

610827158

UNITED NATIONS ECONOMIC COMMISSION FOR AFRICA

REPORT ON A MISSION TO THE  
REPUBLIC OF ZIMBABWE  
(4 - 11 November 1980)

by

Djamal E. Benzine  
Regional Adviser in Data Processing  
Statistics Division

December 1980,  
Addis Ababa

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

## 1. Introduction and Terms of Reference

The mission to the Republic of Zimbabwe was in response to several requests from the Government, to advise on the data processing component of a manpower survey to be undertaken early next year, and to assist in the redaction of a project document to be submitted to UNDP for financing.

In the course of the mission, I had discussions with the following persons to whom I wish to express my appreciation for their co-operation:

### Government of Zimbabwe

- Mr. Murerwa Permanent Secretary, Ministry of Manpower Planning and Development
- Mr. Mandaza Director, Ministry of Manpower Planning and Development
- Mr. Mothobe Director, Ministry of Manpower Planning and Development
- Mr. Mandishona Director of Central Statistical Office
- Mr. D. Stoltz Manager, Computer Bureau, Ministry of Finance (Treasury)

### UN/UNDP

- Mr. Onitiri Resident Representative
- Mr. M. Berrezoug Deputy Resident Representative
- Mr. Friscic UN/DTCD (on mission in Zimbabwe)
- Mr. Ziyambi Programme Officer

## 2. The Survey

The survey will be carried out by the Ministry of Manpower Planning with the assistance of the Central Statistical Office (CSO) and other government offices, agencies and bodies. The objectives of this survey are:

- the assessment of the current stock of skilled and semi-skilled manpower in Zimbabwe
- the identification of skilled and semi-skilled Zimbabweans potentially available from abroad
- the occupational profile for the skilled and semi-skilled workers
- the assessment of the need for education training and upgrading of manpower, especially for African workers.

More comprehensive and detailed information on the employment structure as a whole including the rate of employment and other relevant statistics will be provided by a population census the Government intends to carry out. However, it was felt that top priority should be given to a manpower survey, taking into account the more urgent need for data on skilled and semi-skilled manpower during this transitional period.

The survey will cover:

- (a) Establishments: - all public sector employment
- all private establishments with 100 employees or more
- a 5% sample of all other formal establishments

This might generate about 15,000 establishments in Zimbabwe to be included in the survey.

- (b) A representative sample of 100,000 skilled and semi-skilled workers will be drawn out of the total list of establishments.

The sampling scheme will be designed in collaboration with the Central Statistical Office.

Questionnaires on the establishments will be sent to the employers and will be followed by personal interviews. Data on individual characteristics concerning the employees selected in the sample will also be collected through questionnaires sent to them, followed by personal interviews. The field work is expected to take approximately 4 months with the following personnel:

- 12 Supervisors (one for each province)
- 40 Team leaders
- 300 Enumerators

The questionnaires have not been finalized yet but the following information is considered as the minimum necessary:

(i) for each establishment: name of firm, economic sector, region, locality, total employment by sex, race and wage bill. This data is to be classified in a way which allows them to be tied in with CSO records for national accounts and employment returns.

(ii) for each worker selected in the sample: firm, sector, nationality, race, sex, age, job experience in years since completing formal education, training, education possessed (measured in years and qualifications), training possessed (measured in months, years and qualifications), earnings (including cash and non-cash payments).

A pilot survey will be carried out to test the questionnaire and the definitions, particularly the terms "skilled" and "semi-skilled".

### 3. General Recommendations on Processing the Survey

The processing of the data should be considered as an important feature of the survey plan, and therefore the work on data processing should begin early, as soon as the tabulations and codes have been drawn up. Codes should be drawn up in consultation between statisticians and data processing staff, and be based on the tabulation programme and on a realistic assessment on what further tabulations are likely to be required, if needed for more detailed analysis.

If there is sufficient knowledge to set up precoded boxes on the questionnaires that can be keyed directly, for closed and semi-closed questions, this will cut down the time required for coding. Besides the preliminary work mentioned above, the data processing operations will include the following activities:

#### Receipt and control

The data processing unit should have a list of the units to be interviewed, in order to maintain a basic control over the field work to ensure that all the questionnaires are received for all the survey units.

Non interview adjustment

One way of adjustment for non-interviews - if any - of eligible units, would consist in replacing them by similar one selected at random within the same sample segment. The interviewed unit replacing a non-interview unit will have therefore to be keyed twice during the data entry phase.

Manual editing

The purpose of editing is to detect and correct errors. In the case of conflicting data, it is often very difficult to tell which of the conflicting entries contains the error. This is why the best place for editing is in the field: one can always return to reinterview the informant and correct the error with certainty. Besides a manual edit to be carried out in the field by the enumerators themselves and by their supervisors, further manual edit will be performed in office, especially for important topics.

Coding

This operation will consist in recording numbers by checking precoded boxes, and in assigning codes to informations collected "in clear" on the field, such as occupation and industry. Coding should be done on the questionnaires themselves - which have to be designed accordingly - instead of on a separate coding sheet. It will be faster and more accurate. Because of the complexity of coding certain items, verification of codes will have to be carried out on a 100% basis. It is important to establish production standards for the coding operations (number of questionnaires to be coded per day and the number of permissible errors). On the coding lists, there should be indications on how to treat "not stated" and "not applicable" to prevent the two categories from being grouped together.

Data keying and verification

The data entry and some editing (validation) will be carried out "on line" on the equipment of the Ministry of Manpower Planning and Development: a system INFOREX 5000 with 64 K. bytes of core, 4.9 Megabytes discs, 5 VDU 1 tape and 1 printer. Verification will be done on a 100% basis. Data will be recorded on tapes via disc. The questionnaires have not been finalized yet, but on the basis of an expected count of 15,000 establishments and 100,000 personal records, a very rough estimate of the data entry workload could be:

Data Entry Workload

<u>Personal Records</u>	<u>Number of Characters</u>	<u>Keystrokes</u>
100,000	80	8,000,000
<u>Establishment Records</u>		
15,000	200	3,000,000
Total		11,000,000
	+10%	1,100,000
		12,100,000
<u>Verification</u>	100%	12,100,000
		24,200,000

Assuming an average speed of 8,000 keystrokes per hour and 35 working hours per week, 24,200,000 keystrokes x  $\frac{\text{hour}}{8,000}$  x  $\frac{\text{week}}{35}$  = 86 machine and operator weeks.

With 4 machines, the work will be done in 21.5 weeks with 1 shift or 11 weeks with 2 shifts (8 operators).

#### Computer edit

In addition to the manual edits to be carried out on the field and in office, and the machine editing to be performed "on line" during the data entry phase (to detect omissions of required entries, and impossible entries), it is essential that a complete computer edit on the tapes produced, be performed. A software package can be used for this purpose and I recommend COBOL-CONCOR, distributed (free of charge) by the International Statistical Programmes Centre, US Bureau of the Census. A request for a version of COBOL-CONCOR can be sent through United Nations Statistical Office (UNCO) or directly to the US Bureau of the Census.

Inter-record checking and detection of inconsistencies between related items will be the main objective of the computer editing. This kind of checking cannot be carried out on the INFOREX 5000 during the data entry stage. Corrections can be done either automatically by computer or manually by referring to the questionnaires.

Clear edit specifications should therefore be prepared by the manpower expert in collaboration with the data processing staff. If automatic corrections are to be done by computer, clear imputations will have to be programmed. The errors detected and corrected should be documented in such a way to keep the users aware of the quality of the data they will be working with.

#### Weighting

A basic weight (the reciprocal of the sampling ratio) will have to be entered in each record, so that the tables can be computed directly for the whole population under study.

#### Recoding and file restructuration

The next step will be recoding. It is an operation consisting in creating new variables by grouping or combining values. Reducing long distribution of values into a shorter number of categories is always necessary for variables such as age, profession or economic activity (industry), if the original codes are given in detail. During this step, it is recommended to reorganize the data by creating sub-files for tabulation by topics.

#### Tabulation

The final stage in data processing will be the tabulation. The required tables will be produced on the edited and restructured files. Tabulation packages such as COCENTS, XTALLY, RGCP will reduce the programming time. Except for data entry, all the data processing operations will be performed on the equipment of the Ministry of Finance (Treasury). The configurations are given in Annex.

The editing package (COBOL-CONCOR) and one of the tabulation packages will have to be installed at the computer centre for processing the survey. COCENTS developed by the US Bureau of the Census and XTALLY, produced and distributed by UNSO have been used for a long time by Statistical Offices to process censuses and statistical surveys. Both packages are distributed free of charge. RGSP (Rothamsted General Survey Programme) developed at Rothamsted Experimental Station 1/ in United Kingdom is available only on an annual licence fee and the signing of an agreement. COCENTS can be installed on any machine with a COBOL compiler and XTALLY requires an RPG II compiler. RGSP requires at least 128 K. bytes of core and FORTRAN-H compiler.

#### Data processing expertise

A data processing expert is needed for the project. The only data processing staff at the Ministry of Manpower Planning and Development is the personnel working on the INFOREX 5000 system (4 operators and one supervisor). Their experience is limited to data entry and file management on that equipment. The expert is therefore expected, in co-operation with the manpower expert, to assist the Ministry of Manpower Planning and Development in all the data processing activities in connexion with the survey including:

- Assisting in the questionnaire design to insure its processibility
- Training and supervising coders and operators to perform data entry and validation of survey data on the INFOREX 5000 system
- Selecting appropriate software packages for data editing and tabulation and assisting in their installation at the Computer Centre of the treasury
- Fulfill the duties of systems analyst/programmer as required.

The services of the expert will be needed for at least 6 months, but it is recommended that he joins the project as soon as possible and stays in post until the data is completely processed.

---

1/ Rothamsted Experimental Station  
Harpenden  
Hertfordshire AL5 2JQ  
ENGLAND

ANNEX

Data processing equipment at the Computer Centre  
of the Ministry of Finance (Treasury)

CONFIGURATION (1)

Device type

CPU	Data general eclipse C/350 (installed Dec. 1979) 1 megabyte
Discs	2 units 192 megabytes each
Printer	1 printer
Tape	1 tape drive
Terminals	29 terminals (VDU and printers) throughout the country

CONFIGURATION (2)

Device type

CPU	NCR 615 (installed 1972(?)) 380 kilobytes
Discs	4 - 658 discs - 200 megabytes each 2 - 657 discs - 30 megabytes each 6 - 655 discs - 5 megabytes each
Tape drives	3 - 9 tracks 1600 BPI 2 - 7 tracks 200,556,800 BPI phase mode
Printers	2 - model 640 1500/3000 lines per minute 1 - model 646

Data entry equipment:

There are 2 systems MDS 1200 with 14 key stations