

**ASEAN Meeting on 2010 Round of  
Population and Housing Censuses**

**31 July – 2 August 2006**

**Siem Reap, Cambodia**

Country Paper for Singapore

# **COUNTRY PAPER FOR SINGAPORE**

## **CENSUS OF POPULATION**

### **I Introduction**

The Singapore Department of Statistics (DOS) is the national statistical office responsible for conducting the decennial census of population in Singapore. Since the 1990s, the Department has identified three important exogenous trends as having a profound influence on the collection of social and demographic statistics in the census. First, the demand for comprehensive data on the population on a timely basis has been increasing. Secondly, the advances in Information Technology (IT) including the wide spread use of Internet, data warehousing software and integrated call-centre technologies have opened up new possibilities in data collection and capture. Thirdly, the development of public databases in the 1980s meant that a large amount of administrative data could be matched, captured and used for statistical purposes.

This paper describes the latest Census 2000 in Singapore and the adoption of IT advances and innovations in census operations to facilitate the conduct of the census.

### **II History of Census-Taking**

Census taking in Singapore is in line with the recommendation of the United Nations (UN) that a national census be taken at least once every 10 years.

Singapore's first census was taken in April 1871 as part of the Straits Settlements Census (see Box 1). Since then, regular censuses were undertaken at ten-year intervals up to 1931. The Second World War delayed the next censuses till 1947 and 1957. Singapore's first population census after independence was conducted in 1970. The next three censuses were conducted in 1980, 1990 and 2000.

In the 1980 and earlier Censuses, a traditional fieldwork approach was adopted. In the first stage, houses and other physical dwelling tenements were physically numbered to ensure complete coverage. The second stage involved a large number of field interviewers visiting the households to collect the information and to record them on paper forms. The third stage was conducted during or slightly after census reference day to confirm the validity of entries earlier recorded in the census forms. In the 1990 Census, basic demographic and personal particulars of individuals were obtained from government databases and

pre-printed on census forms for verification with respondents. Field interview remained the main method of data collection.

Singapore conducted a register-based Census, for the first time, for the 2000 Population Census. Using this approach, the full population count and basic profile of the population was obtained from the Department's Household Registration Database (HRD). Additional data items were collected from a sample survey of 20 per cent of dwelling units. These additional data items included marital status and fertility, education and language spoken at home, economic characteristics, mode of transport to school/work, overseas travel, as well as housing and household characteristics.

Box 1: Significant Events in Singapore's History of Census-Taking	
1871	First Census was taken for Singapore.
1880	First Census Ordinance in Singapore was enacted.
1957	The Singapore Department of Statistics was, for the first time, responsible for the taking of the census.
1970	A post-enumeration survey (PES) was conducted for the first time. The PES provided an independent means of assessing the coverage of the census and the quality of the information collected. The PES sample was in addition to the 10% sample enumeration conducted jointly with the main census.
1973	The Census Act 1973 was enacted.
1980	A 20 per cent sample enumeration was conducted jointly with the main census.
1990	Basic personal particulars were obtained from government databases and pre-printed on census forms for verification with respondents. There was no sample enumeration.
2000	First register-based census was conducted and de jure population concept adopted. Additional data items were obtained from a 20% sample enumeration.

### III Census of Population 2000

#### The Tri-Modal Data Collection Strategy

Singapore adopted a fully integrated tri-modal data collection system for the first time in the Census 2000 to obtain detailed information of the Singapore population not available from any administrative database. This tri-modal system comprised internet, computer-assisted telephone interviewing and field enumeration with a database at the backend.

For enumeration under the tri-modal data collection system, the 20% sample comprising some 218,000 dwelling units was split into six batches. The

enumeration of these six batches was done using the staggered workflow of two weeks between each batch as shown below. This staggered workflow was designed to optimise the use of time and scarce resources.

Batch	Weeks									
	1	2	3	4	5	6	7	8	9	----->
1	Internet	Internet	Internet CATI	Internet CATI	Internet CATI	Field- work	Field- work	Field- work	Field- work	
2			Internet	Internet	Internet CATI	Internet CATI	Internet CATI	Field- work	Field- work	Field- work
3					Internet	Internet	Internet CATI	Internet CATI	Internet CATI	Field- work
⋮ ▼ 6										

Respondents in a batch were first sent a notification letter informing them that they had been selected to participate in the sample Census enumeration. Following this, the respondents were then given 2 weeks to submit their returns via the Internet.

After the two-week period, the CATI system then began calling respondents in the batch that had not completed their returns. At this juncture respondents had the option of completing their returns via the telephone or through the Internet. For respondents who had not completed their returns by the end of the fourth week, a reminder letter was sent. This reminder letter also alerted respondents that if their returns were not submitted by the end of the fifth week, a fieldworker would visit them at their home to assist them in completing their returns.

At the end of the fifth week, Internet and CATI submissions for the respondents would close. Fieldwork for the respondents in the batch would then begin. This fieldwork would continue for four weeks until all respondents had completed their returns.

### *Internet Enumeration*

Singapore was among the first few countries in the world to collect Census information from households via the Internet in the Census 2000. With Internet enumeration, respondents enjoyed greater privacy, as their information would not be revealed to an interviewer, but transmitted directly to the Department's

database. Furthermore, the form-filling experience over the Internet was a positive and interactive one. When the respondent logged in to the Census website, some basic data already available in the pre-census database were displayed. User-friendly help features and explanatory notes were provided instantaneously when required. For further convenience, partially completed questionnaires could be saved by respondents and retrieved at a later time for completion.

From an operational perspective, Internet enumeration had many advantages. Most of the data collected from the Internet would already be electronically coded, thus reducing data entry and coding at the back end. Furthermore, there were substantial manpower savings since interviewers were not required to “canvass” information from the population.

### *Computer-Assisted Telephone Interviewing (CATI)*

Unlike Internet, CATI was a tried and tested data collection strategy, having been deployed for the mid-decade General Household Survey (GHS) in 1995. The CATI System for Census 2000 incorporated automatic branching of questions and on-line verification checks. A major improvement was the introduction of a fully automated dialing system in Census 2000. This automated dialing system searched through and dialed numbers on a phone list based on a set of “priority” rules built into the system. When a respondent picked up the phone, the line was automatically transferred to CATI interviewer who then interviewed the respondent. If the number was busy or if there was no response, the system would search for the next available phone number. By using this system, the CATI for Census 2000 was able to screen through over 10,000 phone numbers each day in a “round-robin” manner using an average of only 120 CATI interviewers.

### *Fieldwork*

Households were scheduled for fieldwork if they could not be contacted by CATI after a fixed number of telephone attempts. Fieldworkers visited these remaining households to conduct face-to-face interviews.

A computerised Fieldwork System (FWS) was developed to keep track of the movement of cases, enumerators' workload and status of the cases assigned to them. The workflow for typical fieldwork cases is as follows:

- a) household records with partial information obtained via the Internet and CATI as well as those not enumerated yet were transferred to the FWS upon Internet/CATI cut-off date;

- b) field supervisors allocated these household cases to enumerators;
- c) for greater efficiency, available information was downloaded and printed onto the Household and Individual forms at the regional offices to facilitate interviews;
- d) after obtaining the other missing information, interviewers/supervisors updated the completion status onto the FWS when they returned to their regional offices. The forms were transported to Census Headquarters to be electronically scanned

The register-based approach and the tri-modal data collection system significantly reduced the total number of enumerators from about 2,200 and 3,500 enumerators in 1980 and 1990 respectively to about 400 enumerators in 2000.

### Confidentiality and Security of Information

The tri-modal data collection strategy involved the collection of personal and household information via the Internet and CATI where face-to-face contact with respondents was avoided. To ensure confidentiality, all selected households received a notification letter with a house Identification Number (ID) and unique, randomly generated password. Using the house ID, password and the UINs of two members, respondents were able to log-on and retrieve their household record in the database via the Census website. The checking of the password was performed in a secure manner with Privylink. This used the password as a key to generate a random sequence at the respondent's computer. This random sequence was transmitted over the Internet. As the respondent's password was not sent across the Internet, the password could not be intercepted and read. At the server end, the random sequence received was decrypted with a key server. When the decrypted sequence matched, the respondent was authenticated and granted access.

All personal information provided by respondents was 128-bit encrypted before transmission over the Internet. This protected the information from unauthorised interception. To protect the information from hacking, a Demilitarised Zone (DMZ) utilising two layers of computer firewalls was set up to protect the on-line database in which the information was stored. These security measures were subjected to the most stringent tests and conformed to Infocomm Development Authority of Singapore (IDA)'s computer security requirements.

## Data Processing

Data collected via Internet and CATI were captured in the Census database directly without the need for data entry. However, additional processing had to be undertaken for data obtained from fieldwork. The completed census questionnaires obtained from fieldwork were scanned and data items were captured by Optical Mark Recognition (OMR), Optical Character Recognition (OCR) and Intelligent Character Recognition (ICR). These data items included sex, marital status, ethnic group and dwelling type. The combination of high-end scanning technology and recognition software resulted in time and cost savings as fewer data fields needed to be manually entered by data entry personnel. For data items where OMR/OCR/ICR cannot be applied, data entry was performed. Examples of such data items were highest qualification attained, major field of study and country of usual residence.

Once data were captured in the Census database, records with working persons were routed for coding of occupation and industry. The application software called ACE (Advanced Coding Environment) was specially designed and developed to perform computer-assisted coding. The use of ACE to perform industry and occupation coding represents a significant change from the past when coders had to flip through hundreds of pages of the industry and occupation dictionaries to search for the appropriate codes.

After the data had been coded, the records went through auto-correction, auto-detection, on-line verification and editing and duplicate checking.

## **IV Future Plans**

For future censuses, the Singapore Department of Statistics will continue to adopt advances in information technology as well as to capitalise on administrative data so as to reduce respondent burden and improve operational procedures. Personal Digital Assistants (PDAs), which had been successfully used for face-to-face interviews in the latest mid-decade General Household Survey 2005, would most likely be used in the forthcoming Census 2010. The use of PDAs would enhance the tri-modal data collection approach and help to improve field operations.