

# Combining Survey and Administrative Data for Singapore's Census of Population 2000

Bee-Geok Leow

*Singapore Department of Statistics  
100 High Street, #05-01, The Treasury,  
Singapore 179434, Singapore  
leow\_bee\_geok@singstat.gov.sg*

Eng-Chuan Koh

*Singapore Department of Statistics  
100 High Street, #05-01, The Treasury,  
Singapore 179434, Singapore  
koh\_eng\_chuan@singstat.gov.sg*

## 1 Introduction

Singapore adopted the register-based approach in its latest Census of Population 2000, moving its concept of population coverage towards a de jure count. Three key issues were considered carefully before the Singapore Department of Statistics finally implemented this new approach:

- a. The quality of administrative data in Singapore is sufficiently high to produce an accurate count of the population and its basic characteristics.
- b. The legal environment and data confidentiality practices in Singapore permit the sharing of non-sensitive administrative information.
- c. The cost savings in adopting this approach are substantial.

This paper studies the conceptual change in the definition of population, explains steps taken to improve the quality of statistics from the administrative records, and summarizes the difficulties and benefits of using a register-based approach combined with a survey to provide benchmark Census data.

## 2 Past Censuses and the De Facto Population Count

With the acceleration of the computerized programs in the public sector in the 1980s, there was a significant increase in the use of administrative databases in Singapore within the public sector in the late 1980s. Data from selected public databases were first used on a large-scale basis in the Population Census 1990. Enumerators visiting each household had pre-filled basic information on every household member from the merged database printed onto questionnaires.

The verification of information and enumerating everyone in the field meant that the count was still a de facto count. However, the experience gained gave positive indication that a register-based Census could be conducted in Singapore.

## 3 The Household Registration Database and the De Jure Population Count

The improved IT capabilities for linkages between various statistical systems and administrative sources in the mid-nineties led to the development of an integrated database system, known as the Singapore Department of Statistics' Household Registration Database (HRD).

The HRD formed the core database on population for the conduct of the register-based Census 2000. It provided the basic count of individuals and overall profile of the population.

This was made possible because every resident in Singapore had been issued with a personal unique identification number (UIN), and every foreigner staying or working in Singapore was issued a unique foreign identification number (FIN). This UIN/FIN was used as a key to match relevant government databases containing administrative information on individuals.

Each individual's address was checked against the National Database on Dwellings (NDD) to determine validity. The definition of "usual residence" was applied for all citizens and PRs with valid local addresses.

#### **4 Verification of Records and Data from HRD**

As the HRD formed the core database for Census 2000, it was important for it to be clean. Checks were performed to ensure that the quality of records from HRD were of a very high standard for use.

##### Systematic Error Corrections for Database

Comprehensive validation and verification rules were built into the database's ongoing updating cycles. Individual records that failed the validation checks were verified and corrected.

##### Address Checks

A special team of officers dealt with addresses in monthly updating cycles. Invalid addresses were investigated and checked with other published information or thorough fieldwork to ensure up-to-date information.

##### Comparisons with Official Population Estimates

Population figures based on HRD records for the years 1995 to 1999 were generated and compared with those obtained from the General Household Survey in 1995 and official population estimates. The differences and possible explanations were studied and corrective actions taken.

#### **5 Use of 20 Percent Survey for Census 2000**

The Census 2000 included a total of 54 data items, decided after consultation with key planning ministries. Eight items were obtained directly from the HRD. The remaining items not available from the administrative source were collected from a 20 percent sample of the population. The survey data were expanded to population size and together with the HRD data, provided the comprehensive Population Census 2000 results.

#### **6 Problems of Matching Survey and Administrative Data**

##### Possible Problem of Continuity with Change in Concept

The change from a traditional Census to a register-based Census entailed a change in concept from a de facto to a de jure count. With more Singaporeans working, studying or vacationing overseas, and more foreigners working or staying long term in Singapore, the distortions in excluding usual residents could have been significant for a small country. Comparisons revealed that the difference was as yet not significant.

### Database linkages

Different government agencies had different database platforms. Much preparatory work was carried out to overcome the differences when HRD was first set up in 1996.

### Erroneous Administrative Results / Coverage Errors

Administrative records also contained errors. The Department identified and resolved most of these prior to the start of Census 2000.

### Classifications - Occupation, Industry and Education

Changed administrative definitions or classifications caused difficulties in maintaining consistent definitions across different government agencies, and across time. National standards were developed through inter-Ministry consultations and the adoption of these national standards by the different agencies promoted.

### Confidentiality and Security

The use of an integrated database through merging of particulars meant that breaches in data security would have widespread consequences. In addition to emphasizing the legal protection powers of the Census Act, much care was exercised in the use of administrative and technical measures to protect the information obtained.

### Specific Reference Time Point

The conduct of the 20 percent survey within the Census, which relied on the database's records with an earlier cut-off date, would require survey respondents to update their information between the cut-off date and the actual Census day. Some information such as marital status, address etc. might change and needed to be updated in the field. The reference date was maintained as the Census Day for all information.

## **7 Benefits of the Register-Based Approach**

### Timeliness of Data Release

The register-based approach and integrated use of technology to provide a holistic solution to the entire workflow in data collection, processing and publication, reduced the turnaround time in the output of Census data.

### Reduced Respondent Burden

With detailed information canvassed from 20 percent of the population, four-fifths of the population was left undisturbed for Census 2000.

### Variety of Data Enlarged

The register-based approach also provided the facilities for a variety of administrative databases to be matched to Census database, thereby increasing the volume of data on the population without interviewing them.

## Cost and Manpower Reduction

It would have cost \$70 million for Singapore to do a traditional Census. This register-based approach coupled with new technological innovations shrunk the entire Census bill to only \$24 million. Only 600 persons were employed, compared with the estimated 6,000 for a traditional approach.

## **8 Population Estimates and Censuses for Future Years**

From 2001, the annual population estimates will be derived from the HRD, which captures changes in vital events and net migration through regular data merging with official sources. This would minimize the inaccuracies of inter-censal estimates based on the traditional approach, which would only get updated every ten years when the next Census is conducted. In addition, the estimates would provide basic profile of the population such as age, ethnic group, sex, house-type and other information, previously not possible with the traditional approach.

Future Censuses in Singapore would be reduced in scale and size with no compromise on data quality. This approach will minimize large shocks to the statistical system in terms of planning, logistics and manpower needs, and reduce respondent burden. The Singapore Department of Statistics envisages that the use of the HRD for future Censuses coupled with smaller scale surveys would continue with greater benefits to both data users and producers.

## **RESUME**

*Le recensement de la population de Singapour l'an 2000 adopte une approche à basée sur le registre pour compter sa population dans sa totalité. Cela est rendu possible car chaque habitant à Singapour est délivré d'un numéro unique d'identification personnel (UIN), et chaque étranger domicilié ou travaillant à Singapour a une passe avec un numéro unique d'identification d'étranger (FIN). Le profil de base de la population est ainsi engendré à partir de la Base de Données du Greffe de l'Etat Civil [Household Registration Database] (HRD) développé par le Bureau des Statistiques de Singapour. En outre, les caractéristiques socio-économiques détaillées de la population nécessaire à la planification et la conception politique sont exploitées à partir d'une enquête sur 20% de la population. Avec cette base de données, d'autres informations administratives importantes relatives aux ménages ou aux foyers y sont rajoutées pour élargir et enrichir ses domaines et son utilité. Cette approche implique que :*

- 1. Le compte à 100% et le profil de base sera disponible sous peu après la date de référence du recensement.*
- 2. La charge-répondant est fortement réduite puisque seul 20% de la population est contacté.*
- 3. L'association des données d'enquête et administratives donnent une variété d'informations en agrégation et à jour sur la population et sur les ménages.*
- 4. Cela se montre être une bonne approche et un moyen efficace d'associer les données provenant des enquêtes et des données administratives.*

*Ce document évalue de plus la qualité des données de la HRD (la Base de Données du Greffe de l'Etat Civil) et l'exploitation des sources supplémentaires de données administratives pour enrichir et élargir la base de données du recensement.*

8 May 2001