

For information

xxxx/2004-05

Legislative Council Panel on Financial Affairs

Computer Equipment and Services for the 2006 Population By-census in the Census and Statistics Department

PURPOSE

This paper briefs Members on the Administration's proposal to create a new commitment of \$32.314 million for acquiring the computer equipment and services required for the 2006 Population By-census (06BC) to be conducted by the Census and Statistics Department (C&SD).

BACKGROUND

2. It is an established practice in Hong Kong to conduct a population census every ten years and a by-census in the middle of the intercensal period for the purpose of obtaining up-to-date benchmark information on the social, economic and demographic characteristics of the population. Population censuses were conducted in 1961, 1971, 1981, 1991 and 2001 and population by-censuses in 1966, 1976, 1986 and 1996. In May 1999, the Finance Committee approved the creation of a commitment for acquiring the computer system and services for the 2001 Population Census (01C) computer system.

3. The plan of the 06BC was discussed at the meeting of the Legislative Council (LegCo) Panel on Financial Affairs on 2 April 2004, to which all LegCo members were invited to attend. No major comments were raised at the meeting.

JUSTIFICATION

The Proposed Computer System for the 06BC

4. The 06BC will comprise a detailed enquiry on one-tenth of the population, using a Long Form questionnaire (containing over 50 questions in a form of about 12 pages). It is a large-scale and complex operation involving recruiting and training of some 5 000 temporary field workers; enumerating about 220 000 households during the 18-day data collection period from 15 July to 1 August 2006; and processing of the completed questionnaires within a short period of time. These constraints can only be overcome with adequate computer support in various activities of the 06BC. The computer system will enable the release of 06BC results in the earliest instance to meet the increasing call for timely population statistics from both the Government and the community at large.

5. We have completed an in-house study to ascertain the computer requirements of the 06BC, in particular the feasibility of re-using the computer systems developed in previous population censuses/by-censuses to support the fieldwork operation and data processing work of the 06BC; and exploring the provision of more electronic options to improve the services to the public. The study concludes that the 01C computer system can be re-used and suitably enhanced to meet the requirements of the 06BC.

Support from Existing System

6. The following major functions of the 01C computer system will be retained for use for the operation of the 06BC -

- (a) to support the pre-fieldwork activities, including -
 - preparation and allocation of fieldwork assignments;
 - maintenance and production of digital maps;
 - printing of voluminous householder letters and assignment lists; and
 - recruitment, selection, training, deployment and payment of temporary field workers;
- (b) to provide monitoring mechanism to control and check the amount and quality of outputs of temporary field workers;
- (c) to facilitate the capturing of data items in the Long Form questionnaires and provide functions of editing and validating data collected; and
- (d) to facilitate the timely dissemination of 06BC results.

Inadequacies in the 01C Computer System

7. The 01C computer system is, however, inadequate in supporting in full the fieldwork operation and the data processing work of the 06BC in the following ways -

- (a) most of the computer equipment items used in the 01C have already been redeployed, as planned, for use in other computer projects of the C&SD after the completion of the 01C, in order to maximise the utilisation of the resources. To redeploy these items again would cause disruption to other on-going projects;
- (b) owing to fast evolving information technology, many of the hardware and software models used in the 01C computer system have become obsolete. Modification of the 01C application systems to enable their compatibility with the computer equipment lately available from the market is required;
- (c) the scopes of the 01C and 06BC are different. A detailed enquiry on a large sample of population will be adopted in the 06BC, while the 01C comprised a simple enumeration on the entire population and a detailed enquiry on a sample of the entire population. Changes in application functions of the 01C system, particularly the statistical estimation methods for the 06BC data, are needed; and
- (d) there are some differences in the content, features, structure and format between the digital maps used in the 01C and the 06BC. Adaptation to the latest digital maps provided by the Lands Department, by way of enhancement of the 01C system, is therefore necessary.

8. In the light of the inadequacies in the 01C computer system, the Commissioner for Census and Statistics, with the support of the Secretary for Financial Services and the Treasury and the Secretary for Commerce, Industry and Technology, proposes to acquire the necessary computer equipment and to hire services for making adaptation and enhancement where needed, to enable the 01C computer system to be re-used for handling the work of the 06BC.

New Functions

9. To provide better services to the public and to improve the efficiency of the 06BC operation, the following new functions/features will be added to the proposed computer system -

- (a) ***Improving data accuracy***
We will adopt the Intelligent Character Recognition technology to capture data on selected data fields of handwritten characters on the questionnaires. This would help to further avoid transcription error and repeated checking, hence improving data accuracy.
- (b) ***Improving efficiency of recruitment and training process***
We will establish on-line notice boards on the Internet for expediting communication with temporary field workers, who can then respond to the 06BC Office in a more timely manner, thus improving the efficiency of the recruitment and training process.
- (c) ***Improving field operational efficiency***
We will establish an electronic mail system to facilitate communication and timely transmission of the large amount of information between the 06BC Office (comprising five user sections) and 19 field centres (to be set up during the data collection period in the selected school premises). With improved communication, the field operation can be managed more efficiently and the enumeration progress monitored more closely.
- (d) ***Providing e-booking and e-questionnaires***
The system will enable respondents to make appointments for interviews and to provide their data in electronic form. This will provide added convenience to respondents and enhance the effectiveness of the 06BC operation.

10. Details of the 12 computer sub-systems of the proposed 06BC computer system and their functions are set out at Enclosure.

Encl.

Future Use of the 06BC Computer System

11. To maximise the return of investment in the 06BC computer system, the following sub-systems will be retained for use after the 06BC operation -

- (a) Register of Quarters Sub-system (RQS) - to facilitate continuous updating of a complete list of quarters and building details in Hong Kong which serves as the sampling frame for future population censuses/by-censuses and other household surveys;
- (b) Sampling Sub-system (SAMS) - for sample selection for future population censuses/by-censuses and other household surveys;
- (c) Digital Mapping Sub-system (DMS) - for maintaining digital maps for supporting population censuses/by-censuses and other household surveys;

and

- (d) Statistics Dissemination Sub-system (SDS) - for the ongoing dissemination of the 06BC results for various planning and analytical uses.

12. As in the past, we will re-deploy the equipment and facilities of the rest of the sub-systems of the 06BC to other operational needs in the C&SD with a view to maximising the utilisation of the resources available after the completion of the 06BC.

Cost and Benefit Analysis

13. Given the large scale and complex operation of the 06BC, it is impossible to conduct the by-census without efficient computer support. While it is the experience of statistical offices around the world that an efficient computer system is essential for such a large statistical operation, it is however not possible to quantify precisely the productivity gains and savings arising directly from the computer system. The commonly used cost and benefit analysis with an estimated payback period is therefore not entirely applicable in this case. The proposed system will ensure that the 06BC is conducted smoothly and statistics of good quality are available for use by the large number of users in both the public and private sectors.

14. The proposed 06BC computer system is designed to meet all the operational, timing and data quality requirements of the 06BC project that are essential for its success. It is designed based on the systems inherited from the 01C with computer facilities to be acquired initially for this dedicated purpose and subsequently to be re-deployed for use in other projects in the C&SD to ensure maximum cost-effectiveness. In addition, the proposed computer system will achieve notional savings of \$2,007,000 in terms of one-off cost avoidance for additional posts and administration cost savings.

FINANCIAL IMPLICATIONS

Non-recurrent cost

15. We estimate that implementation of the proposal will incur a non-recurrent cost of \$32.314 million over a five-year period from 2004-05 to 2008-09 for the acquisition of computer hardware and software and related services. Detailed breakdown is as follows –

| Non-recurrent cost | 2004 -05 | 2005 -06 | 2006 -07 | 2007 -08 | 2008 -09 | Total |
|-----------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|-------------------|
| | \$million | \$million | \$million | \$million | \$million | \$ million |
| (a) Hardware and Software | 1.662 | 4.545 | 0.585 | 0.553 | 0.037 | 7.382 |
| (b) Site Preparation | 1.641 | 0.946 | - | - | - | 2.587 |
| (c) Implementation Services | 0.278 | 4.490 | 5.298 | 0.064 | - | 10.130 |
| (d) Contract Staff | 1.438 | 5.026 | 3.544 | 0.112 | - | 10.120 |
| (e) Miscellaneous | 0.525 | 0.648 | 0.473 | 0.215 | - | 1.861 |
| (f) Contingency | 0.033 | 0.105 | 0.096 | - | - | 0.234 |
| Total | 5.577 | 15.760 | 9.996 | 0.944 | 0.037 | 32.314 |

16. As regards paragraph 15(a), the cost is for the acquisition of computer hardware, software and networking equipment. The hardware will include 2 host computers, about 140 personal computer workstations, peripherals and data communication facilities. The software will include operating systems, database management systems, application development tools, etc.

17. As regards paragraph 15(b), the cost is for site preparation. It includes the installation of trunks, power sockets and cabling work at the various sections of the 06BC Office and the setting up of a temporary data centre at the Central Processing Area to facilitate the data processing work of the 06BC.

18. As regards paragraph 15(c), the cost is for hiring of services for equipment installation, system configuration, system customisation and the data capturing services of the completed questionnaires.

19. As regards paragraph 15(d), the cost is for hiring of contract staff services for the system implementation work.

20. As regards paragraph 15(e), the cost is for rental of data communication link, acquisition of consumables and staff training.

21. As regards paragraph 15(f), the cost represents a 5% contingency on selected cost items set out in paragraphs 15(a) to (e).

Recurrent cost

22. The cost required for maintaining the computer equipment during the project period is included in the non-recurrent cost of the project.

23. Starting from 2007-08, staff effort will be incurred to perform the on-going system support. This requirement will be met by internal staff deployment in the C&SD.

IMPLEMENTATION PLAN

24. Subject to approval of funding, we plan to adopt the following implementation schedule:

| <u>Major Activities</u> | <u>Target Completion Date</u> |
|---|-------------------------------|
| (a) Formation of Project Teams | October 2004 |
| (b) Tendering and Procurement ¹ | May 2005 |
| (c) Site Preparation | |
| Phase I – By-census User Sections | January 2005 |
| Phase II - Central Processing Area | January 2006 |
| (d) System Design and Implementation ² | |
| Phase I – Sampling and Block-cutting | January 2006 |
| Phase II – Field Operation | May 2006 |
| Phase III – Data Processing | July 2006 |
| Phase IV – Statistics Dissemination | November 2006 |
| (e) By-census Field Operation | August 2006 |
| (f) Post-implementation Monitoring ³ and Support | February 2007 and onwards |

¹ Tendering for acquiring vendor service to automate the data capture of questionnaires will be completed by May 2005.

² System Design and Implementation for each related sub-system will be completed by phases from August 2005 to November 2006 to cope with the scheduled production date. Coverage of sub-systems in the different phases is set out at Enclosure.

³ Post-implementation monitoring and support for each related sub-system will be completed by phases from July 2006 to February 2007 to cope with the scheduled production date. For the sub-systems RQS, SAMS, DMS and SDS, on-going maintenance and support is required.

WAY FORWARD

25. We plan to submit the proposal to the Finance Committee for approval on 25 June 2004.

**Financial Services and the Treasury Bureau
June 2004**

**Proposed Computer Sub-systems
for the 2006 Population By-census**

| Ref. | Sub-systems | Main Functions |
|---|----------------------------------|---|
| <i>Phase I – Sampling and Block-cutting</i> | | |
| 1. | Registers of Quarters Sub-system | <ul style="list-style-type: none"> • Facilitate updating of quarters and buildings with additional details and better management control. • Provide a frame of quarters and building details for survey sampling and statistical listings. • Provide a quarters frame for sampling in support of the 2006 Population By-census (06BC) fieldwork operation. |
| 2. | Sampling Sub-system | <ul style="list-style-type: none"> • Extract samples for the special class enumeration in the 06BC. • Select samples of quarters in built-up areas from the sampling frame. • Produce related documents and printouts to support field enumeration. • Update the sampling history information for the sampled quarters. • Support the production of assignment lists for the 06BC fieldwork operation. |
| 3. | Block-cutting Sub-system | <ul style="list-style-type: none"> • Distribute the assignments to enumerators evenly before the 06BC operation. • Prepare the assignment lists and related documents before the 06BC operation. |

| Ref. | Sub-systems | Main Functions |
|--|-----------------------------------|---|
| 4. | Digital Mapping Sub-system | <ul style="list-style-type: none">• Make use of the digital maps of the Lands Department.• Enhance the map layer of built-up areas with buildings overlapping on the base maps of the Lands Department to facilitate the fieldwork operation for the 06BC.• Provide customised tools for address searching and map production.• Produce 5 000 sets of maps in support of the 06BC fieldwork operation. |
| <i>Phase II – Field Operation</i> | | |
| 5. | Enumerator Information Sub-system | <ul style="list-style-type: none">• Capture and maintain around 50 000 individual applications of temporary field workers.• Automate the recruitment and training processes by compiling training schedules and short-listing those qualified applicants for training and appointment.• Assign fieldwork assignments to each enumerator.• Maintain data for effecting the payment of temporary field workers in various stages.• Print householder letters and produce other 06BC related documents.• Support communication between 06BC offices and temporary field workers through electronic means. |
| 6. | Hot-line Enquiry Sub-system | <ul style="list-style-type: none">• Support public enquiry on fieldwork operation.• Provide online facilities to check identity of enumerators and information of field centres.• Support booking appointment for enumeration through electronic means. |

| Ref. | Sub-systems | Main Functions |
|---|--|---|
| 7. | Fieldwork Control Sub-system | <ul style="list-style-type: none"> • Facilitate field operation progress monitoring and control for 19 field centres. • Support quality check during fieldwork operation. • Provide data for compiling preliminary population estimates. • Control the despatch of Self-Administered Questionnaire left with each non-contact household at the end of the field operation. • Facilitate communication between field centres and headquarters through electronic means. • Provide control on data collection through enumeration and electronic means. |
| <i>Phase III – Data Processing</i> | | |
| 8. | Questionnaire Tracking Sub-system | <ul style="list-style-type: none"> • Keep track of the movement of questionnaires among different processing sites, so as to alert officers responsible when the questionnaires are undelivered or missing during delivery. • Facilitate the progress monitoring of individual data processing activities. |
| 9. | Data Input Sub-system | <ul style="list-style-type: none"> • Provide and support data collection through electronic questionnaires. • Capture the data on the completed hardcopy questionnaires by adopting the Intelligent Character Recognition/Optical Mark Recognition technology for subsequent processing. • Facilitate the coding of industry, occupation and addresses data items on the completed hardcopy questionnaires. • Integrate data collected via hardcopy questionnaires, electronic questionnaires and self-administered questionnaires. |
| 10. | Data Validation and Editing Sub-system | <ul style="list-style-type: none"> • Validate the data obtained from the hardcopy questionnaires, such as identifying record omission, record duplication, incorrect hierarchical order and incorrect data consistency, etc. • Provide facility to print and correct the identified data errors. |

| Ref. | Sub-systems | Main Functions |
|---|---------------------------------------|---|
| 11. | Imputation and Grossing-up Sub-system | <ul style="list-style-type: none">• Perform imputation to erroneous data and missing data items.• Gross up the data to produce estimates for the whole population.• Perform imputation and gross up the data. |
| <i>Phase IV – Statistics Dissemination</i> | | |
| 12. | Statistics Dissemination Sub-system | <ul style="list-style-type: none">• Prepare statistical tables.• Facilitate statistics dissemination via various means, such as publications, CD-ROMs and tabulation.• Perform statistical analyses and desktop publishing. |