched and, if not, whether it was missed or unresolved and the reason why.
- To record the identifying information for the Census questionnaire against which the PEC questionnaire was matched.

In the matching operation, neither matcher had any idea about how the other matcher had matched the persons or households in a particular EA. However, during the matching exercise it was found that there was no one-to-one correspondence between A and B matchers all the times. After the completion of two independent matches, a 'Reviewer' was assigned the responsibility to examine the matching of two matchers and compile the final match status.

Matching of EAs

The first stage of matching an EA involved locating the Census questionnaires corresponding to the PEC EA. This was often not as straightforward as was expected. The process should have just involved obtaining the box/packet of questionnaires for the census EA corresponding to the PEC EA. However, the PEC EAs were based on the areas as determined prior to the Census and any changes made during the actual Census enumeration were not included in the original listings so that EA numbers did not always correspond with the Census numbers. For example, during the Census, some EAs were renumbered, split, combined or had boundaries altered. Thus, the Census questionnaires for an EA with a different number, or for a combination of EAs, were required for matching to a PEC EA.

Attempts were made in the BBS Headquarters to resolve these problems but sometimes it was necessary for the staff to look at maps or other boxes of census questionnaires to find the corresponding EA or EAs. In some cases, it was even necessary to talk to the relevant district/regional Statistical Officers to try to determine what had actually happened. Thus, it took an unusually lengthy time for some of the EAs to locate the corresponding Census EAs. This affected about 60 of the PEC EAs, although ultimately matching could be done as required.

The problems with correspondence of boundaries between Census and PEC EAs no doubt made matching difficult but did not have any impact on the calculated undercount. The methodology was based on the match status of each household and person. If a household and its occupants were enumerated as part of another EA in the census, thorough searching procedures adopted during the matching operation ensured that they were not erroneously treated as undercount. If there were any doubts about

whether they were enumerated, every effort was made to resolve such cases in a satisfactory manner.

Sometimes, even when corresponding EAs were found, other problems meant that a visit to the sample area was necessary. For example, there were cases in some urban areas when a different numbering system appeared to have been used, with holding numbers used in the Census and street numbers used in the PEC, and the relationship between the numbers was determined by re-visiting the EAs.

Matching of Households

Once the corresponding census questionnaires were located, the next task was to match at the household level. This was done by comparing the address listings in the enumerators' summary books for the Census and PEC to identify the corresponding households. Where this was inconclusive, the questionnaires were compared to see if a match could be found based on names and household structure.

Each household was classified as *matched*, *missed* or *unresolved* on the matching sheet. Where a household was classed as missed or unresolved, the reason for this was recorded if known. In areas with formal/accurate addresses, matching was a relatively straightforward process and it was not difficult to identify a corresponding household or to confirm whether a household was missed in the Census. However, in areas without formal addresses, especially in urban areas, matching was more complex. Difficulties arose when the names of household members did not match or were not unique or when the composition of households had changed. Sometimes it was impossible to confirm whether a particular household and its members were enumerated or not. In these cases, the household was classed as unresolved, and the decision was taken on the basis of results of the follow-up visit.

Sometimes a corresponding questionnaire was found but the household present at the time of the census was completely different. This arose as the original household had left and a new household moved in between the Census and the PEC.

Matching of Persons

Once the corresponding household was identified, a match status was allocated to each person on the PEC questionnaire. In most cases, it was possible to identify a match on the basis of name. However, this was difficult in cases where, for example, a different name or initials appeared to have been used. In these cases, a judgment on whether or not a person was matched was made based on a number of variables including age, marital status, and gender. These variables did not have to be exactly the same for a match to be made as often, particularly for age, since the responses differed slightly.

Usually when a corresponding Census questionnaire was identified, all members of the PEC household were classified as enumerated or missed. However, there were still some situations where it was necessary to allocate a code of unresolved for a person on the PEC questionnaire. For example, where some characteristics of a person on the Census questionnaire were similar but others were different, and the structure of the households did not assist in indicating whether the persons matched. In such cases, the unresolved category was used.

Similarly, if a person was on the Census questionnaire but not on the PEC questionnaire, normally such persons were considered as potential over-count if the household was the same in the Census and PEC, and there was no out-mover. However, it was not possible to draw this conclusion easily for all cases because of potential undercount by PEC. Undercount/missing of people either by PEC or by Census were duly recoded during the matching operation.

2.11 Second Phase of PEC Data Capture: Field Follow-Up Survey

There were a number of steps involved in capturing the data through the follow-up survey. As already mentioned, two different matchers independently matched the PEC and Census households and members, while the Reviewer gave the final decision regarding match status of each person. However, a number of errors were discovered on the file, relating particularly to the numbering of questionnaires which resulted in distorted household counts, and to the manner in which the data was captured for some questions. As a result, a final stage of data checks through field follow-up survey (i.e. PEC-B data collection) and corrections was necessary before an adequate data set was available for the calculation of the final undercount rate and the establishing of final population estimates.

After the first stage of matching was completed, the field reconciliation visit was made to obtain additional information to help resolve 'suspicious' cases. The purpose of this visit was to minimize the net matching error (i.e. the difference between erroneous matches and erroneous non-matches). During the follow-up survey, additional steps were followed to determine match status of unresolved cases. For example, if a particular household in PEC was not found in a particular Census EA, a search was conducted in neighboring EAs as it was possible that some households may have been enumerated as part of the neighboring EAs both during Census and PEC. This occurred in situations because of (i) poorly defined EA boundaries or where EA boundaries were wrongly interpreted by field staff, especially during Census; and (ii) errors in Census coding of EA numbers. It was also ascertained whether the missed person was present or not present on the Census night. Some households and persons initially categorized as missed in the Census were ultimately matched on the basis of reconciliation visits.

The follow-up visit was directed to non-matched persons and households. A total of 61 EAs--23 rural EAs and 38 urban EAs--were included in reconciliation visits. These visits

14

gave an opportunity to identify erroneous census enumerations and the resolution of doubtful cases in order to arrive at a realistic match status for each Census and PEC element.

Chapter 3

Estimating Gross and Net Coverage Errors

3.1 Evaluating Coverage Errors

Dual System of Estimation

The dual system of estimation is used in PEC to estimate gross and net coverage errors. In the dual system, data from PEC is matched with those obtained by the Census to arrive at the true population. The resulting tally can be represented in a contingency or a two by two table. It symbolically shows inputs into the Dual System Estimation of the True Population (Table 3.1).

Table 3.1: Tally of a Two-by-Two Observation

	In census	Out of census	
In PEC	$\tilde{N}_{11}$	$\tilde{N}_{12}$	$\tilde{N}_{1+}$
Out of PEC	$\tilde{N}_{21}$	$\tilde{N}_{22}$	$\tilde{N}_{2+}$
Total	$\tilde{N}_{+1}$	$\tilde{N}_{+2}$	$\tilde{N}_{++}$

where,

$\tilde{N}_{11}$ is estimate of the number of people counted in both Census and PEC
$\tilde{N}_{12}$ is estimate of the number of people counted only in PEC
$\tilde{N}_{21}$ is estimate of the number of people counted only in Census
$\tilde{N}_{22}$ is estimate of the number of people missed by both Census and PEC
$\tilde{N}_{1+}$ is estimate of the total number of people counted in PEC
$\tilde{N}_{1+}$ is total number of people counted correctly in Census (erroneous inclusions factored out)
$\tilde{N}_{++}$ is estimate of the total number of people
$$\tilde{N}_{++} = (\tilde{N}_{1+})(\tilde{N}_{+1}) / (\tilde{N}_{11})$$

that is, number of people counted in Census divided by number of people counted in PEC)/ number of people counted in both Census and PEC

The Dual System Estimate raises the corrected Census total (where erroneous enumerations are subtracted from Census population) by the total estimate of the number of people in PEC divided by the estimate of the PEC number that matched to the Census.

Two-Way Matching

A two way match was conducted between Census records and the PEC to identify omissions and erroneous inclusions. The matching exercise, in addition, produced an estimate of the matched population.

The Dual System Estimation was followed to estimate the True Population of persons in households. The Candrasekaran   Deming estimator, assuming independence, is expressed as follows:

$$\bar{\bar{N}}_{++} = (\bar{\bar{N}}_{+1})(N_{1+})\big/(\bar{\bar{N}}_{11})$$

3.2 Correctly Enumerated Persons

According to Mule (2008), the concept "*correctly enumerated*" has four aspects, namely, appropriateness, uniqueness, completeness, and geographic correctness.

Appropriateness: This means that the person should be included in the Census. For example, since the census date was 15 March 2011, people who died before or are born after this date are not part of the "*target census population*". Indeed, records of erroneous enumerations like fictitious people are not part of the "*target population*".

 Uniqueness: This refers to the need to measure the number of people included in the Census and not necessarily in Census records. If there are duplicate records, the count of records must be reduced for the purpose of the Dual System of Estimation.

Completeness: This means that the Census record must be sufficient to identify a person. If records lack sufficient identification information, it will be difficult to ascertain whether such a person was appropriately and uniquely included in the Census. It may also not be possible to determine whether the person was included in the PEC. It seems that the lack of sufficient information on addresses poses a special challenge in countries like Bangladesh.

Geographic correctness: All those persons must be included in the Census and in the EAs who are supposed to be included. Enumerations outside the EA that are enumerated in the Census are not considered *correctly* included in the census for Dual System Estimation.

In order to come up with a good Dual System Estimates, every effort was made during the PEC operation to ascertain correctly the enumerated population.

Proportion of People Captured in the Census

Having defined the set of correctly enumerated persons, the next step in the Dual System Estimation is to estimate census coverage. The following formula has been used for this purpose.

The Census coverage rate (which is called the match rate) = $\bar{N}_{+11/}$ $\bar{N}_{1+}$ that is, matched population relative to the PEC population.

3.3 Steps in Dual System Estimation

A number of census coverage estimates have been calculated from the Census and PEC returns for selected population parameters. To identify all the elements that are essential in making Dual System Estimates, the following symbols have been used:

o = total number of non movers (estimated from PEC sample)
p = estimated total number of out movers (from PEC sample)
q = estimated total number of in movers (from PEC sample)
r = estimated total number of matched non movers (based on matched cases between Census and PEC sample)
s = total number of matched out movers (based on matched cases between Census and PEC sample)
t = estimated total number of matched in movers (from PEC sample)
u = total number of erroneous inclusions in the population (from Census sample)
v = total number of census cases correctly enumerated in Census but missed in PEC (from Census sample)

Using the above symbols, Table 3.2 presents the data that has been used to calculate the coverage and omission rates for rural area for the Population and Housing Census of 2011.

Table 3.2: PEC and Matched Results: Rural Areas

Symbol	Items	Values
o	Non movers	66,316
p	Out movers	1,378
q	In movers	2,216
r	Matched non movers	63,887
s	Matched out movers	1,192
t	Matched in movers	1,917
m	Matched population	65,804
u	Erroneous inclusions	128
v	Correctly enumerated	3,031

Matched Population = matched non-movers + estimated matched in-movers (see, Dauphin and Canamucio, 1993:31)

$\bar{\bar{N}}_{11}$ = r + t = 63,887+1,917=65,804

A. Census Population Estimate: Rural Area

The census estimate is obtained as follows:

Census Population = matched non-movers + matched out-movers + population erroneously included in Census +population correctly enumerated in Census but missed in PEC

Census Population =r +s+u+v = 63,887+1,192+128+3,031=68,238

PEC Sample Estimate of Total Population

PEC Population = number of non-movers + in-movers

PEC Population = o+q= 66,316+2,216=68,532

Estimating Census Omissions

For calculating Census omissions, following steps have been used:

Subtracting the matched population from the PEC estimated population,

Census Omissions = 68,532 – 65,804
　　　　　　　　= 2,728

The census omission rate is the missed population relative to the PEC population estimate.

Omission Rate = omission*100/PEC population
= 2,728*100/68,532
= 3.980 percent

Coverage Rate is the matched population between the Census and PEC relative to PEC population. It is a complement of the omission rate.

Coverage Rate = $\frac{Matched\ population}{PEC\ Population}$ x 100

$$=65,804*100/68,532$$

=96.019 percent

Erroneous Inclusions and True Population

True population is calculated from the Census population after correcting for erroneous inclusions and divided by the coverage rate.

True Population = (Census population- erroneous inclusions)/coverage rate
=(68,238-128)/0.9602
=70,933

The erroneous inclusions, as earlier stated, include fabrications, out of scope, geographic misallocations, and similar factors. One of the objectives of the PEC is to provide an estimate of erroneous inclusions. This is obtained by dividing the number of erroneous persons by true population, and then multiplying by 100. This facilitates the correction in the Dual System Estimate of the True Population.

Erroneous Inclusion Rate = erroneously enumerated persons*100/true population
= 128*100/70,933
= 0.18 percent

Net Coverage Error

This is the difference between what should have been counted, that is, the True Population and what was actually counted in the census.

Net Coverage Error = True population – Census population
= 70,933 – 68,238= 2,695

Net Coverage Rate

This measure is the total net error relative to the Dual System Estimate of the True Population.
It is an important indicator of the quality of census coverage.

Net Error Rate = (True population – Census population)*100/True population

$$= (70{,}933\text{-}68{,}238)\text{*}100/70{,}933$$

$$= 3.799 \text{ percent}$$

This measure is, at times, approximated by taking the difference between the Omission Rate and the Erroneous Inclusion Rate.

Gross Coverage Error

This can be used as an indicator of the operational quality of the Census enumerations. It is the sum of omissions and erroneous inclusions.

Gross Coverage Error = Omissions + Erroneous inclusions
= 2,728+ 128= 2,856

Gross Coverage Error Rate

This is the absolute gross error relative to the Census enumerated population.

Gross Coverage Error Rate per Unit of Enumeration = (Omissions + Erroneous inclusions)*100/Census population

= (2,856*100)/ 68,238= 4.19 percent

Sometimes it can be approximated by summing up the Omission Rate and the Erroneous Rate.

Using the methodology as adopted above for the rural areas, the rates for urban areas-separately for Municipality, Other Urban and City Corporation as well as for rural areas and Bangladesh as a whole are presented in Table 3.3:

B. Estimating census omissions: Urban Area/Municipality

$$Coverage\ rate = \frac{Matched\ population}{PEC\ Population} \times 100$$

$$= 30{,}666 * 100 / 32{,}429$$

$$= 94.564\%$$

Omission rate $= (100 - 94.564)\%$

$$= 5.436\ \%$$

Erroneous Inclusions and its Rate
True Population

This is calculated from the census population after correcting for erroneous inclusions and divided by coverage rate.

True Population = (Census population- Erroneous inclusions)/Coverage rate
=(31,568-60)/.9456
=33,320
Erroneous inclusion rate = Erroneously enumerated persons*100/True population

$$= 60*100/33{,}320$$

$$= 0.18\ percent$$

Net error rate = omission rate-erroneous inclusion rate

$$= (5.436 - 0.18)\%$$

$$= 5.256\ \%$$

C. Estimating census omissions: Other Urban

$$Coverage\ rate = \frac{Matched\ population}{PEC\ Population} \times 100$$

$$= 10{,}654 * 100 / 11147$$

$$= 95.577\%$$

Omission rate $= (100 - 95.577)\%$

$$= 4.423\%$$

Erroneous Inclusions and its rate
True Population

True Population = (Census population- Erroneous inclusions)/Coverage rate
=(11,034-40)/.95577
=11,503
Erroneous inclusion rate = Erroneously enumerated persons*100/True population

$$= 40*100/11{,}503$$

$$= 0.35\%$$

Net error rate = omission rate-erroneous inclusion rate
         = (4.423-.35)%
         =4.073%

D. Estimating census omissions: City Corporation

Coverage rate = $\dfrac{Matched\,population}{PEC\,Population}$ *x 100*

         $=27,129*100\big/28,276$

         =95.943%
Omission rate=(100-95.943)%
         =4.057%

Erroneous Inclusions and its rate
True Population

True Population = (Census population- Erroneous inclusions)/Coverage rate
=(27,197-57)/0.95943
 =28,287
*Erroneous inclusion rate = Erroneously enumerated persons*100/True population*
              *= 57*100/28,287*

              = 0.20%
Net error rate = omission rate-erroneous inclusion rate
         = (4.057-0.20)%
         =3.857%

Table 3.3: Census Coverage Estimates by Rural- Urban Category

Area	Type of estimate			Net error rate of PEC 2001 (conducted by BBS)
	Coverage rate	Omission rate	Net error rate*	
Bangladesh	95.851	4.158	3.971	4.98
Rural	96.019	3.980	3.799	4.54
Municipality/urban	94.564	5.436	5.256	5.81
Other urban	95.577	4.423	4.073	3.73
City Corporation	95.943	4.057	3.857	7.67

Note: *For calculating net error rate for Bangladesh as a whole, the following weights have been used: 78.3 percent weight to rural area, and   21.7 percent weight to urban category with 7.98 percent weight given to City Corporation, 10.94 percent to Municipality, 2.79 percent to other urban category.

Chapter 4

Evaluation of Content Error

The evaluation of the content error in the Census involves the estimation of variance and bias components of total error in a census statistic. Content errors are errors in recording characteristics of persons who are enumerated both in Census and PEC. These errors may arise out of mistakes in data processing, interviewer's bias, respondents' bias, unclear questionnaire and misreporting. It is also known as response error, estimated only for matched persons and for selected variables, such as age, marital status, and education level.

The only response error component commonly estimated in PEC is variability and not bias. The variability can be measured by four different indicators, namely, **rate of agreement**, **net difference rate**, **index of inconsistency**, and **gross difference rate**. For the PEC of 2011, content error has been evaluated with respect to selected characteristics, namely age, sex, marital status, literacy, source of drinking water, and type of toilet facility.

4.1 Rate of Agreement

The rate of agreement indicates the degree to which the information given in the census matched with the information obtained in PEC for a matched set of individuals/ households. A low rate of agreement implies a high degree of variability and vice versa. Thus, the level of agreement between the responses of the census and the PEC provides the indication of quality of census information.

The rate of agreement (RA) is given by:

$$RA = \frac{1}{n}\sum_{i=1}^{s} y_{ii} \times 100$$

where y_{ii} = number of cases where i was given as response in both the census and the PEC

s = total number of response categories for characteristic X

n = total number of reported cases in both census and PEC

Rate of Agreement by Characteristics

At the national level, the rate of agreement is highest for the category 'sex' at 98.0 percent. This is expected because sex of individuals, in most cases, is apparent from their names. The 2 percent variation in sex might have arisen from enumerators' mis-recording or confusion or ambiguity in names that are shared by both sexes (i.e. Kazal, Shaheen, Ferdous), or during data entry. A similar high rate of agreement (95 percent) was found in the PEC of 2001 in Bangladesh.

Other variables resulted in lower rates of agreement than sex (i.e. marital status, source of drinking water, type of toilet facility and age). The rate of agreement for 'marital status' is 97 percent which is also high. This is because marital status is a simple question and there is little scope for misunderstanding/misinterpretation of marital status of a person. The category 'source of drinking water' ranked third (89.9 percent) in terms of rate of agreement between the census and PEC reporting. Marital status and source of drinking water, which are fairly stable variables, had high rates of agreement at 97 percent and 89.9 percent respectively in the PEC of 2011. The corresponding figures for the PEC of 2001 were: 86.9 percent for marital status and 85.7 percent for source of drinking water.

The level of agreement shows a somewhat lower rate for literacy at 87.4 percent and the lowest level of agreement was found for age (68 percent) closely followed by type of toilet facility (sanitary or others) at the household (69.5 percent). As expected, the rate of agreement is primarily influenced by the extent to which the respondent can provide a precise answer to a particular question. Regarding age, most respondents are not aware of their exact age since birth registration is not strictly followed. Thus, variation in reported age may be quite high, particularly for older age groups. For the latter, it may be partly explained by the fact that there is no clear-cut definition of what exactly constitutes a sanitary latrine. In some cases, water sealed latrines may be categorized as sanitary; while in some other instances, the same may be considered as non-sanitary. In the absence of a universal definition and lack of knowledge regarding sanitary toilet facility, the problem may be particularly serious in rural areas; the interpretation may vary depending on the respondent and also the interviewer due to ambiguity inherent in the definition.

The levels of disagreement that have been found are mostly due to difficulty in understanding a question and mis-recording/categorization by interviewers in census versus PEC. It needs to be mentioned here that the rate of agreement was also found to be the lowest (57.5 percent) for age in the PEC of 2001.

Gross Difference Rate (GDR)

The gross difference rate (GDR) is the complement of the rate of agreement. It measures the percentage of responses reported/recorded differently to questions asked in Census enumeration and PEC. The formula used is:

GDR = 100 – RA.

Net Difference Rate (NDR)

The net difference rate (NDR) is the difference between the number of cases in the census and the number of cases in the PEC that fall under each response category relative to the total number of reported persons in both the Census and PEC in all response categories. The NDR approximates the level of under-reporting or over-reporting for each response in the Census and the PEC relative to the total number of matched persons in all response categories.

The NDR is calculated using the formula:

$$NDR = \frac{Y_{ci} - Y_{pi}}{N_m} \times 100$$

where,
Yci = Un-weighted Census number of cases in the ith category,
Ypi = Un-weighted PEC number of cases in the ith category,
Nm = Un-weighted number of matched cases.

This is a measure of bias only when the re-interview is considered more accurate than the original response.

Table 4.6 shows the NDR of various characteristics. It can be seen that, on average, almost all categories under sex, marital status and literacy are either under- or over-reported by less than one percent. However, with respect to toilet facility and source of drinking water, most cases are over- or under- reported by about two to four percent.

Sex is highly accurately reported, with a tendency of females being under-reported while males are over- reported. In the categories of marital status of household members, only currently married are over- reported by 0.40 percent, while the never married, widowed, and divorced categories are under-reported varying from 0.05 percent to 0.25 percent. Similarly, with respect to literacy, illiterates are under-reported by 0.56 percent; by contrast, literates and not reported categories are over-reported by 0.21percent and 0.34 percent respectively.

Table 4.6 also shows that 'tap' and 'tube-well' under source of drinking water category and 'type of toilet facility' being used are the two categories which are over- or under-reported on average by one to four percent among all the categories probed by the PEC. The high NDR of 'other type' under the type of toilet categories can be partly attributed to 'less than clear-cut' definition of what exactly constitutes a sanitary toilet or otherwise. This pattern could be due to some respondents' misunderstanding of the definition of the two categories and reporting one for the other. Tube-well as the main source of drinking water is under-reported by more than three percent and 'tap' and 'pond/canal' are over-reported by one to two percent.

4.2 Index of Inconsistency

The index of Inconsistency (I) is the ratio of the Simple Response Variance (SRV) to the total variance for a given item. It is computed for each response category i according to the formula below:

$$I = \frac{(y.i + yi. - 2yii)}{\frac{1}{n}[y.i(n - yi.) + yi.(n - y.i)]} \times 100$$

Aggregate index of inconsistency (for categories of characteristics)

The Aggregate Index of Inconsistency (IAG) is a summary measure of the index of Inconsistency (that is, or all the response categories of the characteristic as a whole). The computation formula is as follows:

$$IAG = \frac{(n - \sum_i^c yii)}{n - \frac{1}{n}\sum_i^c y.iyi.} \times 100$$

At national level very little inconsistency is observed in the reporting of sex with the index of Inconsistency being only 4 percent. Under marital status category, the aggregate index of inconsistency is more than five per cent. A review of all indices showed, on the average, much higher aggregate indices for toilet type, source of drinking water and literacy. The low index of inconsistency for sex and marital status may be explained by the fact that sex of a person is always known and marital status rarely changes.

Table 4.6 gives the index of Inconsistency for all PEC variables. For the marital status categories, the Index of Inconsistency for never married category is 4.55 percent. Equally low level of inconsistency is observed for the currently married category. However, for the other categories of marital status, the indices are high-11.43 percent for widowed, and 26.43 percent for divorced categories respectively.

On the other hand, type of toilet facility and sources of drinking water show much higher level of inconsistency (ranging from 25 percent to 55 percent).These very high indices are mainly due to the low numbers of consistent persons in these categories. For example, 'other sanitary' shows the highest level of inconsistency of 54.89 percent. This means that nearly 55 percent of responses for this category for the matched records were inconsistent. The table shows a decreasing level of the indices for 'sanitary' and 'non-sanitary' categories.

The Index of Inconsistency for the literacy level shows the literates having lower level of inconsistency of 20 percent, which is followed by 22 percent for the illiterates. Considering the source of drinking water, tap water showed the lowest level of inconsistency of 25 percent followed by tube-well (29 percent). The pond/canal category has the highest rates of 53 percent.

Table 4.1: Response Variance for Sex: Comparison of Census Enumerator and PEC Enumerator Reporting for 137910 Match Persons

As reported by the PEC Enumerator	As reported by the Census Enumerator		
	Total	Male	Female
Total	100.00	50.07	49.93
Male	49.92	49.00	0.01
Female	50.08	0.01	49.02

Table 4.2: Response Variance for Marital Status 10 Years and Above: Comparison of Census Enumerator and PEC Enumerator Reporting for 137910 Match Persons

As reported by the PEC Enumerator	As reported by the Census Enumerator				
	Total	Never married	Currently married	Widowed	Divorced
Total	100.00	49.57	48.60	3.42	0.41
Never married.	49.62	48.46	1.09	0.06	0.02
Currently married	46.19	0.99	45.02	0.16	0.03
Widowed	3.68	0.09	0.41	3.16	0.02
Divorced	0.51	0.03	0.08	0.05	0.34

Table 4.3: Response Variance for Source of Drinking Water: Comparison of Census Enumerator and PEC Enumerator Reporting for 31294 Match Households

As reported by the PEC Enumerator	As reported by the Census Enumerator				
	Total	Tap	Tube-well	Pond/Canal	Others
Total	100.00	20.67	76.33	3.00	0
Tap	18.44	15.67	2.66	0.12	0
Tube-well	79.53	4.89	72.97	1.66	0
Pond/Canal	2.03	0.11	0.70	1.22	0
Others	0	0	0	0	0

Table 4.4: Response Variance for Literacy 6 years and above: Comparison of Census Enumerator and PEC Enumerator Reporting for 130700 Match persons

As reported by the PEC Enumerator	As reported by the Census Enumerator			
	Total	Literate	Illiterate	Not reported
Total	100.00	50.28	48.39	1.33
Literate	50.07	44.77	4.94	0.35
Illiterate	48.95	0.28	42.70	0.98
Not reported	0.98	0.23	0.75	0

Table 4.5: Response Variance for Toilet Facility: Comparison of Census Enumerator and PEC Enumerator Reporting for 31294 Match households

As reported by the PEC Enumerator	As reported by the Census Enumerator			
	Total	Sanitary(water sealed)	Sanitary	Non-sanitary
Total	100.0	29.12	40.18	30.70
Sanitary (water sealed)	31.46	21.46	7.68	2.31
Sanitary (without water seal)	35.49	5.33	24.89	5.26
Non-sanitary	33.05	2.33	7.60	23.13

Table 4.6: Net Difference Rate and Aggregate Index of Inconsistency by different Characteristics

Characteristics	Number of Cases in Census	Number of Cases in PEC	Net Difference Rate	Index of Inconsistency
Sex				*3.96*
Male	69050	68841	0.15	3.96
Female	68860	69069	-0.15	3.96
Marital Status				*5.64*
Never married	68362	68433	-0.05	4.55
Currently married	64261	63705	0.40	5.43
Widowed	7420	5077	-0.25	11.43
Divorced	567	695	-0.09	26.43
Source of Drinking Water				*28.61*
Tap	6469	5771	2.23	24.68
Tube-well	23886	24887	-3.19	28.77
Pond/canal	939	636	0.96	52.77
Literacy				*24.5*
Literate	65715	65436	0.21	20.05
Illiterate	63248	63984	-0.56	22.29
Not reported	14329	13872	0.34	
Toilet Facility				*45.94*
Sanitary(with water seal)	9113	9845	-2.33	41.78
Other sanitary (without water seal)	12573	11105	4.69	54.89
Non-sanitary	9608	10344	-2.35	40.28
Age				*36.14*

Chapter 5

Implications of the Results and Lessons Learned

The PEC findings are supposed to give credibility to the Population Census 2011 results and would guide the users to better interpret the census results. However, best practices of conducting PEC could not be strictly adhered to in all respects during the process. For instance, due to time and resource constraints the same organization (BBS) that did the census also provided the personnel for the fieldwork, data matching, data analysis and field reconciliation for the PEC. This arrangement partially compromised the independence of PEC from the census, a cardinal assumption behind conducting the PEC.

If the results of PEC are to be credible, the independence of PEC from census should be observed at all stages of PEC implementation; such as fieldwork, data matching, data management and field reconciliation. The lesson here is that independence of PEC from census is vital and should be maintained at all costs. This can be achieved if the PEC unit in BBS is manned by the personnel who do not participate in census activities. Alternatively, an organization outside of BBS can be contracted to plan and implement the PEC, from the beginning to the end. Adequate funds for PEC should be allocated in advance to ensure smooth running of the activities of the survey.

The lesson learned is that if the budget and manpower for different stages of PEC is not planned, secured and strictly followed, it would cause delays in the implementation of different stages of PEC. In order to successfully implement the PEC by an independent body outside BBS, the outside organization should be well equipped to face the challenges in planning the exercise-- how to deal with the problems of fieldwork after a census, should have trained manpower to implement the matching procedures including data collection and field reconciliation, with adequate manpower and logistic facilities to carry out data processing, tabulation and analysis for write up.