

UNITED NATIONS
ECONOMIC COMMISSION FOR AFRICA

~~b18~~
b11050871

REPORT ON A MISSION TO ETHIOPIA
(28 December 1982; 6, 10 and 11 January 1983)

By
D. Benzine
Regional Adviser in Data Processing
Statistics Division

and

E.A. Colecraft
Regional Adviser in Demographic Statistics
Statistics Division

January 1983
Addis Ababa

This report is the sole responsibility of the writers. It has been submitted to the Department of Technical Co-operation for Development, United Nations, which may alter or supplement the findings.

1. Introduction

The mission was undertaken with the concurrence of the Government to advise on the draft questionnaires for the 1983 Population and Housing Census and also the data processing requirements for the census.

Discussions were held on four afternoons (28 December, 1982; 6, 10 and 11 January 1983) with the following officers of the Central Statistics Office:

Ato Mitik Beyene, Manager;
Ato Abdulahi Hassen, Head, Population and Social Statistics;
Ato Makonnen Tekle-Haimanot, Team Leader, Social Statistics Experts;
Ato Habtemariam Tesfaghiorghis, Team Leader, Demography and Housing Statistics Expert;
Ato Hadgu Bariagaber, Team Leader, Vital Statistics Experts.

The Regional Adviser in data processing also held discussions with Mr. Jean Pierre Gogan, Director of SERIC.

2. Census time-table

A detailed time-table for the census has been prepared. It is expected that there will be a pilot census in April 1983 to test the census questionnaires and enumeration procedures. The main population and housing census is scheduled to be conducted in December 1983.

3. Enumeration procedures

The census is expected to be a de jure count of the population. It is also the intention of the census organizers to reduce office coding of the census information as much as possible. Most of the coding is therefore expected to be done by Enumerators during the census enumeration. It was pointed out that the recording of codes only by Enumerators in the field might have an adverse effect on the quality of the census data. It was therefore decided, after some discussion, that Enumerators should first record the responses in words and then enter the appropriate codes. It is our view that the amended procedure will make it possible to verify the coding which will be done by Enumerators. The coding of geographic identification particulars, occupation and industry will, however be done in the office.

4. Census questionnaires

Two types of questionnaires will be used for the census enumeration. These are: rural population and housing census questionnaire; and urban population and housing census questionnaire. The first draft of both types of questionnaires have been prepared and they formed the bases of our discussion during the mission. The two questionnaires are similar except that the rural questionnaire contains a limited number of items on housing characteristics.

The following are the items on the rural population and housing questionnaire: geographical identification particulars, serial number of household, name, relationship to head of household, sex, age, religion, ethnic group, language, place of birth, disability status, marital status, ability to read and write, highest grade completed,

activity status during most of last 12 months, occupation, industry, children ever born alive and number of births in the last 12 months to women 10 years and over, particulars of deaths in the household during last 12 months, total household members, material for construction of walls of housing unit, number of rooms in the housing unit, availability of room used exclusively as kitchen and whether livestock spend the night in the same room with members of household.

The main difference between the rural and urban questionnaires is in respect of the reference period for economic items and the housing characteristics. The urban questionnaire contains much more detailed questions on housing conditions. The additional items are: type of housing unit, type of structure use of housing unit (building), material for construction of roof, material for construction of floor, material for construction of ceiling, type of kitchen, type of tenure, amount of rent, source of water supply, type of bathing facilities, type of lighting, type of fuel, availability of radio, telephone and television in housing unit, and total number of households in housing unit.

In the case of the economic characteristics the rural questionnaire will be used to investigate the usual economic activity of the rural population (activity during most of last 12 months) whilst the urban questionnaire will be used to enquire into current activity of the urban population (activity during last seven days).

The format of the questionnaires as well as the items on the questionnaires were discussed with the census officials and a number of suggestions made. The suggestions which will be taken into consideration in the revision of the draft questionnaires include the following:

Arrangement of items on questionnaires: The mission suggested that as far as possible questions which are common to both rural and urban questionnaires should have the same column numbers on the questionnaires so that they might be recorded on the same positions on the magnetic media during data processing. Such an arrangement will permit the tabulation of common variables at the national level without unnecessary extra work.

Population count: It is our view that in addition to the de jure population for which detailed information would be collected, some limited information be also collected for the de facto population. This could be done by recording the names and residential status of both usual household members and their visitors (during the reference period) on the household questionnaire.

Education: The mission recommended that "school attendance" should be investigated in addition to "highest grade completed".

Place of birth: It was pointed out that though the question on place of birth is fairly simple and can be easily asked by Enumerators, it provides limited data for internal migration analysis. It was suggested that if space would permit the inclusion of an additional question then "duration of residence" should be investigated in addition to the question on place of birth. Otherwise "place of residence at a fixed prior date" may be the most satisfactory single question to obtain adequate data for migration analysis.

Economic characteristics: It is the view of the mission that the use of two different concepts of economic activity i.e. usual activity in rural areas and current activity in urban areas may make it difficult to compare the economic characteristics data for the two sectors. Further, usual activity is more difficult to investigate and it may present problems for the Enumerators in the field.

The main reason for considering the concept of usual activity for the rural area is the seasonal nature of agricultural activity. The problem of seasonality of certain types of occupation in either rural or urban areas, however, may be overcome with the inclusion of "had job but did not work" in the categories under activity status. The employed persons will then be both those who worked during the short reference period and those who had regular jobs but did not actually work for one reason or the other.

The issue was discussed at length and it was decided that both concepts would be used during the pilot census and a final decision taken in the light of the findings.

Housing conditions: It is the intention of the census organizers to use the housing unit as the unit of investigation for the housing conditions. Thus in cases where two or more households live in one housing unit, the information on housing conditions will be recorded only on the questionnaire for the first household interviewed. It is the view of the mission that household should be used as an additional unit of investigation for the housing census. This will permit the cross-classification of household characteristics by housing conditions.

Tabulation programme. It was suggested that the first draft of the tabulation programme should be prepared as soon as possible. It was also pointed out that it would be desirable for the tabulation programme to be prepared before the final revision of the census questionnaires.

Data entry workload estimates: In a previous visit to the CSO by the ECA regional data processing adviser, the census data entry workload was assessed (Ref: note for the file, 21 April 1982).

Taking into account the new draft questionnaires prepared, these estimates were revised during the mission and the details are given in Annex I.

5. Data processing equipment:

It may be recalled that the CSO which was using a rented NCR V-8455 computer has decided to replace it by a new Hewlett-Packard 3000 series 44.

CSO has had considerable problems with the use of the NCR V-8455. The system also does not support FORTRAN and the Statistical software packages required by CSO.

The HP computer is more reliable (1) and there are similar equipments in Addis Ababa at ECA, ILCA and Ethiopian Airlines. This equipment is serviced by SERIC, which has now a local representation in Addis Ababa. This computer, which was purchased with funds from UN agencies (UNICEF, UNDP and FAO) (as executing agency) has just been installed in a new site with the following configuration.

CSO computer: Present configuration:

- 1 HP 3000 Series 44 CPU with 1 Megabyte of Store (32440 B)
- 2 Disk drives of 404 Megabyte each (fixed) 7935 H
- 2 Magnetic tape decks 1600 Bpi (7970 E)
- 2 Line printers 300 lpm 132 cpl (BDS)
- 12 Workstations with cables (2622 A and 13232 C)
- 3 Terminal controllers (30018 A/44)
- 1 Current regulator (30018 A)
- Compilers (COBOL, FORTRAN, RPG and Basic)

Cost: US \$241,433 (excluding transport and staff training charges)

The present equipment at CSO cannot handle all the census work, especially data entry. It is therefore recommended to upgrade the HP 3000/44 configuration.

Given the population size and the data content, the number of workstations required to complete the data entry task depends on factors such as the percentage of verification, the number of shifts of operation and the time allowed for completing the task.

There are now 12 workstations available: one for console, and the others are used for interactive programming and data entry. The HP 3000/44 system can support up to 60 (point to point) workstations, that is 48 more. With 48 workstations being used exclusively for census data entry, the work can be completed in 15 months if work is done in two shifts, the operators attain a speed of 8000 key depressions per hour per operator and the percentage of records verified is about 50 per cent (2).

(1) The manufacturer and the servicing agency guarantee by contract that the equipment will be in working condition for 99 per cent of the time.

(2) The extent of verification may vary from 100 per cent at the beginning to around 20 per cent when the staff become more trained.

It is however suggested that the CSO should purchase only 24 additional workstations and rent 24, when the census records are ready for keying. If all 48 workstations were purchased the CSO will have data entry facility far in excess of its needs after completion of the census work.

In order to cut down the response time, when using simultaneously all the workstations, it is necessary to extend for that period the main memory to its maximum (4 Megabytes), and to use the GESMASQUE software.

The proposal for upgrading the HP 3000/44 to cope with the census work is given in Annex II.

Annex III gives the part of equipment that is to be purchased in 1983 Megabyte of core storage, 24 workstations, 2 terminal controllers and at the request of CSO, one removable disk drive (75 Mb) and five diskpcks. The Hardware cost is US \$ 152,060. The freight and handling charges are approximately 3 per cent of the hardware cost: US \$4,562.

Because of voltage fluctuations at CSO (3) it is recommended to acquire also a 16 KVA voltage regulator. The CSO uses presently one 7.5 KVA IMUNELEC voltage regulator which will not be sufficient after upgrading the system. A 16 KVA voltage regulator costs approximately US \$6,000 according to SERIC, which proposes the model "ENERGIE".

The computer should be ordered from SERIC (4) which is the exclusive distributor for Ethiopia, and from which the present system is purchased. According to the Government any other arrangement may lead to unnecessary complications and delays and probably extra cost in the long run.

If the order is placed now, the equipment will be ready by June 1983 to process the pilot census planned for April 1983.

The Government is willing to take care of the monthly maintenance charges of this equipment which represent approximately 1 per cent of the hardware cost.

Annex IV gives details of the equipment that can be rented. The purchase price is US \$131,129 but it can be rented for US \$6,000 per month. Since this part will be used for the main census planned for December 1983, the renting period will start from January 1984 for approximately 15 months. The cost will be therefore US \$72,000 for 1984 and US \$18,000 for 1985.

-
- (3) According to measurements carried out by NCR, the voltage at CSO is not a constant 220 volts, but suggests and says from 0 to 242 volts with impulses as high as 1,512 volts.
 - (4) SERIC 16-18 rue Barbes, 92126 Montrouge CEDEX, France, Phone 6571300, Telex 203.034, SERIC has a local representation for servicing called "SERIC ETHIOPIA".

6. The Site

The present data entry room is not large enough to house all the workstations required for the census.

Another room adjacent to the present one is therefore required for the extra workstations. The cost of site preparation will be borne by the Government.

7. Data processing expertise

Although there are some programmers at CSO (5), it is not considered that the persons at present in the data processing section of the CSO have the necessary experience and expertise to process the 1983 population census.

A census data processing expert should therefore be recruited internationally to join the project preferably in April 1983. A job description for this post (Annex V) has been prepared and forwarded to DTCD.

At the time of the mission, a data processing consultant, M. Mick Pearce (from WFS) was on a 2 month assignment at CSO, mainly to assist in processing an agricultural survey.

8. Software

It is expected to use the HP software VIEW and GESMASQUE for data entry.

COBOL-CONCOR will be used for data validation and error correction. There exists an HP version at ECA and the Regional Adviser on data processing will assist in installing it on the CSO computer. For tabulation, CENTS IV is preferred.

UNSO is requested to get from the US Bureau of the census in Washington DC a copy of the CENTS system and the corresponding documentation and forward it to the CSO preferably via ECA pouch.

(5) One CSO programmer, Mr. Teshome Teklu, has attended the COBOL-CONCOR workshop held at ECA, Addis Ababa, in January 1983.

ANNEX I

CENTRAL STATISTICS OFFICE
ADDIS ABABA, ETHIOPIA

Revised workload estimates for the 1983 population census.

Assumptions	Total	Rural	Urban	Remarks
Population size	35,000,000	30,450,000	4,550,000	87%-13%
Number of households	7,777,778	6,766,667	1,011,111	4,5 persons per household
Population 10 years and over	22,750,000	19,792,500	2,957,500	65%
Females 10 years and over	11,375,000	9,896,250	1,478,750	32.5%
Deaths per year	700,000	609,000	91,000	20%

A. RURAL FORM

Characteristics	Field size	Number of cases	Number of keystrokes
<u>Households</u>			
Identification	18	6,766,667	121,800,006
Household characteristics	10	6,766,667	67,666,670
<u>Individuals</u>			
Identification automatic reproduction	1	30,450,000	30,450,000
Characteristics common to all persons	14	30,450,000	426,300,000
Characteristics for persons 10 years and over	10	19,792,500	197,925,000
Females 10 years and over	10	9,896,250	98,962,500
Deaths	3	609,000	1,827,000
			<hr/> 944,931,176

B. URBAN FORM

Characteristics	Field size	N° of cases	N° of key depressions
<u>Households</u>			
Identification	21	1,011,111	21,233,331
Household characteristics	35	1,011,111	
<u>Individuals</u>			
Identification automatic reproduction	1	4,550,000	4,550,000
Characteristics common to all persons	18	4,550,000	81,900,000
Persons 10 years and over	7	2,957,500	20,702,500
Females 10 years and over	10	1,478,750	14,787,500
Deaths	3	91,000	273,000
			<hr/>
			143,446,331

Total Rural + Urban: 944,931,176 + 143,446,331 = 1,088,377,507

Assuming 50% verification: 1,088,337,507 x 1.5 = 1,632,506,261
+ 10% (miskeying, correction etc....) 163,250,626
1,795,756,887

Assuming 2 shifts of operation 35 working hours per week, a keying speed of 8,000 key depressions per hour and 48 workstations, the data entry work can be carried out in:

$$\frac{1,795,756,887}{2 \times 35 \times 8,000 \times 48} = 66.81$$

approximately 15 months

ANNEX II

PROPOSAL FOR UPGRADING THE HP 3000/44
SYSTEM TO COPE WITH THE CENSUS WORK:

	Cost in US \$
48 Terminals 2622A with cables	145,920
3 Mb CPU memory	60,000
4 Terminal controllers ATP	32,400
A.T.P. controllers	3,900
Second memory controller (for memory extension) (300 94A)	1,969
1 Removable disk drive of 75 Mb (7 925)	29,000
5 Diskpack 75 Mb each	10,000
	<hr/>
	283,189

Software: GESMASQUE

ANNEX III

PART OF EQUIPMENT TO BE PURCHASED FOR UPGRADING
THE CSO HP 3000 SERIES 44 COMPUTER

	Cost in US \$
24 Terminal Block Mode 2622A	72,960
1 Megabyte of core storage	20,000
2 Terminal controllers ATP	20,100
1 Removable disk drive 75 Mb	29,000
5 Diskpacks 75 Mb	10,000
	<hr/>
	152,060
+ 3% freight, insurance and handling charge	4,561.8
	<hr/>
	156,621.8

This equipment should be ordered from SERIC,

16-18 Rue Barbes

92126 Montrouge CEDEX FRANCE

Phone 6571300

Telex 203.034

which is the exclusive distributor for Ethiopia.

SERIC has a local representation for servicing called "SERIC Ethiopia"

ANNEX IV

PART OF EQUIPMENT
 THAT CAN BE RENTED
 (From January 1984)

	Cost of the Equip. in US \$
24 Terminals Block mode 2622 A	72,960
2 Mb of CPP memory	40,000
2 Terminal controllers ATP	16,200
⊙ Second memory controller (for memory extension) (30094A)	1,969
	<hr/> 131,129
<p>The cost for renting the above is 6,000 US \$ per month 6 month to be paid in advance</p>	
<p>Cost for 12 month rental $12 \times 6,000 = 72,000$</p>	
<p>Cost for 15 month: $15 \times 6,000 = 90,000$</p>	

ANNEX V

REQUEST FROM THE GOVERNMENT OF
ETHIOPIA

JOB DESCRIPTION

Post title: Census Data Processing Adviser

Duration: One year, with possibility of extension

Date required: April 1983

Duty Stations: Addis Ababa

Duties: Under the guidance of the Manager, CSO and in close co-operation with other United Nations experts and national counterparts, the data processing adviser will be expected to organise the census data processing activities and assist in all aspects of that work.

In particular, the adviser will:

1. Develop systems for the control of data flow, file creation, validation and error correction and for tabulation and subsequent access to the data;
2. Train operators to perform data entry and validation of census data on direct-data-entry devices;
3. Install at the CSO computer, software packages for the validation, error correction and tabulation of census data and train personnel of the Central Statistical Office in their use;
4. Write programmes for data editing, correction, imputation, tabulation and analysis using adequate software;
5. Perform any other function as necessary to assist in the effective processing of the census data.

Qualifications:

1. A University degree in Computer Science or advanced training in Computer Methods;
2. At least five years experience and expertise in statistical data processing with knowledge of census or survey microdata processing techniques;
3. Good knowledge of the main programming languages including COBOL and FORTRAN
4. Thorough familiarity and experience with Software packages such as CONCOR, COCENTS and CENTS IV
5. Familiarity with HP3000 Computer Systems

Language: English

Background information: The Government has decided to undertake a General Population Census in December 1983, the first of its kind in Ethiopia. Data will be collected through two types of questionnaires:

- one to be administered to the rural population (approximately 32 million persons in 7,800,000 households) and
- another to be filled for urban population (approximately 3 million persons). Data will be entered through workstations directly attached to the CSO computer: HP 3000/44 with tapes and disks. The United Nations Fund for Population Activities will provide assistance for the 1983 census mainly in the form of advisory services, equipment and supplies.