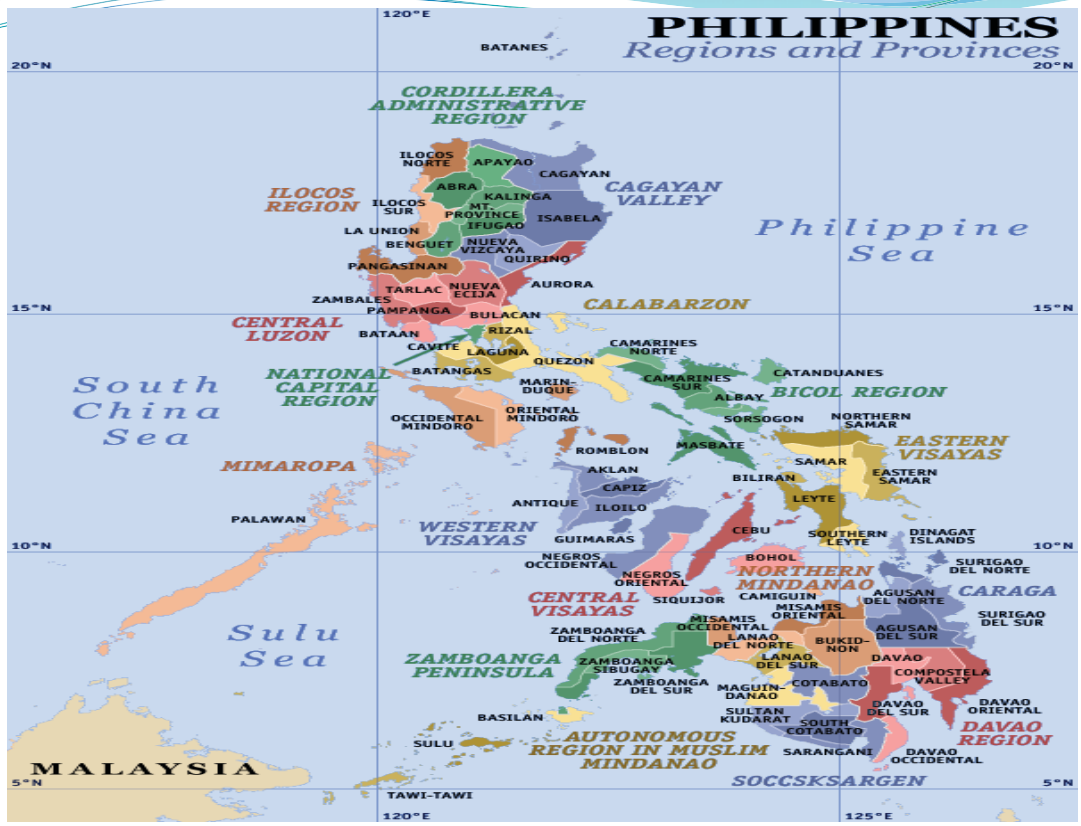
Census-taking in the Philippines

Gene V. Lorica, Philippine Statistics Authority

glorica@census.gov.ph

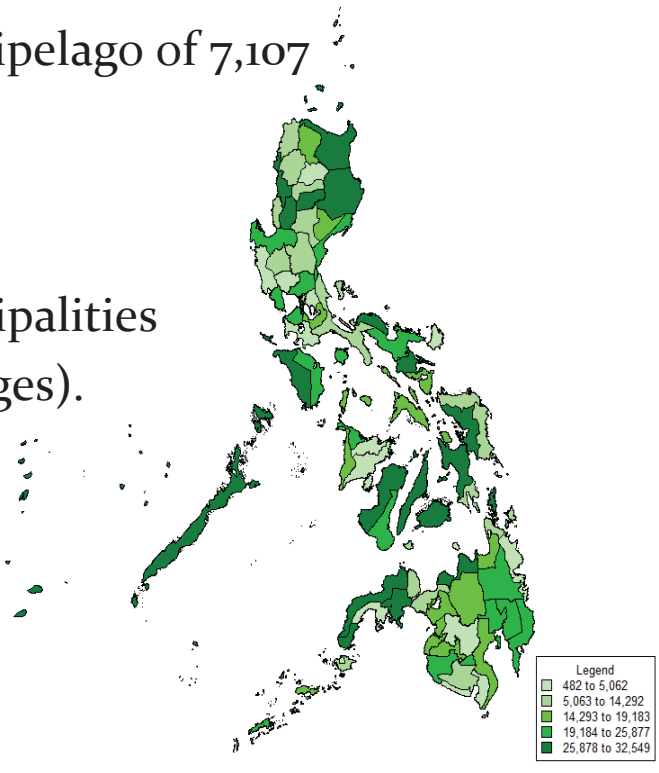
Census-taking in the Philippines

- Outline
  - About the Philippines
  - Census-taking Challenges
  - Recent Censuses
  - Direction (2015 Census of Population)



About the Philippines

- The Philippines is an archipelago of 7,107 islands
- 17 administrative regions
- 81 provinces
- 144 cities and 1,490 municipalities
- 42,028 *barangays* (or villages).



Census-taking Challenges

- Due to its topography, transporting documents from the Central Office in Manila down to the municipality is costly and difficult. It also restricts supervision of the enumerator in some areas
- Required by law that official count of the population be proclaimed 6 months after the field operation
- ICT infrastructure is still way below compared to what developed country has

Census-taking Challenges

- Only 9 percent of the households has subscription to an Internet Service Provider (ISP)
- Only 7.5 percent of the households has landline telephone although 79% of households has at least one member with cell phone
- Some of the areas have no existing road network and households in these areas have no distinct addresses

Census-taking Challenges

- Internal Revenue Allotment (IRA) share of local government is based on the population – this made some local officials influenced census operation by attempting to bloat the number of people in their area thru padding
- Most of those who live in exclusive subdivisions, high-end condominiums, and Filipino-Chinese in China town were uncooperative
- No population registry of the citizens

Recent Censuses

- 2000 Census of Population and Housing
- 2002 Census of Agriculture and Fisheries
  - a survey with very large sample
- 2007 Census of Population (Mid-Decade)
- 2010 Census of Population and Housing
- 2012 Census of Agriculture and Fisheries
  - the first 100% Census of Agriculture and Fisheries after 3 decades

2000 Census of Population and Housing

Census Method Used (Field enumeration)

- House-to-house visit in the area but if no competent respondent is available at the time of visit a revisit was scheduled or in some cases a self-administered questionnaire (SAQ) was provided and collected at a later scheduled date
- Every 5th household in an enumeration area (20% sample) were interviewed using the long form (CPH Form 3) and the rest using the short form (CPH Form 2)

2000 Census of Population and Housing

Census Method Used (Field enumeration)

- Interviewing was done simultaneously with listing of households
- At the end of each day, enumerators were required to write the summary of interviewed households in their accomplishment report form (CPH Form 10)

2000 Census of Population and Housing

Census Method Used (Manual Data Processing)

- Forms were bundled by EA and correctness of the geographic and household ID in each questionnaire were verified
- Coding was done
- Due to time constraints, data scrutiny was discontinued and the forms were sent to their respective Census Processing Centre (CPC) for scanning and data capture

2000 Census of Population and Housing

Census Method Used (Machine Data Processing)

- At the end of census period, all CPH Form 10 were encoded and tabulated for the preliminary count and eventually used as the official count after verification of entries with their respective entries in the Listing Form (CPH Form 1)
- Used document imaging system developed by a private company for the scanning (forms with drop-out colours) and data capture of the main forms using optical mark recognition (OMR) and intelligent character recognition (ICR)

2000 Census of Population and Housing

Problems (Field Operation)

- Could not provide the NSO management with 'real-time' status of the field operation
- Late discovery of under and over-coverage in some enumeration areas, i.e., long after the field enumeration

Problems (Cost)

- Census form is much more expensive than ordinary questionnaire
- Investment in software and hardware is quite high

2000 Census of Population and Housing

Problems (Machine Processing)

- Because the system was designed for data capture of documents in general but not specifically for survey/ census processing where forms must be processed by folio/batch, it resulted to difficulty in consolidating the data by enumeration area
- Data entry verification had to be done (using Centry of IMPS) for all extracted write-in data because it was discovered that the supposedly intelligent character recognition (ICR) functionality is not that intelligent as claimed by the software vendor

2000 Census of Population and Housing

Problems (Machine Processing)

- Has to resort to extensive automatic data editing due to the following:

  a) manual scrutiny was not done for all forms

  b) data processing was done in just 6 processing centre

  c) online data cleaning could not be done because the out-sourced system was not designed for that

2007 Census of Population

Census Method Used (Field enumeration)

- House-to-house visit in the area but if no competent respondent is available at the time of visit a revisit was scheduled or in some cases a self-administered questionnaire (SAQ) was provided and collected at a later scheduled date.

- Interviewing was done simultaneously with listing of households

2007 Census of Population

Census Method Used (Field enumeration)

- At the end of each day, enumerators were required to write the summary of interviewed households in their Accomplishment Report form (CP Form 10)
- At least once a week, CP Form 10 were submitted to Census Area Supervisor (CAS) who in turn send the summary data to the Quick Count System GSM modem installed at the CO thru text (SMS) before submitting the Forms to the Provincial Office (PO)

2007 Census of Population

Census Method Used (Manual Processing)

- Forms were bundled by EA and correctness of the geographic and household ID in each questionnaire were verified
- Coding was done
- Data scrutiny was done

2007 Census of Population

Census Method Used (Machine Processing)

- Every week, CP Form 10 summary data received and compiled at the CO were sent to their respective POs for verification of data with their corresponding CP Form 10.
- At the end of the census period, data from CP Form 10 were tabulated for preliminary count. The data was also used to validate the completeness of the data processing done (main forms) in each enumeration area

2007 Census of Population

Census Method Used (Machine Processing)

- Scanning and interpretation of OMR fields were done at the Regional Offices (ROs)
- Image files (of scanned forms) and interpreted data batch files were sent to the Central Office (CO) for data encoding of write-in entries using Key-from-Image program (initial version) and for online data cleaning using CsPro DE application program which was programmed to enable viewing the corresponding image of the current case in a separate window
- Official count of population was generated using the captured OMR data in each form

2007 Census of Population

Beneficial Effect (Field Operation)

- NSO management was provided with most current status of field operation at the start of each week
- Some areas were discovered to have been padding the population while the field enumeration is on-going
- Under and over enumeration in some areas were discovered right after the field operation in the area is declared as completed
- Areas deemed to be under-enumerated were subjected to saturation drive

2007 Census of Population

Beneficial Effect (Machine Processing)

- Faster data capture compared to conventional data entry method
- Was able to come up with the official count in time for what the law requires (6 months after the field operation)

Beneficial Effect (Cost)

- Saved cost on forms since the data processing did not require special form - requirement is ordinary paper of at least 100 gsm

2007 Census of Population

Problems (Machine Processing)

- Although data cleaning was done without referring back to the questionnaire, the use of CsPro DE application requires several mouse clicks and/or keystrokes to switch from CsPro data entry (DE) screen to the image window and vice-versa (an arduous task)
- Could not reflect the corrections in the image file, hence, after the data processing there are data in the file that is not the same as their corresponding entries in the image file

2007 Census of Population

Problems (Machine Processing)

- Automatic data editing resulted to significant number (but of lesser extent compared to the previous census) of entries imputed due to the following:

a) Some items were left blank during online editing since the data in the scan image file was illegible (not readable)
b) Since the questionnaires were left at the Regional Office, it would take time to refer back to the questionnaire
c) Some trivial data inconsistencies were accepted since the data editor at the CO are not familiar with the situation at the source area
d) Possibly because there was no data entry verification done

2010 Census of Population and Housing

Census Method Used (Field enumeration)

- House-to-house visit in the area but if no competent respondent is available at the time of visit a revisit was scheduled or in some cases a self-administered questionnaire (SAQ) was provided and collected at a later scheduled date.
- Every 5th household in an enumeration area (20% sample) were interviewed using the long form (CPH Form 3) and the rest using the short form (CPH Form 2)
- Interviewing was done simultaneously with listing of households

2010 Census of Population and Housing

Census Method Used (Field enumeration)

- At the end of each day, enumerators were required to write the summary of interviewed households in their Accomplishment Report form (CPH Form 10)
- At least once a week, CP Form 10 were submitted to their Team Supervisor (TS) who in turn send the summary data to the Progress Monitoring System GSM modem installed at the CO thru text (SMS) before submitting the Forms to DSO/CAS who in turn submits the collected forms to Provincial Office (PO)

CP Form 10 (CPF-10)

CP Form 10
(Accomplish in duplicate.)

Republic of the Philippines
PHILIPPINE STATISTICS AUTHORITY

2015 Census of Population
EN'S ACCOMPLISHMENT/PROGRESS MONITORING REPORT
(FOR PILOT CENSUS USE ONLY)

PROVINCE 14
CITY/MUNICIPALITY 20
BARANGAY 028
ENUMERATION AREA NUMBER 0100
EN'S CODE 130
STATUS OF ENUMERATION 1 (1 – Ongoing, 2 – Completed)

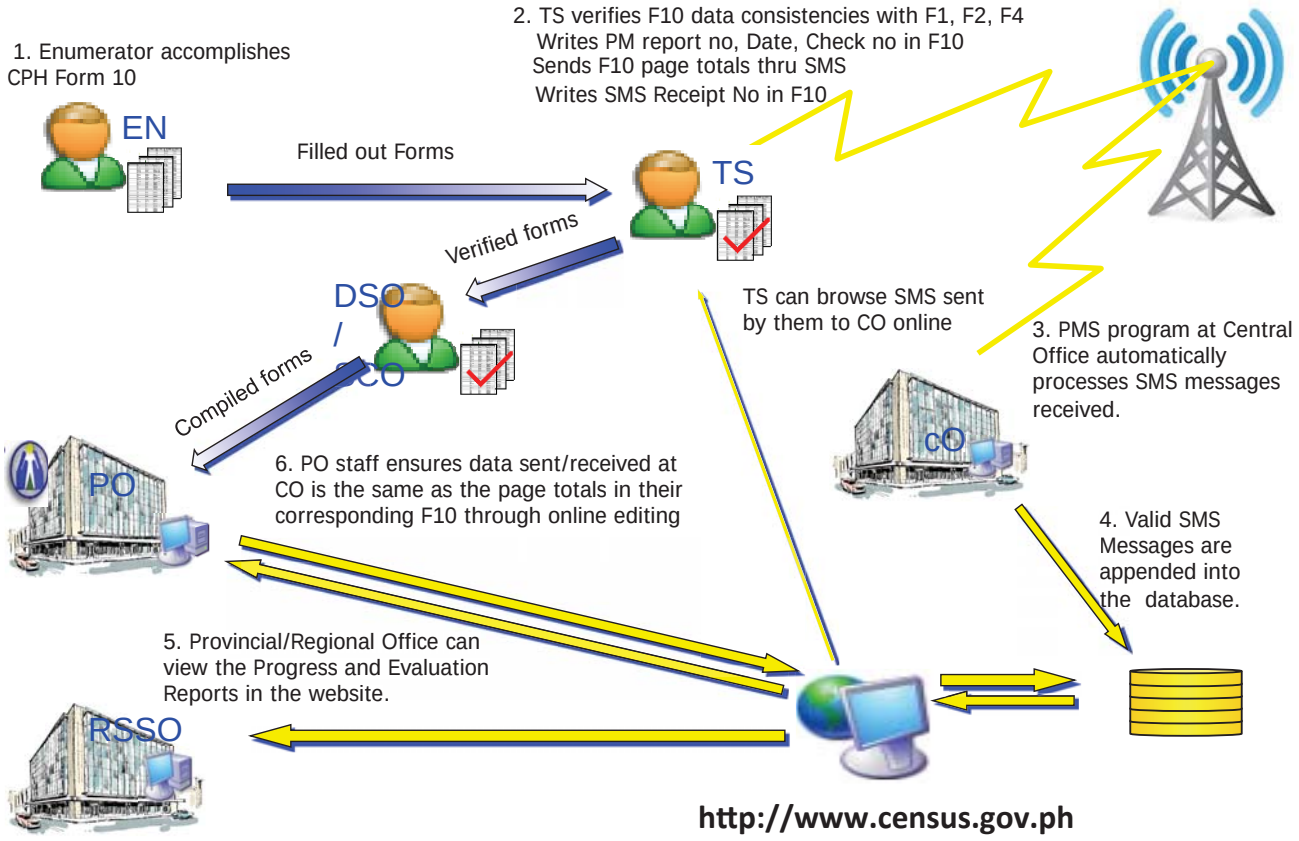
This portion is to be filled out only by the Team Supervisor (TS).
PM Report No. 1
Date (mm/dd) 0808
Check No. 0212
SMS Receipt No.

INSTRUCTION This form should be accomplished by the Enumerator everyday corresponding to the successful interviews of household and institution. Entries in this form should come from CP Form 1.

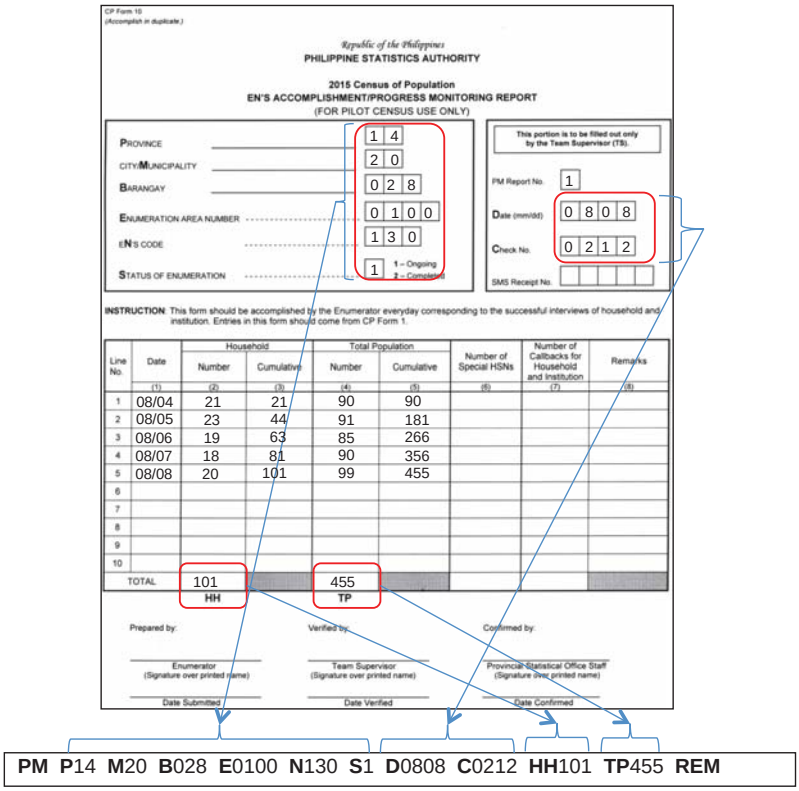
Line No.	Date	Household		Total Population		Number of Special HSNs	Number of Callbacks for Household and Institution	Remarks
		Number	Cumulative	Number	Cumulative			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1	08/04	21	21	90	90			
2	08/05	23	44	91	181			
3	08/06	19	63	85	266			
4	08/07	18	81	90	356			
5	08/08	20	101	99	455			
6								
7								
8								
9								
10								
TOTAL		101 HH		455 TP				

Prepared by: Enumerator (Signature over printed name) — Date Submitted
Verified by: Team Supervisor (Signature over printed name) — Date Verified
Confirmed by: Provincial Statistical Office Staff (Signature over printed name) — Date Confirmed

Progress Monitoring System



CP Form 10 (F10)



Republic of the Philippines
PHILIPPINE STATISTICS AUTHORITY

2015 Census of Population
EN'S ACCOMPLISHMENT/PROGRESS MONITORING REPORT
(FOR PILOT CENSUS USE ONLY)

PROVINCE 14
CITY/MUNICIPALITY 20
BARANGAY 028
ENUMERATION AREA NUMBER 0100
EN'S CODE 130
STATUS OF ENUMERATION 1 (1 – Ongoing, 2 – Complete)

This portion is to be filled out only by the Team Supervisor (TS).
PM Report No. 1
Date (mm/dd) 0808
Check No. 0212
SMS Receipt No.

INSTRUCTION: This form should be accomplished by the Enumerator everyday corresponding to the successful interviews of household and institution. Entries in this form should come from CP Form 1.

Line No.	Date	Household Number	Household Cumulative	Total Population Number	Total Population Cumulative	Number of Special HSNs	Number of Callbacks for Household and Institution	Remarks
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1	08/04	21	21	90	90			
2	08/05	23	44	91	181			
3	08/06	19	63	85	266			
4	08/07	18	81	90	356			
5	08/08	20	101	99	455			
6								
7								
8								
9								
10								
TOTAL		101 HH		455 TP				

Prepared by: Enumerator (Signature over printed name) — Date Submitted
Verified by: Team Supervisor (Signature over printed name) — Date Verified
Confirmed by: Provincial Statistical Office Staff (Signature over printed name) — Date Confirmed

PM P14 M20 B028 E0100 N130 S1 D0808 C0212 HH101 TP455 REM

2010 Census of Population and Housing

Census Method Used (Manual Processing)

- Forms were bundled by EA and correctness of the geographic and household ID in each questionnaire were verified
- Coding was done
- Data scrutiny was done

2010 Census of Population and Housing

Census Method Used (Machine Processing)

- Scanning and interpretation of OMR fields were done at their designated Scan Stations (either ROs or POs)
- Image files and interpreted data batch files were sent to their respective PO for data encoding and sample key-verification of write-in entries, and data cleaning using Key-from-Image program (improved version)

2010 Census of Population and Housing

Census Method Used (Machine Processing)

- Every week, CP Form 10 summary data received and compiled at the CO were sent back to their respective POs for verification of data with their corresponding CP Form 10. At the end of the census period, data were tabulated for preliminary count. The data was also used to validate the completeness of the data processing done in each enumeration area
- Official population counts were generated using the clean data

2010 Census of Population and Housing

Beneficial Effect (Field Operation)

- NSO management was provided with most current status of field operation at the start of each week
- Some areas were discovered to have been padding the population while the field enumeration is on-going
- Under and over enumeration in some areas were discovered right after the field operation in the area has been declared as completed
- Areas deemed to be under enumerated were subjected to saturation drive

2010 Census of Population and Housing

Beneficial Effect (Machine Processing)

- Data capture and data cleaning were much faster than the previous census
- The number of imputed data in automated data editing was greatly reduced to almost insignificant level, i.e., out of hundreds of thousand cases, only hundreds or even tens of cases were imputed by the program

2015 Census of Population

Census Method that will be used (Field enumeration)

- The same method as in 2007 Census of Population
- Use of self-administered e-Questionnaire in areas that are known to have households (most of them) that are difficult to interview
- Listing of households will be done before interviewing

2015 Census of Population

Census Method that will be used (Manual Processing)

- The same processing as in 2010 Census of Population except that coding and data scrutiny will be done during data cleaning in the machine data processing phase using Key-from-Image program – may still change depending on the result of the Pilot Census

2015 Census of Population

Census Method that will be used (Machine Processing)

- Scanning and interpretation of OMR fields will be done at their designated Scan Stations (ROs or POs)
- Scan image files and interpreted data batch files will be sent to their respective PO for data entry and sample key-verification of write-in entries, coding, and data cleaning using Key-from-Image program (full-fledged version)
- Preliminary count will be generated from verified 'texted' data
- Official population counts will be generated using the clean data file

Demonstration

- Preparing Key-from-Image Template File
- Data Entry using Key-from-Image program
- Data Cleaning using Key-from-Image program

Concluding Remarks

- This presentation focuses on the workaround for the inherent problem in using document imaging data capture technique because there are other statistical offices that still opted to or will still use this technology. I hope that this presentation would make them aware that there is an existing tool that can be used to do online cleaning of captured data to minimize “mindless” editing/ imputation by the automatic edit/impute program

Concluding Remarks

- The PSA will continue to improve the census-taking method and procedure to come up with the timely release of realistic population counts and characteristics by improving data collection and speeding up the data processing so that data could be fully utilized the soonest time possible

Concluding Remarks

- The craze nowadays in statistical offices all over the world is the use of handheld device for surveys/censuses which is reminiscent of the document imaging data capture technique more than a decade ago. With our disappointing experience in implementing new but untested technology in our 2000 Census of Population and Housing, we are now in a wait-and-see attitude towards using tablet PCs for household-based surveys/censuses. It is best to wait for true and honest report about the experience of other countries who tried it

Concluding Remarks

- Although we are now using the gadget (Android Tablet) for collecting prices of items in CPI market basket, we still have not used it for our household surveys due to the following apprehensions:

  One question one screen data collection system which is the easiest to develop in an Android environment may not be applicable to household based surveys/censuses. There is a big possibility that valid but wrong entries would be entered to a wrong person when asking and entering data about each person in a household, and for large-sized households, there is a great possibility of member(s) of the household being unintentionally missed in listing them

Concluding Remarks

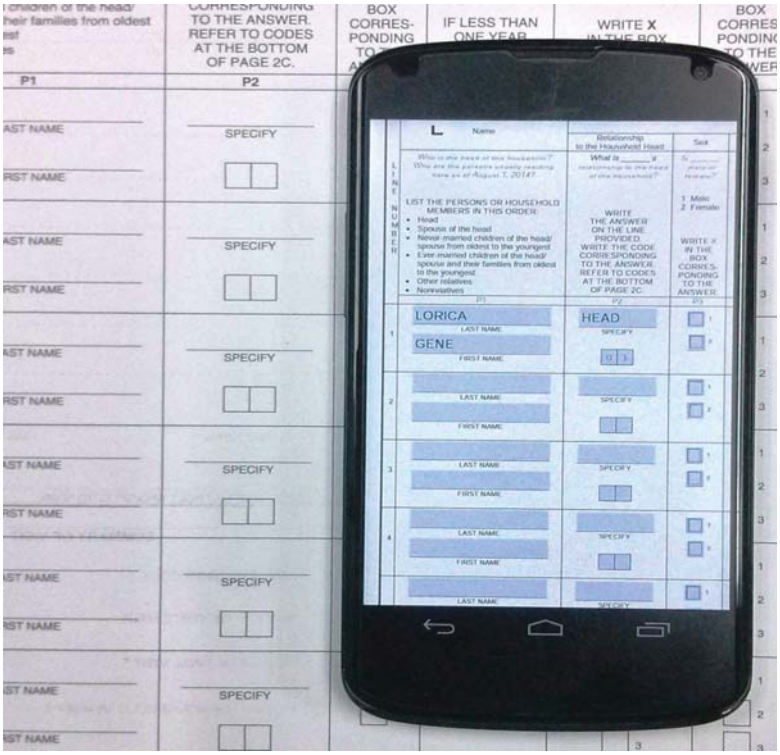
Apprehensions (continued):

Due to high cost of investment needed, failure will be very expensive

There is a possibility of conflict with existing law in the country regarding privacy/confidentiality when names of respondents and members of households are entered electronically

Entering responses in a tablet must have the 'feel and touch' of filling-out actual paper form – the same as the 'fillable' portable document format (pdf) file

Sample 'fillable' pdf in Android Tablet



Sample ‘fillable’ form using **Key-from-Image** program in Windows Tablet