**For discussion**
**on 25 June 2004**

**FCR(2004-05)19**

# ITEM FOR FINANCE COMMITTEE

**CAPITAL WORKS RESERVE FUND**
**HEAD 710 - COMPUTERISATION**
**Census and Statistics Department**
**New Subhead "Computer Equipment and Services for the 2006 Population By-census"**

Members are invited to approve the creation of a new commitment of $32,314,000 for acquiring computer equipment and services for the 2006 Population By-census.

## PROBLEM

The computer system inherited from the 2001 Population Census (01C) in the Census and Statistics Department (C&SD) is inadequate to support the fieldwork operation and data processing work of the 2006 Population By-census (06BC).

## PROPOSAL

2. 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 computer equipment and services for making adaptation and enhancement to the 01C computer system for handling the work of the 06BC.

**/JUSTIFICATION .....**

## JUSTIFICATION

### The Proposed Computer System for the 06BC

3. 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 the recruitment and training of some 5 000 temporary field workers; the enumeration of about 220 000 households during the 18-day data collection period from 15 July to 1 August 2006; and the processing of completed questionnaires within a short period of time. The exercise can only be conducted cost-effectively with adequate computer support. The computer system will also 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.

4. 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. We have also explored the provision of more electronic options to improve the services to the public. The study concludes that although some components of the 01C computer system can be re-used, some adaptation and enhancement would be necessary to meet the requirements of the 06BC.

### Support from Existing System

5. The 01C computer system will be re-used in 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 fieldwork 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 .....**

**Inadequacies in the 01C Computer System**

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

(a) most of the computer equipment items used in the 01C have already been re-deployed, 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 resources. To re-deploy these items again would cause disruption to other on-going projects;

(b) owing to the 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 computer 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 06BC. Adaptation to the latest digital maps provided by the Lands Department, by way of enhancement of the 01C computer system, is therefore necessary.

**New Functions**

7. Apart from addressing the inadequacies in paragraph 6 above, we propose to include the following new functions/features in the proposed computer system to provide better services to the public and to improve the efficiency of the 06BC operation -

(a) ***Improving data accuracy***

We will adopt the Intelligent Character Recognition technology to capture handwritten characters on selected data fields of the questionnaires. This would help to further avoid transcription errors and repeated checking, hence improving data accuracy.

/(b) .....

(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 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 selected school premises). With improved communication, the field operation can be managed more efficiently and the enumeration progress monitored more closely.

(d) ***Providing electronic booking and electronic questionnaire***

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.

8. Details of the 12 computer sub-systems of the proposed 06BC computer system, incorporating existing and new functions, are set out at the Enclosure.

Encl.

**Future Use of the 06BC Computer System**

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

(a) Register of Quarters Sub-system - 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 - for sample selection for future population censuses/by-censuses and other household surveys;

(c) Digital Mapping Sub-system - for maintaining digital maps for supporting population censuses/by-censuses and other household surveys; and

/(d) .....

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

10. 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 after the completion of the 06BC, with a view to maximising the utilisation of resources available.

**Cost and Benefit Analysis**

11. Given the large scale and complex operation of the 06BC, it is impossible to conduct the by-census without efficient computer support. For this reason, and same as the case for the 01C computer system, it is 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. This notwithstanding, it is the experience of statistical offices around the world that an efficient computer system is essential for such a large statistical operation. The proposed computer system will ensure that the 06BC is conducted smoothly and that statistics of good quality are available for use by the large number of users in both the public and private sectors.

12. The proposed computer system will achieve notional savings of $2,007,000 (including one-off cost avoidance for additional posts of $1,040,000 and administration cost savings of $967,000). These are additional costs, over and above the current project estimate, that would have been incurred had the 01C computer system not been enhanced for re-use.

**FINANCIAL IMPLICATIONS**

**Non-recurrent cost**

13. We estimate that implementation of the proposal will incur a non-recurrent cost of $33,665,000 (of which $1,351,000 will be met from existing resources of the C&SD) over a five-year period from 2004-05 to 2008-09 for the acquisition of computer hardware, software and related services. Detailed breakdown is as follows -

**/Non .....**

| | Non-recurrent cost | 2004 -05 $'000 | 2005 -06 $'000 | 2006 -07 $'000 | 2007 -08 $'000 | 2008 -09 $'000 | Total $'000 |
|---|---|---|---|---|---|---|---|
| (a) | Hardware and software | 1,662 | 4,545 | 585 | 553 | 37 | 7,382 |
| (b) | Site preparation | 1,641 | 946 | - | - | - | 2,587 |
| (c) | Implementation services | 278 | 4,490 | 5,298 | 64 | - | 10,130 |
| (d) | Contract staff | 1,438 | 5,026 | 3,544 | 112 | - | 10,120 |
| (e) | Miscellaneous | 525 | 648 | 473 | 215 | - | 1,861 |
| (f) | Contingency | 33 | 105 | 96 | - | - | 234 |
| | **Total non-recurrent funding requirement** | **5,577** | **15,760** | **9,996** | **944** | **37** | **32,314** |
| (g) | Staff cost | 394 | 676 | 281 | - | - | 1,351 |
| | **Total non-recurrent in-house staff cost** | **394** | **676** | **281** | **-** | **-** | **1,351** |
| | **Total non-recurrent cost** | **5,971** | **16,436** | **10,277** | **944** | **37** | **33,665** |

14. As regards paragraph 13(a), the cost is for the acquisition of computer hardware, software and network equipment. The hardware will include two host computers, plus about 140 personal computer workstations, peripherals and data communication facilities. The software will include operating systems, database management systems, application development tools, etc.

15. As regards paragraph 13(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 to facilitate the data processing work of the 06BC.

16. As regards paragraph 13(c), the cost is for hiring of services for equipment installation, system configuration, system customisation and the data capturing services of the completed questionnaires.

/17. .....

17. As regards paragraph 13(d), the cost is for hiring of contract staff services to provide support in project planning and monitoring, liaison with users, other Government departments and vendors, procurement, implementation and data conversion. It includes 48 man-months of Project Manager, 134 man-months of Systems Analyst, 103 man-months of Programmer, 2 man-months of Research Manager, 53 man-months of Statistical Assistant, and 44 man-months of General Clerk.

18. As regards paragraph 13(e), the cost is for rental of data communication link, acquisition of consumables and staff training.

19. As regards paragraph 13(f), the cost represents a 5% contingency calculated on the basis of contract staff cost in relation to the parts on project planning and monitoring, and system implementation.

20. As regards paragraph 13(g), the cost represents 24 man-months of one Analyst/Programmer I to provide support in the implementation of system enhancement. The C&SD will absorb the requirement from within its own resources.

**Recurrent cost**

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

22. Starting from 2007-08, staff effort of one Analyst/Programmer I 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**

23. Subject to approval of funding, we plan to adopt the following implementation schedule -

/**Major** .....

| | Major Activities | Target Completion Date |
|---|---|---|
| (a) | Formation of Project Teams | October 2004 |
| (b) | Tendering and Procurement | 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 | |
| | 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 and Support | February 2007 and onwards |

## BACKGROUND INFORMATION

24. 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 01C.

25. We consulted the Legislative Council Panel on Financial Affairs on the plan of the 06BC on 2 April 2004. No major comments were raised at the meeting. We also informed the same Panel of this proposal on computer equipment and services by circulation on 7 June 2004.

---

Financial Services and the Treasury Bureau
June 2004

**Enclosure to FCR(2004-05)19**

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

| Ref. | Sub-systems | Main Functions |
|---|---|---|
| ***Phase I – Sampling and Block-cutting*** | | |
| 1. | Register of Quarters Sub-system | • Facilitate updating of quarters and buildings with additional details and better management control.<br>• Provide a frame of quarters and building details for survey sampling and statistical listings.<br>• Provide a quarters frame for sampling in support of the 2006 Population By-census (06BC) fieldwork operation. |
| 2. | Sampling Sub-system | • Extract samples for the special class enumeration in the 06BC.<br>• Select samples of quarters in built-up areas from the sampling frame.<br>• Produce related documents and printouts to support field enumeration.<br>• Update the sampling history information for the sampled quarters.<br>• Support the production of assignment lists for the 06BC fieldwork operation. |
| 3. | Block-cutting Sub-system | • Distribute the assignments to enumerators evenly before the 06BC operation.<br>• Prepare the assignment lists and related documents before the 06BC operation. |
| 4. | Digital Mapping Sub-system | • Make use of the digital maps of the Lands Department.<br>• 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.<br>• Provide customised tools for address searching and map production.<br>• Produce 5 000 sets of maps in support of the 06BC fieldwork operation. |

| Ref. | Sub-systems | Main Functions |
|---|---|---|
| ***Phase II – Field Operation*** | | |
| 5. | Enumerator Information Sub-system | • Capture and maintain around 50 000 individual applications of temporary field workers.<br>• Automate the recruitment and training processes by compiling training schedules and short-listing those qualified applicants for training and appointment.<br>• Assign fieldwork assignments to each enumerator.<br>• Maintain data for effecting the payment of temporary field workers in various stages.<br>• Print householder letters and produce other 06BC related documents.<br>• Support communication between 06BC offices and temporary field workers through electronic means. |
| 6. | Hot-line Enquiry Sub-system | • Support public enquiry on fieldwork operation.<br>• Provide online facilities to check identity of enumerators and information of field centres.<br>• Support booking appointment for enumeration through electronic means. |
| 7. | Fieldwork Control Sub-system | • Facilitate field operation progress monitoring and control for 19 field centres.<br>• Support quality check during fieldwork operation.<br>• Provide data for compiling preliminary population estimates.<br>• Control the despatch of self-administered questionnaire left with each non-contact household at the end of the field operation.<br>• Facilitate communication between field centres and headquarters through electronic means.<br>• Provide control on data collection through enumeration and electronic means. |

| Ref. | Sub-systems | Main Functions |
|---|---|---|
| ***Phase III –Data Processing*** | | |
| 8. | Questionnaire Tracking Sub-system | • 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.<br>• Facilitate the progress monitoring of individual data processing activities. |
| 9. | Data Input Sub-system | • Provide and support data collection through electronic questionnaires.<br>• Capture the data on the completed hardcopy questionnaires by adopting the Intelligent Character Recognition/Optical Mark Recognition technology for subsequent processing.<br>• Facilitate the coding of industry, occupation and addresses data items on the completed hardcopy questionnaires.<br>• Integrate data collected via hardcopy questionnaires, electronic questionnaires and self-administered questionnaires. |
| 10. | Data Validation and Editing Sub-system | • Validate the data obtained from the hardcopy questionnaires, such as identifying record omission, record duplication, incorrect hierarchical order and incorrect data consistency, etc.<br>• Provide facility to print and correct the identified data errors. |
| 11. | Imputation and Grossing-up Sub-system | • Perform imputation to erroneous data and missing data items.<br>• Gross up the data to produce estimates for the whole population.<br>• Perform imputation and gross up the data. |

| Ref. | Sub-systems | Main Functions |
|---|---|---|
| ***Phase IV – Statistics Dissemination*** | | |
| 12. | Statistics Dissemination Sub-system | • Prepare statistical tables.<br>• Facilitate statistics dissemination via various means, such as publications, CD-ROMs and tabulation.<br>• Perform statistical analyses and desktop publishing. |