

SECRETARIAT OF THE PACIFIC COMMUNITY

REGIONAL MEETING OF HEADS OF PLANNING AND HEADS OF STATISTICS (HOPS)
(Noumea, New Caledonia, 22–26 July 2013)

Session 5: CAPI - Computer Assisted Personal Interviews

(Document presented by the Secretariat of the Pacific Community)

EXECUTIVE SUMMARY

1. The *Ten Year Pacific Statistics Strategy* (TYPSS) objective 5 refers to the introduction of new and innovative statistical tools and systems. Since 2000, several such systems have been introduced, such as automated data capture (scanning) already referred to in Session 1, and the development of new survey methodologies. This paper refers to a third innovation – the field testing of CAPI technology during the 2012 Tuvalu census to test its comparative effectiveness and strengths relative to traditional paper-based interviews prior to the start of the 2020 Round of censuses, commencing with Kiribati and Palau in 2015.

Key achievements / Lessons Learned

2. Data Quality: all surveys and censuses require the capture of good quality data and the CAPI system provides the facility and capability to collect this.
3. The use of logic/consistency checks in the system assures a high degree of correctness in the data, and the option to run batch edits after the completion of each household gives more credibility to the data collected as data are verified in real-time (while enumerator is still with the household).
4. The cost-benefits of using the tool is more favourable in terms of printing costs, data entry costs, coding and training costs.
5. The timeliness of processing is another advantage with processing and data editing time much faster compared to manual data entry.

Recommendations

6. Given demonstrated gains in data quality and overall gains in enumeration effectiveness and efficiency, and with enumerators quickly adapting to this new technology, this technology could be readily implemented during the 2020 round of censuses, pending a careful country-by-country consideration of cost-benefits compared to traditional paper-based interviews.

Introduction

1. TYPSS objective 5 refers to the introduction of new and innovative statistical tools and systems. Since 2000, several such systems have been introduced, such as automated data capture (scanning) already referred to in Session 1, and the development of new survey methodologies. This paper refers to a third innovation - the field testing of CAPI technology during the 2012 Tuvalu census to test its comparative effectiveness and strengths relative to traditional paper-based interviews prior to the start of the 2020 Round of censuses, commencing with Kiribati and Palau in 2015.
2. CAPI stands *computer assisted personal interviewing*, a surveying technique that uses a computer-based questionnaire, where census or survey enumerators record people's responses straight into a PDA/tablet/notebook; some countries have recently also made use of mobile phones for their data collection.
3. With many countries having successfully introduced automated data capture (electronic scanning of forms) during the 2010 round of censuses to speed up data entry and improve data quality, Tuvalu, as the last PIC to undertake a census in November 2012, provided a good opportunity to field-test the CAPI methodology, and evaluate its usefulness and viability for use across the region, prior to the start of the 2020 round in 2015.
4. Four small laptop/tablets (*DirAction*) were purchased and used for the testing of the CAPI Application with the assistance of a consultant, Mr Ruben Hume, who had vast experience with survey CAPI applications and DHS and Census data processing. The application was tested in two census enumeration areas (EAs) on Funafuti, where two teams were selected to undertake field enumeration with the assistance of the SPC data processing staff who accompanied them during the trial.

Key achievements

5. The most important feature of CAPI is the QUALITY of data being captured. With all the in-built checks, questions and instructions in the application assures a high degree of correctness in the data being captured.
6. All necessary logic/consistency checks, filtering and skips are incorporated in the application which really helps the enumerators where they don't have to memorise filtering/skip questions. They don't even have to worry about logics in the responses as the system does it for them.
7. Questions and instructions for each question are shown on screen. This saves printing cost if the enumeration is done through paper questionnaires. It also minimises the hassle of carrying these questionnaires around within an EA. One can imagine doing a remote EA where households are far apart and infrastructure is not well established.
8. Use of listing and names of household members for referencing questions. It was a coincidence that for the Tuvalu Census, the listing of the households was done together with the actual census enumeration which allows for more consistency and efficiency in the data collection. The system uses the listing information to cross-check against what is being captured during the interview and it uses the person's names to reference in the questions.
9. Batch edits are executed after the completion of each household. At the completion of each household, the enumerators then select an option in the menu to run the batch edit for that household. Any errors detected are verified instantly there with the household members. Data are fully edited after the completion of each batch editing.
10. *Timeliness in data capture*. Fast to complete a household compared to manual filling of questionnaires (paper interview).

11. Since all information is captured during the interview, *there is no need for double entry*, as long as all the necessary internal consistency/logic checks are included in the system.
12. *The need for imputation is minimised* (with this process ideally becoming obsolete), as data verification takes place in real-time, while the interview is proceeding.
13. Summary information and cross-tabulations for provisional/preliminary results can be easily generated from the data being collected.
14. Cost effective. Saves a lot of printing cost, data entry cost, coding and less training time required of filling questionnaires.
15. CAPI versus Scanning. The most compelling argument in favour of utilizing CAPI is a noted improvement in data quality, as illustrate in completed survey and census forms.

Main challenges

16. *Availability of IT support*: With the use of computer/tablets requires the availability of technological support in terms of:
 - a. Networking and internet access for sharing/transferring of data and application files.
 - b. IT support for troubleshooting of hardware/software. Technical assistance to be readily available if there is a problem with the system. It is suggested to involve staff from the IT Unit as well as officers with some computing and CSPro background.
17. *Power*: Batteries need to be charged especially if used in the outer and remote areas where electricity is not available. One solution is to purchase a spare battery for each computer, where one could be charged while the other is being used.
18. *Backing up of data*: The provision of external drives or USBs for the enumerators to backup the data. Having internet access would be more reliable by transferring data to a central location.
19. *Automatic coding of occupation and industry codes*: This facility is not available in the current version of CSPro and one solution is to train the enumerators on coding techniques and the system uses 'pull-down' menus for them to select the codes from the list.
20. *Remoteness of EAs*: The issue of remoteness will always be an issue when using technologies and one option is to do a dual mode of collection where the CAPI system could be used in the areas with good technological infrastructure and use the manual data collection (paper questionnaires) in these remote areas.

Proposed way forward

21. It is anticipated that National Statistics office (NSO) will be introduced with new data collection tools that will improve data collection and dissemination. The adoption of these new technologies will largely improve the quality of data being collected as well as the timeliness of collection, processing and dissemination of information.
22. Usage of new technologies means new skills, knowledge and experience, hence, the need to build capacity and provide the necessary training to NSOs on the usage of these technologies, data processing software and IT.

Recommendations for consideration by HOPS

23. Given demonstrated gains in data quality and overall gains in enumeration effectiveness and efficiency, and with enumerators quickly adapting to this new technology, this technology could be readily implemented during the 2020 round of censuses, pending a careful country-by-country consideration of cost-benefits compared to traditional paper-based interviews.