F0 007 909 543

REPUBLIC OF BOTSWANA

1995/6 LABOUR FORCE SURVEY

TECHNICAL & OPERATIONAL REPORT

Price P10.00

CENTRAL STATISTICS OFFICE

Printed by the Department of Printing and Publishing Services, Gaborone

1995/6 LABOUR FORCE SURVEY TECHNICAL REPORT

Published by

Central Statistics Office
Private Bag 0024, Gaborone
Phone: 352200 Fax 352201

Contact Statistician: K Makhumalo
Labour Statistics Unit (Old Industrial Site)
Phone 306409 Fax 306235

Printed by and obtainable from
Department of Printing & Publishing Services
March 1998

COPYRIGHT RESERVED

Extracts may be published if
source is duly acknowledged.

PREFACE

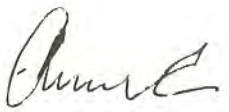
This is the second survey of this nature to have been conducted in Botswana. The first Labour Force Survey (LFS) to be conducted was in 1984/5, the results of which were widely used by the government in its policy formulation, the private sector and the international community.

The main objective of the technical report for the 1995/6 LFS is to give the users an insight into the problems encountered and also apprise them on the current state of the methodological and conceptual problems encountered in the designing, processing and analysing the results of this important survey. It is also intended to give them an overview of the expected results of this survey. Furthermore, it is intended to enlighten the users about the concepts of labour force and their applicability to this country. The text is accompanied by charts and tables which will help the users in the understanding of the labour force characteristics.

The preparation of the main report is at an advanced stage, with most of the tables compiled, but it may take sometime to have the main report to be finalised. However, in the interim, a Statistical Brief based on this survey would be released. This will enable interested analysts to have access to the data set.

This exercise was one of the most demanding and challenging to have been undertaken by CSO. It started in late August 1995 and was completed in early September 1996. It involved many people, who helped one way or another in making this exercise a success.

I would therefore like to take this opportunity to express appreciation to all those who participated in this exercise. To mention but a few, the general public who provided the necessary information; interviewers, who tirelessly worked throughout the entire period; district and town/city council officials; and all those whom I have not mentioned but whose input led to the successful completion of the survey.



G M Charumbira

GOVERNMENT STAISTICIAN

CONTENTS

1. BACKGROUND AND SURVEY OBJECTIVES 1

1.1 BACKGROUND 1
1.2 SURVEY OBJECTIVES 1

2. MAJOR CONCEPTS AND DEFINITIONS 1

2.1 SCOPE 1

2.1 COVERAGE RULE 2
2.2 HOUSEHOLD AND HEAD OF HOUSEHOLD DEFINITION 3
2.3 SURVEY PERIOD 4

3. MAJOR ECONOMIC ACTIVITY CONCEPTS 4

3.1 ECONOMICALLY ACTIVE POPULATION 4
3.2 CURRENT ECONOMIC ACTIVITY 5
3.3 USUAL ECONOMIC ACTIVITY 7

4. MAJOR CLASSIFICATIONS USED 9

4.1 STATUS OF EMPLOYMENT 9
4.2 SECTOR OF EMPLOYMENT 10
4.3 OCCUPATION OF EMPLOYMENT 11
4.4 INDUSTRY OF EMPLOYMENT 11
4.5 SUBJECT OF TRAINING 12
4.6 IDENTIFICATION OF THE INFORMAL SECTOR 12

5. SAMPLE DESIGN 13

5.1 STRATIFICATION: 13
5.2 CREATION OF SAMPLE BLOCKS 14
5.3 ALLOCATION OF SAMPLE BLOCKS TO STRATA 14
5.4 FIELD DELINEATION OF BLOCKS 14
5.5 DWELLING SELECTION WITHIN BLOCKS 15

6. QUESTIONNAIRE DESIGN 15

7. PRETEST 16

8. MAIN SURVEY OPERATORS 16

8.1 COLLECTION SPREAD 16
8.2 FIELD INSPECTIONS 17
8.3 SAMPLE LOSS 17

9. DATA PROCESSING 18

9.1 MANUAL EDITING AND CODING 18
9.2 DATA ENTRY 19
9.3 DATA CONTROL 19
9.4 COMPUTER EDITING 19
9.5 DERIVED ITEMS 19
9.6 FREQUENCIES AND CHECK TABLES 19
9.7 DATA EDITING 20
9.8 TABULATIONS 20

10. SAMPLE AND WEIGHTING PROBLEMS AND ADJUSTMENTS 20

10.1 SAMPLE SIZE AND STRATUM ALLOCATION 20
10.2 BLOCK SELECTION AND MEASURES OF SIZE (MoS) 20
10.3 DEMARCATION OF BLOCKS: 21
10.4 LISTONG OF SELECTED BLOCKS 21
10.5 WEIGHTING 22
10.6 BENCHMARK ADJUSTMENTS 23
10.7 HOUSEHOLD DATA 24

10.8 QUARTERLY ESTIMATES 24

11. DATA EVALUATION 24

11.1 TOTAL ECONOMIC ACTIVITY MEASURES 24
11.2 TRAINING CLASSIFICATIONS 27
11.3 NON-CITIZEN ESTIMATES 28
11.4 PERSONS TEMPORARILY ABSENT FROM WORK 29
11.5 UNDEREMPLOYMENT MEASURES 30
11.6 INCOME DATA 31
11.7 INFORMAL SECTOR MEASURES 32
11.8 HOUSEHOLD DATA 33
11.9 QUARTERLY ESTIMATES 34

1. BACKGROUND AND SURVEY OBJECTIVES

1.1 Background

The last Labour Force Survey conducted in Botswana by the Central Statistics Office (CSO) was in 1984/5, over ten years ago. Thereafter, it was decided that there was a need to conduct a regular Labour Force Survey(LFS) as part of CSOs on-going household survey program. The 1995/6 LFS would be a mid-census period update on the economic activity situation of the population. The LFS would however provide much more in-depth information than a census. In future the CSO plans to conduct LFSs at five or six years intervals.

One of the problems about the long gap between LFSs was that the CSO had limited capacity in terms practical experience in the conduct, processing and analysis of LFSs. Labour Force Surveys are not easy surveys since they require good theoretical as well as practical knowledge of their particular requirements. The lack of local expertise was noticeable in all phases of the survey, particularly the training. It is hoped and expected that some of the staff who worked on this LFS will be able to apply their experience on future LFSs if they are more frequent.

1.2 Survey Objectives

The broad objective of the survey was to obtain comprehensive data on the status of the Labour Market prevailing in Botswana.

More detailed objectives were:-

* To provide measures of both current and usual economic activity.

* To obtain a measure of the size of employment in both formal and informal sectors.

* To provide measures of unemployment and underemployment.

The aim of this technical report is to provide guidance to users of the current data and for persons concerned with development of future Labour Force Surveys.

2. MAJOR CONCEPTS AND DEFINITIONS

2.1 Scope

Scope refers to the boundary within the universe of the total population of Botswana which was to be covered by the survey.

Geographic Scope

In early discussions of the Technical Steering Committee, it was proposed to cover almost all the population. However it was established that the master sample of Enumeration Areas (EAs) of CSO only covered private dwellings. There was no sample or sampling frame available for the population in institutional dwellings and no resources were available to build one. Thus the survey had to exclude most of the institutional population. It was possible to include the staff of many institutions but the following were excluded:

- BDF staff living in Barracks
- Guests at hotels and lodges
- Inmates of Prisons

Diplomats

Foreign diplomats resident in Botswana were excluded from the survey.

Non-Citizens

In some countries non-citizens are excluded from LFSs as their size is small and they are often hard to contact. Notwithstanding that, in Botswana the non-citizen population is relatively large and important and it was decided to include them. However non-citizens are often clustered in certain areas of the major towns and the sample was not designed to accurately cover this clustered population.

Age Limit

There was considerable discussion on what age limit to take for the survey. No age groups were totally excluded from the survey but two age restrictions were taken for parts of the questionnaire for this survey. For those less than five years only questions on relationship, age, sex and citizenship were asked. For those aged five to eleven years, school attendance and level only were also asked. Often, for Labour Force Surveys, a lower limit of 15 years or 10 years is taken for the main economic activity questions. For this survey it was decided to take 12 years to conform with the 1991 Population Census. For persons 12 years and over the full range of demographic and economic questions were asked when applicable. In future it should be considered whether to lower the age limit for the economic questions to 10 years to explore the incidence of child labour which is a topic of interest in many countries now. Note that with a limit of 10 years, there is no problem to produce most tables for 12 years and above for census comparability.

2.1 Coverage Rule

The usual basic decision on a coverage rule is whether to adopt a de-facto or de-jure approach. The de-facto approach is to include anyone who slept in the selected household the night before the interview. The de-jure approach is to include persons who usually live in the selected household. The de-facto approach is more precise for fieldwork and commonly used in population censuses. However it is difficult to use household economic data from such an approach. For example, temporarily absent wage earners are excluded from households by such an approach. A de-jure approach is often used for Labour Force Surveys. Botswana has had a continuing problem in surveys of missing persons, particularly males, in the younger adult ages. By missing is meant these age groups are considerably under-represented compared to the census figures and projections

based on the census. It is thought that the coverage rules in the past and a general lack of attention to basic coverage has contributed to this problem.

It was decided to conduct a test for this survey and ask questions about absent usual residents and anyone who slept in the selected dwelling on the previous night i.e. recording both groups. Respondents were recorded in the following categories on the questionnaire:

- Usual Resident present (42,756)
- Usual Resident absent (957)
- Visitor 14 Days and over (654)
- Visitor less than 14 days (620)

The number of records recorded in each group is given in brackets. The aim was for interviewers to probe strongly for the absentee category and the visitor categories. It was hoped to include mobile young men in towns who are hard to record, in the visitors 14 days and over group. In general the high proportion (95 percent) of usual residents present suggests that interviewers did not probe well enough for the other groups. The manual and probably the training, did not emphasise these points. As will be discussed in the evaluation section the average household occupancy was less than that for the 1991 population census.

In theory the usual residents absent should be visitors in other households, temporary guests in hotels, hospitals etc. or out of the country. Tabulations showed that over 50 percent of the residents absent were young students (who either stay at boarding schools or stay with relatives or friends near schools during school terms). This is partly in line with expectations. Visitors in both groups were a mixture of economically active persons and a large proportions of students. To avoid double counting, the students staying with other private households who were recorded as temporarily absent, should be deleted from the visitor count. However they seem to be in both visitor groups and we cannot distinguish them from other students on routine visits with their families. The whole issue obviously still needs more testing and research, and visitors divided if possible into those who would be recorded at their homes and those who would not. Possibly the days limit could be increased to 30 days in future, and those with a permanent residence elsewhere and with someone living there, always included in the less than 30 days group. However care should be taken not to make the coverage rules too complex.

For this survey it was decided to include the first three groups above i.e. all usual residents plus the longer term visitors. There is some possible small double count by this method but as noted most would be students, who are rarely economically active.

2.2 Household and Head of Household Definition

A Household is a group of one or more persons, related or unrelated, living together under the same roof in the same lolwapa "eating from the same pot" or making common provision for food and other living arrangements.

The head of household is any person, male or female, at least 12 years of age, who is regarded by other members of the household as their head. This person may or may not be a blood relative to other members of the household. Such a person should have either spent the last night at the household or returning within 14 days.

A domestic servant was recorded as a separate household.

2.3 Survey Period

The LFS started in late August 1995 and concluded in early September 1996. There was a three-week break over the Xmas/New year period to avoid unusual economic activity in that period. The spread of a LFS over approximately a year has operational advantages. It allows a lower workload in any one month and also allows some corrections to be fed back to the field as they are identified in the early months. Several field circulars were issued in an attempt to correct common errors which were being identified. As noted earlier, the training and manual had been weak on many aspects. However, as is often the case, the field circulars were only partly effective in stopping the common errors.

There are problems of interpretation of a reference date for data spread over a year. The current activity data does not refer to any fixed time during the year and it is definitely wrong to say it refers to the middle of the survey year. It should be taken as a measure of the average economic activity throughout the twelve months. 'On' and 'off' agricultural seasons are represented in the data.

As explained in the sampling section, the sample was distributed approximately equally to the four quarters of the survey 12 months. It is thus possible to produce survey estimates for each of the four quarters. However sample numbers are relatively small in each quarter and there is concern particularly about the urban sample distribution by quarter. It was also not possible to weight the quarters fully independently by quarter (see weighing section), and only rates and percentages can be produced on a quarterly basis. For the main survey report only major aggregate rates have been produced on a quarterly basis and further analysis by quarter should be done with caution.

3. Major Economic Activity Concepts

The conceptual basis of a Labour Force Survey is extremely important. The international recommendations concerning the economically active population as adopted by the 15th International Conference of Labour Statisticians (ICLS) are designed to ensure that LFS results are compatible with the overall framework of economic statistics. ILO Bureau of Statistics assisted with some comments during the design stage which were very useful. The comments below centre on the basic concepts adopted and the variations which were decided on for Botswana (which are allowable under the international recommendations).

3.1 Economically Active Population

This is the key basic definition for the survey and, in general, it was decided to adopt the definition internationally recommended. Economically active persons are those who supply labour for the production of goods and services for the market, barter or for home consumption as defined by the United Nations Systems of National Accounts. The production boundary is very wide under this system and includes paid employment and a wide range of self employed activities but excludes unpaid domestic activities such as minding children, cooking food for own family etc. Common

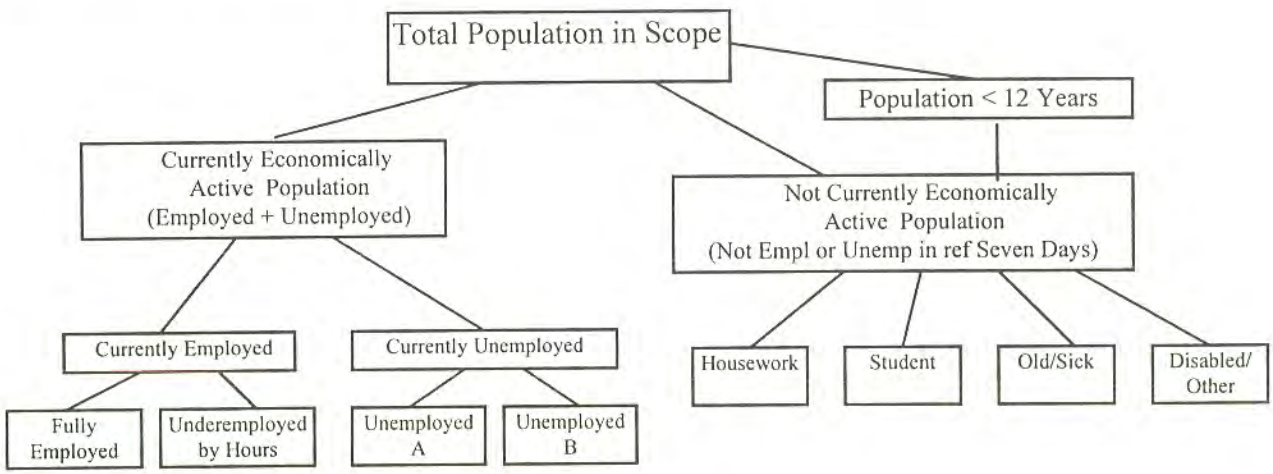
household duties such as fetching water and collecting firewood were excluded, as it was thought that most persons, particularly women would be included in the employed if these activities were included. Very few respondents would then be considered for the unemployed questions. The list of employed activities was drawn up and produced on page 1 of the questionnaire to act as a prompt list for both Question 3 (Usual Activity) and Question 8(Current Activity). The effect of this wide scope of economic activities is to include most of the adult population as economically active (employed or unemployed). The list of economic activities were discussed in the technical working group for the survey.

3.2 Current Economic Activity

International recommendations allow the use of any recent short period e.g. day, week. For the LFS it was decided to use the previous seven days from the date of the survey interview.

The overall current activity model adopted for current economic activity measures is shown diagrammatically below.

Diagram 1- Current Economic Activity Components of the Total Population



Currently Economically Active (The Labour Force) - Persons who were either employed or unemployed, as defined below, in the reference week.

- Employed Persons - Persons who did some work in the reference period either for payment in cash or kind (paid employees) or who were in self employment for profit or family gain, plus persons temporarily absent from these activities but definitely going to return to them (e.g. on leave or sick). Self-employment includes the large number of persons working on their own lands or cattleposts. Unpaid family workers in family businesses are included. Some work was defined as 1 hour or more in the reference seven days. It should be noted that any economic work took priority over all other activities.
- Unemployed Persons - Persons who were not employed as defined above and who stated that they were available for work.

- Currently Not Economically Active (Not in the Labour Force) - Persons who were neither employed nor unemployed in the reference period. This includes persons doing solely unpaid domestic work in their own house, studying or not working because they were sick or retired. Note that persons with these activities are included as economically active if they did any economic activity as defined above, in the reference period.

Underemployed persons by hours are persons who worked less than 35 hours in the reference week for an economic reason (i.e. excluding those on leave, at school and similar non-economic reasons) and who said they were available for more work.

Unemployed A are persons currently unemployed who were not only available for work but had taken active steps to find work in the last 30 days.

Unemployed B are persons currently unemployed who have not taken active steps to find work in the last 30 days (discouraged job seekers).

The following issues were particularly discussed during survey development:

Unemployment - It was decided not to restrict the unemployed to those both without work and actively seeking work. The Botswana Labour Market is largely unorganised. There are few labour exchanges and particularly in rural areas, it is often a meaningless activity to be actively seeking work. It was decided to use the broader and more applicable (and understandable) concept of being available for work. This relaxation of international recommendations is allowed. However it was decided to ask additional questions and produce the sub-sets of the unemployed as given above i.e. A - available and actively seeking, B - available but not actively seeking. Concept A is mostly applicable in urban areas. It was decided to use four weeks as the limit for actively seeking as it is a reasonably short period but also allows more time, given the state of the labour market. As will be seen in the tabulations, just over 50 percent of persons categorised as Unemployed B said they were not actively looking for work as they thought none was available. The Unemployed B rate is thus mostly a measure of discouraged job seekers in the Botswana imperfect labour market.

Underemployment - The rate given in the diagram above and in most tables is what is often called the 'Visible underemployment rate'. It covers persons working less than 35 hours who said they were available to work more hours. The 35 hours limit was taken as it is close to the normal working hours of most government employees. See section 10 E for a discussion of other measures of underemployment which were tried experimentally in this survey.

Temporary absences from employment - This covers persons not working in the reference period but having work from which they are temporarily absent. This is a much discussed issue internationally. Such a person is included in the employed population provided they keep a strong attachment to the job and there is a definite expectation he/she will return to it. The most common categories are persons on leave from a wage job.

For self employed the concept is more difficult. Formal leave is rare for the self employed, particularly small operators. The business or agriculture must be continuing in the person's absence e.g. spouse or relative looking after it. A limit for this group was taken of one month, as there are cases of a person periodically or seasonally engaging in business or agricultural activities and they could confuse this category if a longer period was allowed. Unpaid helpers and casual workers cannot be temporarily absent from work as they have no measurable attachment to the job. It was

appropriate to call such persons not economically active or unemployed as appropriate during such absences.

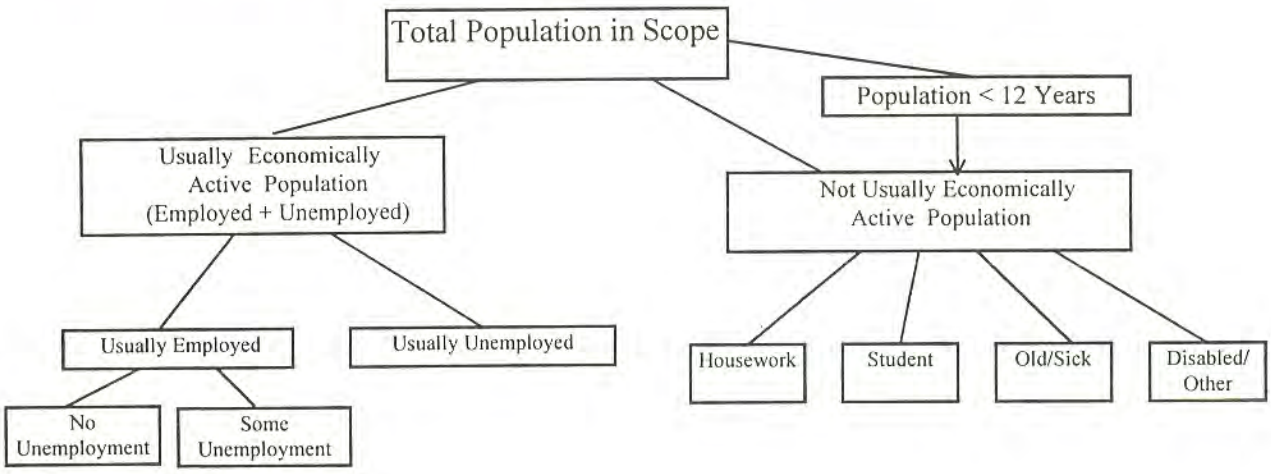
Persons temporarily absent from employment will have zero hours for current hours worked. As noted later in the data evaluation section, the number recorded as temporarily absent from work generally is reasonable.

3.3 Usual Economic Activity

One of the major early decisions in the survey design was to measure usual economic activity as well as current economic activity, as internationally recommended. It was thought that these usual measures would be particularly useful for seasonal activities and this has proved correct. It was decided to use the 12 calendar months up to the end of the calendar month preceding the date of the interview for the survey, as the reference period. Any usual economic activity was one that took more than one hour a week or half a day a month.

The overall usual activity model is shown below with summary explanations following.

Diagram 2- Usually Economically Active Components of the Total Population



Usually Economically Active - Persons who either worked or were available for work for 6 or more of the 12 reference months.

Usually Employed - Persons who were usually economically active and spent half or more of the economically active months, working.

Usually Unemployed - Persons who were usually economically active and spent more than half the economically active months, not working and available for work.

Usually Not Economically Active - Persons who did not work or were not available for work for more than 6 months of the 12 months.

Persons with Some Unemployment - Persons who stated that they were not working but were available for work at least part of a month. All the 'usually unemployed' are in this category.

Some of the 'usually not economically active' also fall in this category but it has been noted that most were young students who were recorded as available for work in school vacations. The usually not economically active with some unemployment have not been shown in the tabulations for this publication. Note that only 'Available' was used as a qualification for these measures and not 'Actively seeking work'. They are thus equivalent to the A plus B rates for current activity.

The derivation of the categories above was quite complex, but was done on the computer once data were fully edited. The key information is the totals in each category from Q3 of the questionnaire. The categories are.

A - no. of months working all month
(Q4 YES was made 12 for A)

B - no. of months working part month and available for work

C - no. of months working part month and not available for work

D - no. of months not working at all and available for work

E - no. of months not working at all and not available for work.

We assumed:-

(i) That part months on average were half a month. (This has been done in other countries)

(ii) That when a person was working part of a month and available for work part of a month, this was half a month each. (See B).

(iii) That when a person is not working and is available for work, it is the whole month (See D).

The major derivations are:-

* Persons Usually Economically Active
$A + D + B + C/2 > \text{or} = 6$

* Person Usually Employed
Economically Active and
$A + (B + C)/2 > \text{or} = D + B/2$

* Persons Usually Unemployed
Economically active and
$A + (B + C)/2 < D + B/2.$

* Persons Usually not Economically Active
$E + C/2 > 6$

The other derivation shown in the diagram is

* Usually Employed with some Unemployment
Usually employed and
B or $D > 0$

Other useful concepts which were derived and tabulated are.

* Persons Employed at all during 12 months
  Economically Active or not Active and
  A, B or C > 0
* Persons Unemployed at all during 12 months
  Economically Active or not Active and
  B or D > 0

4. Major Classifications used

The following major classifications were applied for the LFS:-

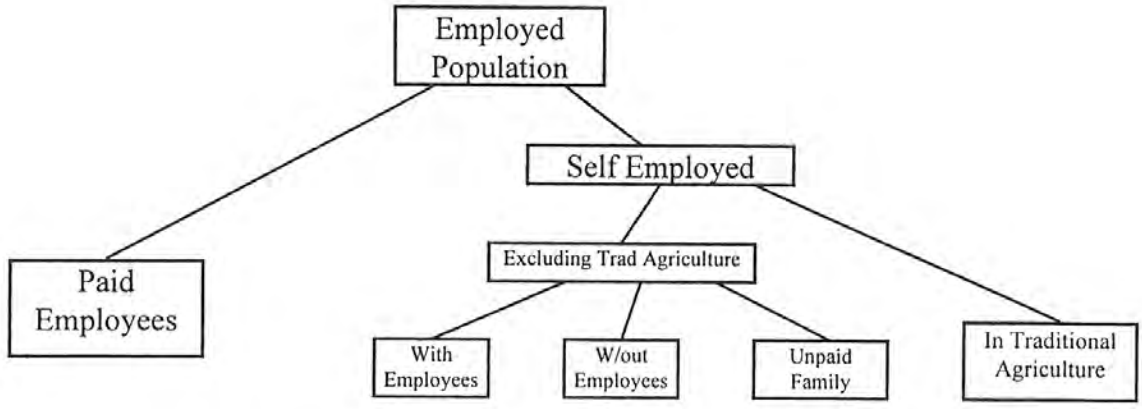
Status in Employment
Sector of Employment
Occupation of Employment
Industry of Employment
Subject of Training

The first four classifications were applied for each currently employed person for their main activity and secondary activity (if any). Sector and Status were also applied for usual activity. Occupation coding was also applied for the previous job of the unemployed. The subject of training classification was applied for all persons with any formal training of at least one month, for employed persons with any on-the-job training and for any training received while working outside the country (for those who have worked outside, mostly in RSA).

Each classification is explained below.

4.1 Status of Employment

A modified status in employment coding was used for the LFS as shown in the diagram and explained below.



- **Paid Employees** are persons who during the reference week performed some work for a wage or salary in cash or kind or were temporarily absent from such work. Permanent, temporary and casual paid employees were also included.

- **Self Employed** (Non Traditional Agriculture) are persons who during the reference week performed some work for profit or family gain (not on their own lands or cattlepost) or were temporarily absent from such work. This includes small and large businessmen working on their own enterprises. The category can be sub-divided into self employment with or without employees.

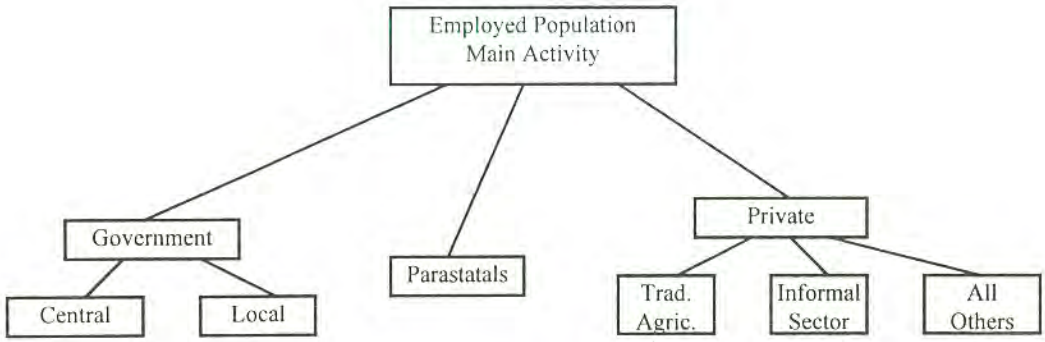
- **Unpaid Family Helpers** (Non Traditional Agriculture) are persons working completely without payments in cash or kind in family enterprises. Such workers on their own lands or cattlepost are not included here.

- **Traditional Agriculture Workers** are persons working on their own lands or cattlepost either in agriculture, livestock or fishing as either self employed persons or as unpaid family helpers.

The major variation of the commonly used international classification of status in employment, is the provision of the special category for traditional agriculture workers. These are often combined with the self employed. However it was decided that the separate category was very desirable for the LFS as the non agricultural self employed can easily be lost in the dominant agricultural group.

4.2 Sector of Employment

A special coding of sector of employment was used for the LFS as shown in the diagram and explained below.



- **Central Government** - Refers to persons working for all Central Government. This includes all Teachers in Government Primary and Secondary schools (including Government-aided and Community Junior) plus food stampers and night watchmen in schools.

- **Local Government** - Includes persons working for District and Urban Councils. There are a few pre-schools under local Government. Note a large number of persons employed on drought relief projects are included under Local Government.

- **Parastatals** - includes the wide range of enterprises owned at least 50 percent by the government, either profit making or non-profit making.

- **Traditional Agriculture** - includes all persons who stated that they were working on their own or family lands or cattlepost either in agriculture, livestock or fishing or as paid employees on similar establishments. Note that paid employees are quite significant.

- **Informal Sector(IS)** - an attempt was made to identify this group of policy interest through a series of questions. The definition is explained in Section 4.6 below. Note that persons employed by private households, not for business purposes, are included in the IS (i.e. Maids, Gardeners and Watchmen), but separately identified.

- **Other Private** - includes persons working for other enterprises not included above. This is broadly often called the private formal sector. It should be noted that persons working for NGOs are also included in this group.

The major innovation in the sector classification was the identification of the informal sector. This classification has proved extremely useful and is widely used in analysis. For usual activity, the informal sector group is not available and has been combined with 'other private'

4.3 Occupation of Employment

BOSCO (Botswana Standard Classification of Occupation) was applied. Copies of this classification are available from the Labour Statistics Unit.

4.4 Industry of Employment

A Botswana adaptation of the International Standard Industrial Classification (BISIC) was applied. Additional codes were particularly added to identify informal sector activities. Copies of the

classification are available from the Labour Statistics Unit. Note that this classification differs slightly from the one used for the Establishment and Enterprise Register (EER). Informal Sector and domestic employees for example are not included on the EER.

4.5 Subject of Training

A special subject of training classification was applied. It is a Botswana adaptation of a similar classification used for the Tanzanian LFS, which in turn was based on an earlier Botswana version. Copies are available from the Labour Statistics Unit.

4.6 Identification of the Informal Sector

There is considerable policy interest in the Informal Sector (IS) in Botswana. Users expressed a very strong interest in the LFS providing separate employment estimates for the IS. There is also a proposal for a separate special informal sector survey, which may be held in 1998.

The definition of the Informal Sector as adopted at the 1993 ILO International Conference of Labour Statisticians (ICLS) is relatively complex but a reasonably practical adaptation is possible and was applied for this LFS. The key points to the ICLS definition are that the informal activity should be part of the 'Household sector' in National Accounts terminology. This means that the activity should not be conducted as a registered company or such that a reasonably complete set of accounts are kept (quasi corporations). Exclusion of agriculture and employment over a particular limit are allowed by the international recommendation and were introduced.

In practice the definition was largely applied by the questionnaire structure as explained below:-

1. All persons working on own or family lands or cattleposts were excluded (Q21 & Q34)

2. Persons employed in the Government (Central or Local), Parastatals or NGOs were excluded (Q24 & Q37).

3. Domestic servants and other private households workers were separately identified in Q24 & Q37 and have been tabulated with the IS but separately identified.

4. Paid employees in Traditional agriculture were also included in Code 5 of Qs24 & 37 but were identifiable by their industry code and excluded from the IS.

5. Enterprises with total employment of more than 10 people were excluded (Q25 & Q 38).

6. Remaining persons were asked if the enterprise was a registered company or with a professional association, and whether the enterprise kept a complete set of accounts. A 'Don't know' category was allowed as employees often did not know this information. These two questions were followed by an interviewer evaluation where they were to assess the enterprise from the earlier answers and any supplementary questions they needed to ask. A guide box was given on the questionnaire. The location of the enterprise was particularly used in this evaluation (Q27 & Q40)

The evaluations and the reasons for it were to be checked by supervisors in the field and in the office. In general the procedures seemed to work reasonably well in practice. It obviously worked better for the self employed who were answering about their own enterprise than for employees.

During processing the IS was restricted to those enterprises with less than five paid employees. This restriction was to be compatible with the employment size groups on the EER, the base for the 'formal sector' employment surveys. It is also thought to be a reasonable upper limit for the IS as commonly accepted. Slightly less than five percent of records were taken out of the IS by this restriction

The 1993 ICLS resolution also recognised a sub-division of the IS into:-

- Informal Own Account enterprises, who can employ unpaid family workers and the occasional paid employees but no continuing paid employees.
- Enterprises of Informal Employers, who can employ continuing paid employees but usually only up to a low size limit e.g. less than five employees.

This classification will be applied in some of the tabulations.

A discussion of the effectiveness of this IS identification is contained in Section 11.

5. SAMPLE DESIGN

The LFS was the first survey to be conducted on a new master sample. Full details of the design of the master sample can be obtained from the Household Surveys Unit of the CSO. However a brief description is given below. The base data for the design were the 1991 Population Census. No complete data were available for dwellings or for households which were vacant, non-contacts or refusals for the census.

5.1 Stratification:

Fourteen strata were formed, five urban and nine rural. The five urban Strata were-

1. Gaborone
2. Francistown
3. Selibe Phikwe
4. Small Towns
5. Urban Villages

The Small towns and Urban Villages strata are in accordance with the 1991 Census classification. The nine rural strata correspond with the nine rural Districts. The samples were selected independently within strata and thus valid results could be produced for each of these strata. Note however that complete District results will need the estimation of any urban small town or urban village component to add to the rural component.

5.2 Creation of sample Blocks

It was decided to create sample blocks of approximately 150 households. This was done by increasing the household count in all the 1991 Census Enumeration Areas (EAs) by 15% for growth and then dividing by 150 to obtain a measure of size. In a relatively few cases, small EAs were combined to form a composite block. The resulting measures of size were then summed to calculate a stratum allocation of the 420 sample blocks

5.3 Allocation of sample Blocks to Strata

The stratum allocation is given below. It was designed to give low standard errors (Ses) for the rural stratum as can be noted, a very high sample in the smallest rural strata.

Table 1 - CHIPS MASTER SAMPLE

	1991Census	Proportional	Final
Stratum	Blocks	Allocation	Allocation
1	421	56	34
2	194	26	29
3	119	16	28
4	139	19	29
5	518	69	36
6	243	32	30
7	70	9	27
8	247	33	30
9	84	11	27
10	706	94	41
11	89	12	27
12	191	25	29
13	57	8	26
14	69	9	27
Total	3147	419	420

A minimum sample of 25 blocks was taken from each stratum and then the remaining blocks were spread by strata. The result was a sample heavily biased to small districts and overall to rural rather than urban areas.

5.4 Field delineation of Blocks

The blocks were selected systematically within strata with a random start, probability proportional to size (PPS) using the Measures of Size (MoS) mentioned earlier. With the blocks selected, Household survey and Census cartographic staff were used to attempt to delineate the blocks in the field. If the selected EA had a MoS of 1 then the EA was the Block. If the selected EA had a Mos of 2 then it had to be divided equally into two parts etc. If possible natural features were to be used to identify boundaries. The selected blocks were listed in mid-1995. All habitable structures were

to be listed and separate dwelling numbers were to be given to externally identifiable households within dwellings (e.g. rows of rooms with external doors rented out).

5.5 Dwelling Selection within Blocks

Household selection was done in the office after block listing, in most cases. The aim was to obtain a sample of approximately 30 dwellings per block. It was acknowledged that there would be some additions for multiple households within dwellings and some sample loss through vacant households and non-response. The aim was to send interviewers in the field with a pre-determined listing of each block. However, in some cases the listing was done late and a sample had to be selected in the field. In such cases a sample interval and random start to apply were supplied to the interviewers.

These are the main features of the sample design. See Section 10 for details of the major problems encountered because of this design and the practical problems in implementing it.

6. QUESTIONNAIRE DESIGN

The Questionnaire design is a key activity in any survey. During questionnaire development, questionnaires from other African countries were consulted, particularly the Tanzania Labour Force questionnaire and valuable feed-back was given by the ILO Bureau of Statistics in Geneva.

The major features of the final questionnaire design are:-

LFS1 - LFS1 is a household questionnaire, one for each household and also serves as a cover for the LFS2s for the individuals.

Page 1 -This page was largely for administrative use in the field and allowed recording of field details particularly details of response problems. The final response status was entered on computer records as well as the total number of persons and the number eligible for LFS2

Page 2 & 3 - These pages contain the basic household listing with key demographic, educational and migration questions. The layout seemed to work well with few problems.

LFS2 -LFS2 is the individual questionnaire and the key questionnaire for the LFS. There was one for each person aged 12 years and over as identified on the LFS1. The basic design was as follows:-

Page 1 - A repeat of the indicative in detail was given here which as noted was too extensive and confusing. Similarly the interviewer activity details were again too detailed and confusing. It would be best if most of this was kept to LFS1. Q1 and Q2 were check questions on age, sex and citizenship. They were already contained on LFS1. It was hoped that LFS2 would be completed with the respondent in person whereas LFS1 could have been done with an informant only. Thus the repeat questions on age, sex and citizenship were meant to improve the quality of these data items. There was no difference but the main use was to assist in identifying LFS2s when the confusing indicative was filled in incorrectly. In future these questions should not be necessary.

Page 2 to 4 - The important detailed check list of economic activities for both usual and current activities was on page one with the first questions on usual activity. The usual economic activity questions then continued up to most of Page 4.

Page 4 and 5 - Major filter questions on current economic activity and details for persons not economically active and first unemployed question.

Page 5,6 & top of 7 - Details for persons currently unemployed.

Page 7 & 8 - Questions for employed persons on their main economic activity

Page 9 & 10 - Questions for employed persons on their secondary economic activity, if any.

Page 10 (bottom), 11, 12 & part of 13 - Questions on hours worked, current and usual but note the hours are for that current activity. Also contained are some experimental questions on underemployment.

Page 11 & 12 - Income questions

Page 12 & 13 - Migration and work in RSA

7. PRETEST

A Pre-test was conducted in March 1995 in Gaborone and surrounding rural areas. Nine EAs were covered with thirty households in each. The pre-test was basically meant to test the questionnaire. Minor corrections were made to the draft questionnaire after the pre-test.

8. MAIN SURVEY OPERATORS

8.1 Collection spread

The Survey sample was spread evenly over the twelve months period. Sample Blocks were allocated systematically by stratum to quarters to attempt to give a balanced representation of sample blocks in each quarter. In a few cases interviewing in a particular block had to be deferred but they were replaced by a similar Block in the same stratum where possible. The small table below gives the final number of person records per quarter. This shows a good balance between the quarters.

Table 2: 1995/6 LFS Number of Person Records per Quarter

Quarter	Count	%
Total	45,616	100.0
1	11,220	24.6
2	11,770	25.8
3	11,644	25.5
4	10,982	24.1

Note for any quarterly data, the four quarters represent the following approximate periods:

1st Qtr - Last calendar Quarter of 1995
2nd Qtr - First calendar Quarter of 1996
3rd Qtr - Second calendar Quarter of 1996
4th Qtr - Third calendar Quarter of 1996

8.2 Field inspections

The Family Health Survey over-lapped with the LFS which was very undesirable and restricted greatly transport availability and time of senior staff for LFS matters. Also the overall Government diversion of transport for the eradication of cattle in the north restricted transport availability. It should be stressed however that a quality control system needs to be established for Household Surveys. The system should be independent of senior household surveys staff and subject matter staff field visits. Principal and Senior Statisticians' field visits can only be rare with their other duties.

8.3 Sample Loss

The response summary below gives the final status of the nearly 13,000 households in the sample by the three major geographic groupings (as used in the 1991 Census). A program was prepared to produce this type of summary for each EA once data had been entered. However data processing problems with missing and duplicate households made the summary meaningless and they were not used for field control. If the data processing problems can be solved, such summaries are very valuable for locating interviewers and teams with high non-contact and vacancy rates. It still needs a field quality control system/officers to go and check on all EAs with suspiciously high sample loss. All vacants, at a minimum, should be checked by supervisors. The vacancy rates below are high. The rural rates may be explainable but the urban rates are high given the pressure on accommodation in urban areas. Non-contacts are also higher than for surveys in most other developing countries.

Summaries as presented below should be produced and published for all surveys in future and research made into what is an acceptable vacancy rate for the different geographic areas of Botswana.

Table 3: All Sampled Households by Geographic area by Final Response Status

	Sample	OK	Non	Refusal	Vacant	Destroyed	Other
NUMBERS	H'holds		Contact				
Urban Towns	3669	3323	122	18	196	9	3
Urban Villages	1078	972	28	3	98	4	1
Villages	8039	6741	105	4	1107	60	5
Grand. Total	12786	11036	255	25	1401	73	9
PERCENTAGE	%	%	%	%	%	%	%
Urban Towns	100	91	3	0	5	0	0
Urban Villages	100	90	3	0	9	0	0
Villages	100	84	1	0	14	1	0
Grand. Total	100	86	2	0	11	1	0

9. DATA PROCESSING

Data processing of a new major survey is always a complex activity. CSO has very little experience in survey processing and very few trained staff. As noted earlier the final questionnaire was very similar to a Tanzanian one and thus it was hoped and expected that many aspects could be copied from Tanzanian experience. This did occur to some extent but many problems were still experienced.

9.1 Manual Editing and Coding

Supervisors were to check the forms in the field but this was rarely done effectively if at all. The main editing and coding was done in the office. Special instructions were prepared and a special training course conducted. The major tasks were:

* Checking completion of workloads.

  Checking each questionnaire for completeness and imputing where necessary

  Coding for BOSCO, BISIC and Subject of Training as discussed on Section 2.

  Checking the `other' categories in questions to see if they could be re-allocated to more suitable categories.

  Checking the "Informal Sector evaluation".

  General checking of sequences, completeness and consistency.

The LFS questionnaire was quite complex and such editing required considerable diligence particularly as little editing was being done in the field. In general the manual editors and coding had many problems. A very high error rate was located in later computer edits. It did seem that many of the editors did not understand the strict rules for editing and coding and applied their own interpretation. Supervision was a problem with often little effective supervision.
In future it is recommended that a special small group be set up for the major tasks of editing and coding under the supervision of Labour Statistics staff and at the same location as Labour Statistics staff.

If this group could be used for all economic activity coding in CSO, experience could be built up in this exacting task. Manual editing and coding continued throughout 1995 and concluded in late 1996

9.2 Data Entry

Data entry was done on independent personal computers (PCs) which were linked to the CSO network. However for some reason, the network could not be made compatible with Dbase4 and in effect each PC operated independently. This did cause problems when corrections were made to the data entry program or the look-up databases for valid EA codes and other codes (occupation etc.). The amendments were often not done to all the versions which created some of the later errors. The data entry staff were given special training for the LFS data entry. Unfortunately there was no data entry supervisor but the programming staff were often available although actually located in a separate building. One household surveys officer located in the same building was also often present to assist with problems. Overall the data entry system was far from ideal and could be improved with little effort.

9.3 Data Control

The necessary techniques for data control are well documented in various UN and other manuals but have yet to be introduced for CSO data processing. This could be done with overall improved management with minimal effort and is obviously needed. The officer who finally controlled the data did a good job.

9.4 Computer Editing

An extensive series of 145 computer edits were run. The Computer edits produced a very high number of errors. Staff of household surveys were used to correct the errors initially but this was very time consuming. As noted earlier, if the manual editing and data entry had been better, the number of errors located at this stage would have been much fewer. Several cycles were necessary before each quarter was clear of these major edits. Finally a small group was used to clear the remaining errors and this technique of a small specialised group would work better in future for both coding and editing.

9.5 Derived Items

The major Labour Force categories (e.g. current and usual employment status) are derived data item from several questions on the LFS2 questionnaire. A program for these derivations was written and applied as the data was clean of all edits. The derivation however again located some errors which had to be fixed. A full list of derivations and how they are derived is given in the separate paper on data file documentation.

9.6 Frequencies and Check Tables

Frequencies of all questions and check tables were run as soon as each quarter's data was clean. These located unusual values which had not been picked up by computer edits. These were again queried and amended as necessary. The UN package, Xtable, was used for most of this work as it

is easily to learn and use and very flexible. Literally hundreds of small tables were produced during this checking.

9.7 Data Editing

The LFS data set was intensely edited as described above and the result is a very clean consistent database. There should be very few unusual values discovered during data analysis. The whole effort though has revealed the very severe weaknesses in the CSO data processing system which needs urgent attention.

9.8 Tabulations

Final weighted tabulations were initially produced using Xtable. Xtable although very limited, is very flexible and easy to learn. Tables were produced using Xtable and brought into the Excel package for other manipulations and to make them attractive for publication.

10.0 SAMPLE and WEIGHTING PROBLEMS and ADJUSTMENTS

10.1 Sample size and Stratum Allocation

A sample of 420 Blocks and it's allocation by stratum was given in Section 5. It was an extremely large sample for Botswana, about 5 percent of the population was to be included (approx. 30 dwellings per Block). In the small rural districts, a quarter of the population were to be taken. The aim was to produce estimates with low standard errors for these rural districts but it did seem to be carried too far. The large sample with a large rural component severely strained CSO's resources. A consultant was employed by Overseas Development Administration to re-examine the sample design and one of his major recommendations was that the sample should be cut and more resources put into quality control. The University of Southampton based consultant also recommended a considerably different stratum allocation which gave higher urban proportion. This would be much more suitable for LFSs as the variety of economic activity is greatest in urban areas. A copy of the consultants report is available in CSO's Household Survey's Unit. For this LFS, the sample strongly represented remote rural areas which have a very limited variety of economic activity. Conversely, the urban sample was low and many of the smaller characteristics by occupation, for example, will have high standard errors.

10.2 Block selection and Measures of Size (MoS)

The basic decision to select Blocks rather than EAs seems to have been based on a false premise that EA sizes in terms of number of households were often large and variable. A quick frequency showed that 98 percent of EAs were likely to have less than 250 dwellings (1991 Census plus 15 percent) and only five were likely to have greater than 350 dwellings. (Very small EAs were combined.) Generally it is no problem for an interviewer to cover 30 dwellings out of 250 (or even 350). Most Botswana villages are fairly tightly clustered so that most samples would have been in easy walking distance. The few large EAs or scattered cattle post or lands areas could have been specially segmented. The segmenting of the EAs into Blocks was a major and unnecessary task for Household surveys in most cases and created major problems. The MoS of 150 also caused a high

degree of rounding (everything less than 224 Dwellings to 1 etc.) and thus the PPS probability became very approximate. The Southampton consultant also recommended that this procedure be discontinued.

10.3 Demarcation of Blocks:

There were many problems in demarcating the Blocks. The worst problems occurred when the 1991 Census EAs did not have clear boundaries themselves. For the 1991 census it was apparently decided to keep the 1981 boundaries often and if necessary, create two EAs within the 1981 boundary. The EAs within the 1981 boundaries did not have geographic boundaries and were only locatable by the census house numbers. Thus the first problem often was to locate the EAs when many house numbers had been removed. Trying to add growth to an EA without geographic boundaries was of course very difficult. If the EA was segmented then it was very difficult to create a block. Most sketch maps of blocks (& EAs) were extremely difficult to follow. Most household survey interviewers have no training in map reading let alone in how to read someone else's sketch map.

10.4 Listing of Selected Blocks

When the first preliminary LFS results were produced in early 1997 for the Budget, it was obvious that there were major problems. Only a few basic unemployment rates were released for Budget purposes. A comparison of the listing total dwellings by block from the 1991 Census, 1991 Census plus 15 percent, the LFS and the FHS showed very large variations both plus and minus. It was decided to produce two quick sets of weights, one based on the ideal in line with 1991 Census data and another on the LFS Listing Totals. Both were benchmarked to age/sex cells to see if this would correct any problems. The two sets of estimates by major characteristics (sector and status) were very different. Listing is very important and a major factor in estimates and the weighting exercise confirmed this point. The differences noted earlier occurred not only in rural areas but also in urban areas. A quick investigation was launched into some of the differences in Gaborone. Boundary identification was definitely a problem but interviewers sometimes took the whole EA rather than the Block. Dwelling definition was also a problem.

This created a major problem as no reasonable weights could be calculated from the existing listings. It was decided to launch a re-listing of all EAs with major differences from the comparisons mentioned earlier. Interviewers were re-trained and new better listing forms were designed. Finally almost all EAs were re-listed in the new format. The Blocks were abandoned and complete EAs were listed. For the problem EAs mentioned earlier, the complete 1981 boundary was re-listed and new EAs were created within that boundary but with clear geographic boundaries. The number of EAs within a 1981 boundary was the same as 1991 but the house numbering was abandoned as a means of demarcation.

This re-listing was being finalised when the Consultant arrived from Southampton. He reviewed the problem and the preliminary results of the re-listing which still showed many differences. After discussions he recommended that EAs with major differences plus or minus from the census data should simply be removed totally from the sample and the weighting adjusted. This was the least biased procedure and the sample was large enough to take this sample loss. A thorough check of the re-listings found many clerical errors and other problems and finally three EAs were totally

removed from the sample. This was in late August and gave little time for new weights and these current estimates.

10.5 Weighting

There are several components to the weighting-

- **From EA Level to Stratum level** - As noted earlier the block listings were not reliable and had to be abandoned. We decided to use the re-listing totals for EAs and assume the sample selected from the blocks was representative of the whole EA. In 56 percent of cases the block was the EA and there had been no segmenting. The actual selection probabilities were no longer valid and an alternate had to be used. It was decided to use a simple weight in each stratum of the 1991 Census total households divided by the sum of the 1991 census households for the selected EAs in that stratum. The selection by MoS had been an approximate PPS sample and this weighting assumed was actually at the detailed level of number of households.

- **From Household Level to EA Level:** - This is a simple weight obtained by dividing the EA listed total dwellings by the number of selected dwellings in that EA. Externally identifiable households within the one structure were taken as dwellings for listing and selection purposes. Additional multiple households can and were identified at the interviewing stage and these were called households rather than dwellings. This is the key component which relies heavily on accurate listing totals for each EA.

- **A non-response adjustment**:-

For LFS, no substitution was allowed for non - response and household questionnaire (LFS1) had to be returned for all households, responding or non-responding. On the front of LFS1 a response code was entered after all field follow up and manual processing. The response code was entered on computer records. The codes are:-

1	Fully Responding
2	Non-contact
4	Refusal
6	Dwelling Vacant
7	Dwelling Destroyed
8	Other

Only Non-contact and refusals were taken as non-response. The other sample loss was effectively taken as zero i.e. no-one lived in these households. With the high vacancy rate this is a worrying assumption but no other alternative was possible. The non-response adjustment was made at the EA level. The adjustment was equal to the total households located in the EA divided by the presumed valid response in that EA (i.e. Code 1 plus Codes 6 to 8 above). In effect non-contacts and refusals are given the characteristics of average valid respondents in the EA. Note that total households were used including any multiple households located during interviewing rather than the original selected dwellings. The overall average non-response adjustment was only 1.022 but there was considerable variation with many 1 (no adjustment) and the maximum 1.31.

- **Overall Weight**: The overall weight was thus a product of the three weights above and was entered on computer records as WEIGHT1

$$WEIGHT1 = Wt1 * Wt2 * Wt3$$

10.6 Benchmark Adjustments

When the above weight, WEIGHT1 was applied to the person level data the results were still below census projection but not to the extent shown with the earlier listing. The problem is that the difference is highly age and sex specific. There is a severe problem of under-counting of younger adult males in all Botswana surveys. As Table 4 below shows, there was a suspected undercount overall of 13 percent of Males and 5 percent of Females. Such rates are not unusual and are very similar to, but higher overall than the 1994 HIES adjustment rates. Economic activity is highly related to age and sex and such adjustments are very common in other countries.

The Benchmark taken was the middle variant population projection for March 1996 which is approximately mid-survey. Deductions were made for the following groups; BDF staff on bases, Guests in hotels and Prisoners, who were not covered in the LFS. None of these groups are very large but they are very age specific and significant proportions of the problem age groups. Estimates of these groups were obtained by using census data as a base and projecting with data from other sources. It was initially proposed to do the adjustment for Rural and Urban areas separately but the resulting ratios would have meant lowering rural estimates considerably particularly for women. It is suspected that the Urban projections are currently too high. It was decided to do the LFS adjustments at the National level and thus show the LFS Urban/Rural pattern rather than adjust it out. Demographic staff have been made aware of this conclusion.

The ratios below were applied to WEIGHT1 from the earlier calculations and are on the Person data files as WEIGHT2. Note that only the weights 12 years and above are important for economic characteristics.

Table 4 : Age/Sex cell Benchmark adjustments to LFS data

Total	Males	Females
0-4Yrs	1.22	1.17
5-9Yrs	1.16	1.17
10-11Yrs	0.94	0.99
12-14Yrs	1.05	1.03
15-19Yrs	1.07	1.07
20-24Yrs	1.12	1.04
25-29Yrs	1.29	1.12
30-34Yrs	1.19	1.07
35-39Yrs	1.24	0.99
40-44Yrs	1.15	0.87
45-49Yrs	1.20	0.92
50-54Yrs	1.29	0.96
55-59Yrs	1.09	1.01
60-64Yrs	1.14	1.16
65+Yrs	0.94	0.81
N/S	0.00	0.00
Total	1.13	1.05

10.7 Household Data

The above sections discussed the person level data which is the major interest in labour force survey data. However household tables are also important. It was decided not to apply any adjustments to the household level data as projections at this level are much more uncertain.

It must be noted that the household tables produced from LFS simply sum the persons in households by various characteristics to the household level. It was not possible to apply any variable age adjustment when doing this. Thus average occupancies on the household data are definitely low by whatever the under-enumeration within households was. Users must be aware of this. As noted it, seemed the young adult ages were the most effected. There is no WEIGHT2 on the Household data file.

10.8 Quarterly Estimates

With the relatively large sample and a good balance of data collection between the quarters, it should be possible to produce reasonable quarterly estimates. It was decided to adjust the quarterly data produced using Weight1, to the same Benchmark for each quarter. The Benchmark was the same as for the overall annual figure. This is similar to the demographic technique of using a stable population model. The resulting weights were put as WEIGHT3 on the data file. The results for some of the major variables are discussed in Section 11.4.

11. DATA EVALUATION

The objective of this section is to critically review aspects of the data with a view to possible improvements to methodology or definitions in future. It will highlight strengths and weaknesses in the data which detailed users should be aware of.

11.1 Total Economic Activity Measures

The first and major interest in a survey's results is in how well it measures the major aggregates of economic and non-economic activity and the major status in employment components of the employed population. The attached table gives a comparison of the aggregate current measures from the 1995/6 LFS, the 1994 HIES and the 1991 Census.

As can be noted, for LFS the total paid employees looks reasonable as does the unemployment rate and the number of not economically active Students. However the overall participation rate in economic activity is low. This is caused by very low numbers in unpaid work in family businesses and work on own lands, cattleposts etc. There seems no doubt these are underestimated compared to the other two data sources and there is a corresponding overestimate in the numbers doing housework (particularly males), and retired etc.

The usual activity measures have produced considerably better measures of the employed population. The concepts and reference periods of usual and current activity are quite different but there still should be a very close relationship between the aggregate totals. As the table below shows, the number usually economically active and employed (403,093) is considerably higher than those currently employed (345,405).

Table 5: 1995/6 Labour Force Survey - Comparison of Current and Usual Economic Activity measures

CURRENT		USUAL ACTIVITY		
ACTIVITY	Total	Act & Employ	Act & Unempl	Not Active
Total	950793	403093	179497	368203
Employed	345405	297503	35001	12901
Unemployed	182703	46283	121542	14878
Not Econ Active	422685	59307	22954	340424

Table 6: COMPARISON OF 1991 CENSUS, 1994 HIES and 1995/6 LFS
POPULATION BY SEX AND ACTIVITY - ADJUSTED TO AGE/SEX CELL BENCHMARKS

1991 CENSUS (12 Years and over)

	TOTAL	Economically Active					Total Economically Active	Participat. Rate (excl UK)	Unempl. Rate	Not Economically Active			Unknown
		Paid Employees	Self Employ	Unpaid Work in Family Business	Unpaid Work Lands,Fm, C' Post	Actively Seeking Work				Housework	Student	Retired	
NUMBERS													
TOTAL	859221	275928	28639	7759	67612	61265	441203	52.4	13.9	182725	190734	26558	18001
Males	402056	175734	14735	4125	44979	31852	271425	69.5	11.7	17617	89677	12082	11255
Females	457165	100194	13904	3634	22633	29413	169778	37.7	17.3	165108	101057	14476	6746
Prop of Total		%	%	%	%	%	%			%	%	%	%
Persons	100	32.1	3.3	0.9	7.9	7.1	51.3	n/a	n/a	21.3	22.2	3.1	2.1
Males	100	43.7	3.7	1.0	11.2	7.9	67.5	n/a	n/a	4.4	22.3	3.0	2.8
Females	100	21.9	3.0	0.8	5.0	6.4	37.1	n/a	n/a	36.1	22.1	3.2	1.5

1994 HIES

	TOTAL	Economically Active					Total Economically Active	Participat. Rate (excl Un)	Unempl. Rate	Not Economically Active			Unknown
		Paid Employees	Self Employ	Unpaid Work in Family Business	Unpaid Work Lands,Fm, C' Post	Actively Seeking Work				Housework	Student	Retired	
NUMBERS													
Persons	908849	261999	47373	15328	67104	107723	499527	55.3	21.6	161763	208865	32766	5928
Males	420501	151278	18751	3285	49347	54959	277620	66.5	19.8	20704	101258	18148	2771
Females	488348	110721	28622	12042	17757	52764	221907	45.7	23.8	141059	107607	14619	3157
Prop of Total		%	%	%	%	%	%			%	%	%	%
Persons	100	28.8	5.2	1.7	7.4	11.9	55.0	n/a	n/a	17.8	23.0	3.6	0.7
Males	100	36.0	4.5	0.8	11.7	13.1	66.0	n/a	n/a	4.9	24.1	4.3	0.7
Females	100	22.7	5.9	2.5	3.6	10.8	45.4	n/a	n/a	28.9	22.0	3.0	0.6

1995/6 LFS

	TOTAL	Economically Active					Total Economically Active	Participat. Rate (excl Un)	Unempl. Rate	Available but not Seeking Work-reason		Not Economically Active			
		Paid Employees	Self Employ	Unpaid Work in Family Business	Work on own Lands,Fm, C' Post	Actively Seeking Work				No work etc.	H'work	Housework	Student	Retired	Unknown
NUMBERS															
Persons	950793	276039	34695	3238	31433	94528	439933	46.3	21.5	59347	28828	134940	208826	71655	7264
Males	443933	154501	15288	1504	18008	45461	234762	52.9	19.4	26588	9899	40369	102852	25875	3588
Females	506860	121538	19407	1734	13425	49067	205171	40.5	23.9	32759	18929	94571	105974	45780	3676
Prop of Total		%	%	%	%	%	%			%	%	%	%	%	%
Persons	100	29.0	3.6	0.3	3.3	9.9	46.3	n/a	n/a	6.2	3.0	14.2	22.0	7.5	0.8
Males	100	34.8	3.4	0.3	4.1	10.2	52.9	n/a	n/a	6.0	2.2	9.1	23.2	5.8	0.8
Females	100	24.0	3.8	0.3	2.6	9.7	40.5	n/a	n/a	6.5	3.7	18.7	20.9	9.0	0.7

There will always be some movements in employment over a year but the table below traces the current non-economic activity for those currently not active but usually employed by their usual activity (Q7B) and for those currently unemployed the simple total (Q7B) as no current categories are available for the unemployed. What it shows is that most of the usually employed (not currently employed) were recorded as working on own lands for Q7B and were recorded as not active and doing household duties currently. There seems to definitely be a problem with home duties as a current activity (non-economic).

There seems no doubt that these persons should have been recorded as economically active. Interviewers did not see the relationship between current and usual activity. Respondents were also confused by the question. This is unfortunate as not only agricultural work seems to be understated but also self employment. Once excluded from the employed category there are none of the details of occupation, industry, sector, time worked etc.. The main problem seems to be between employed and not economically active. The total usually unemployed is very similar to the total Currently unemployed (A & B).

Table 7 : Characteristics of Persons Usually Employed but Currently not Employed

USUAL ACTIVITY	Usually Employed-Current Non-economic Activity							Usually Employed
Q7B Answer	Total	Att Sch	H'hold	Old	Sick	Disable	Other	Currently Unem.
Total	**59307**	**848**	**45548**	**2700**	**8500**	**218**	**1493**	**46283**
Lands etc.	49284	247	40373	2500	5470	78	616	23432
Paid Employees								
Cent Govt	1869	222	1098	30	242	0	277	3513
Local Govt	2299	152	1056	54	830	11	196	3024
Parastatal	304	19	212	0	48	0	25	674
Private Sector	2411	55	985	63	1147	73	88	11197
NGOs	42	0	9	0	0	0	33	93
Private H'holds	1739	0	1021	53	366	48	251	3105
Self Employed								
with Employees	101	0	39	0	62	0	0	282
Without Employees	803	0	512	0	276	8	7	883
Unpaid Helpers/F.Bus.	455	153	243	0	59	0	0	66

11.2 Training Classifications

Training is a topical issue in Botswana currently with the Government committed to spending considerable amounts of money on training, particularly vocational training. There was an attempt to record more details of training for the LFS with this policy interest in mind. Three classifications were used:-

- **Level of training** - (LFS1 Column 11a) This classification gave the most problems and is not really a classification as the codes are not unique. It should definitely not be used in its current format again and needs further research. A field circular was issued after about a month to tell interviewers that the following rules were to be applied for codes 04 to 08.
  - ◊ 04 Elementary Certificate - Courses less than 1 year.
  - ◊ 05 Intermediate Certificate - Courses from 1 to 2 years
  - ◊ 06 Advanced Certificate - Courses 2 years upwards
  - ◊ 07 Certificate - Courses of unknown duration but usually shorter/simpler ones
  - ◊ 08 Diploma - Only for diplomas of 2years plus with shorter ones to be coded as for certificates above.

This was not always applied and did not solve all the problems. The Trade test A is apparently the equivalent of a National Craft certificate and the City and Guilds are exams which are part of the qualification of technicians. The following resulting classification was derived for the data but it is very uncertain:

1. Trade Test C (Code 1)
2. Trade Test B (Code 2)
3. Trade Test A & NCC (Code 3 and 10)
4. Other Technical (Codes 09 & 11 to 13)
5. Elementary & Unknown Certificates (Codes 04 & 07)
6. Intermediate Certificates (Code 05)
7. Advanced Cert. & Diploma (Code 06 & 08)
8. Junior Degree (Code 14)
9. Senior Degree (Code 15)
10. Unknown (Code 99)

The technical qualifications are almost certainly understated and some were recorded under the general certificates etc. Thus overall users should treat this data with great caution.

- **Subject of Training**
- **Institution of Training** – LFS1 Column 11C - Note that two codes for the other category were introduced in the field and for coding

10. Other - Botswana
11. Other - Outside Botswana

Most entries are in 10 but the earlier codes are of some use. For example investigations have already been done of the effect of Vocational(Code 02) and Brigades (Code 05) training on employment prospects. Except for the limited number of institutions listed this approach seems to have worked reasonably well.

11.3 Non-Citizen Estimates

Non-citizen estimates are always difficult from household surveys in developing countries as the non-citizens tend to live in clustered locations and the sample cannot represent such a population well. However the sample in the LFS was relatively large and the estimates below, by country and by current economic activity are reasonable. The number of non-citizens employed is close to employment survey estimates and well above work permit estimates. The number from South Africa is low compared to work permit data and the number from Zimbabwe is high. Given the limitations of the sample these estimates are plausible.

Table 8: Population 12 Years and over by Citizenship by Current Econ. Activity Status

	Current Economic Activity Status			
	Total	Employed	Unemployed	N.E.A.
Total	950793	345405	182703	422685
Botswana	928780	331859	180498	416423
South Africa	3889	2563	212	1114
Zambia	2732	1248	436	1048
Zimbabwe	5384	3048	857	1479
Other Africa	2374	1662	112	600
U.S.A	637	432	120	85
India	2074	1164	114	796
Other Asia	1041	647	273	121
United Kingdom	2835	2033	68	734
Other Europe	1047	749	13	285
Non-Cits Total	22013	13546	2205	6262

** NEA - Not Economically Active*

11.4 Persons Temporarily Absent from work

As noted in section 2, persons temporarily absent from work (with certain set conditions) were to be included as economically active. This question has caused problems in several LFSs in other countries. However it seems to have worked reasonably well in this survey. The table below gives the reasons reported for working low hours by Sector for the temporarily absent. The reasons with zero answers have been left in to show that there were few absurd answers. Overall, 5.1 percent of the employed were temporarily absent from work in the reference weeks, which is a plausible figure. The numbers temporarily absent in Traditional Agriculture and the Informal Sector are relatively high but the distribution of reasons is reasonable. The only reason with a relatively high number which is suspect is 'Can't find more work'. This is not really a reason for being temporarily absent from your main job and possibly such persons, particularly in the Private Sector, are really unemployed.

Table 9: Persons Temporarily Absent from work - Reason given by Main Sector of Employment

Reason for Absence	Total	Central Govt	Local Govt	Parastatals	Private Sector		
					Trad. Agr	Informal	Other
Total	17092	5414	1314	402	2784	3452	3726
Can't find More work	699	3	42	0	173	430	51
No suitable land	0	0	0	0	0	0	0
Lack of Raw Materials	1198	0	0	0	0	828	370
Machinery Breakdown	13	0	0	0	0	13	0
Industrial dispute	24	24	0	0	0	0	0
Disability	0	0	0	0	0	0	0
Illness, Old age	1506	100	192	0	345	473	396
Household commitment	1575	37	50	0	692	615	181
In school or Training	288	149	12	0	7	0	120
Leave, holiday	11356	5027	1018	402	1363	1006	2540
Didn't want More work	25	0	0	0	10	15	0
Other	408	74	0	0	194	72	68

11.5 Underemployment Measures

The basic measures of underemployment discussed so far and shown in most tables are the traditional measures of visible underemployment. Persons working below a certain number of hours (in this case 35 hours) are asked if they would like to work more hours and whether they were available to work these extra hours. The ILO has been looking at alternate measures and some additional questions were included in this survey to test some of the alternatives. In the first difference, all persons were asked if they would like to work more hours (Q52). The result is shown in the table below. The number under-employed increased considerably from the 35 hours limit to having no limit. The right hand side of the table shows that most were wanting additional work at their current job rather than different jobs (Q52). It is notable that the increase is highest in the Other-Private (the Formal Private Sector) but even in Central Govt., there was a sizeable increase.

Table 10 : Number of Persons Under-employed by Differing Conditions & Type of Job Sought

Sector of Employment	Total	UNDER-EMPLOYED		UnderEmp - No limit - Extra Hrs		
		35Hr Limit	No Limit	Curr Job	Diff Job	Both
Total	345405	23348	44118	31778	11164	1176
C-Govt	75479	2392	6430	5066	1180	184
L-Govt	39917	11715	11708	8013	3417	278
Parast	12135	60	1343	1154	104	85
Trad Agric	48657	2737	3538	965	2457	116
Informal	57240	4797	7637	5060	2424	153
Other-Private etc.	111977	1647	13462	11520	1582	360

This is interesting additional information but it does need to be considered whether government policy should be so concerned about persons already working long hours and wanting to work more as it should be about those working low hours. Many wanting additional hours were working 45 hours already. Almost all stated that they wanted more money as their reason (Q55). The characteristics of the different groups does need further research.

In a second set of questions, employed persons were asked if they wished to change jobs (Q57). Some of the results are given below in Table 11. Persons who feel their qualifications are not being used are underemployed in a broader sense. It is interesting to note that over 10 percent are considering self employment and that 'better working conditions' are a major reason given. As with the previous group of questions this topic also needs more cross classification by other variables and more research. Both however give useful additional insights into the under-employment problem.

Table 11 Employed Persons wishing to Change Jobs

A: BY SECTOR OF MAIN EMPLOYMENT

Reason	Total	SECTOR OF EMPLOYMENT					
		C-Govt	L-Govt	Parastatals	Trad Agric	Informal	Other-Private
Total	105725	14650	13434	2804	16409	20715	37713
Better Use of Quals	6882	1217	855	341	1495	1132	1842
Better pay	77352	9458	11237	1541	13387	16766	24963
May lose Job	3860	959	343	258	307	475	1518
Nearer home	1390	422	156	68	228	156	360
Better for health	5347	1066	355	166	361	711	2688
Better W/conditions	10894	1528	488	430	631	1475	6342

B: BY TYPE OF WORK SOUGHT

Reason	Total	Paid	Self Employment	
		Employ.	Small Bus	Agric
Total	105725	94696	9306	1723
Better Use of Qualifications	6882	6197	452	233
Better pay	77352	70489	6138	725
May lose Job	3860	3590	138	132
Nearer home	1390	689	437	264
Better for health	5347	4411	848	88
Better Working conditions	10894	9320	1293	281

11.6 Income Data

Little analysis has been possible at this stage of the income data but the brief tables below give a summary of average earnings by Sector. Income data are difficult on one visit household survey and are often unreliable, especially for the self employed. The gross paid earnings shown below is in reasonable agreement with the employer survey data, when consideration is taken of the groups covered in LFS and not in the Employer Surveys (i.e. domestics, paid employees in traditional agriculture, drought relief employees in Local Government, Informal sector paid employees). The self employed data are of the expected pattern and needs comparison with other sources.

Table 12: Average Monthly Earnings

GROSS PAID INCOME			
BY SECTOR	Total	Male	Female
Total	804	930	645
C-Govt	1123	1227	1012
L-Govt	513	612	426
Parastatal	1340	1482	1045
Pvt Trad. Agric	119	113	160
Informal	202	300	169
Other	891	1032	650
SELF EMPLOYED INCOME (net of expenses)			
BY SECTOR	Total	Male	Female
Total	1371	2198	665
Informal	657	1004	415
Other	4468	5369	2620

11.7 Informal Sector Measures

Informal sector employment can be as a main or secondary activity and as a self employed person or as an employee. Table 13 below gives the total IS employment in the LFS. Of the total, 4,700 were employed in the IS as a secondary activity almost all of whom were self employed. Note, in a few cases a person was in the IS in main and secondary activities and is only counted once below

Table 13 Total Informal Sector Employment

	Total	Males	Females
TOTAL	60501	21071	39430
Self Employed	30951	12742	18209
Paid Employees	26962	6854	20108
Unpaid Helpers	2588	1475	1113

The total above compares with the total for the 1992 USAID/GEMINI study of 54,400 of whom 28,400 were self employed, 17,600 paid employees and 8,400 unpaid. The LFS definitely underestimated unpaid helpers and probably underestimated the smaller self employed activities as well. However the overall comparison is not unreasonable.

The attached table has been included to show an attempted reconciliation between survey/census estimates and the Employment Survey estimates. The Employment Survey estimates have been used for Central Government, Local Government and Parastatals as they are not available for HIES and the census. The drought relief work scheme was not operating in 1991 but was in 1994 at probably a higher level than in 1995/6. There was some labour intensive road works in 1991 which is similar. The conclusion drawn from this data by CSO was that the Employment Surveys were over-estimating Private sector employment in the unknown Strata. Revisions will be made to Employment Survey methods in future.

Table 14: Comparison of Census/Survey and Employment Survey Estimates

	1991 Census (August)	1994 HIES (Avge 12 Mths)	1995/6 LFS (Avge 12 Mths)
Total Paid Employment-Cen&Surv	275,928	261,999	275,670
Paid Employ in Institutions(BDF etc.)	4,000	0	0
A: Comparable Total Paid Empl.	271,928	261,999	275,670
Employment Survey Ests	Sep 1991	Sep 1994	Mch 1996
Central Govt	57,400	66,865	70,077
Local Govt*	13,900	16,037	16,341
Parastatals	12,900	12,937	13,718
B: Govt & Semi G	84,200	95,839	100,136
C = A – B	187,728	166,160	175,534
Cen & Survey ests (Not in ES)			
Domestics	31,617	26,000	21,000
Paid Empl-Trad Agr	20,000	13,000	16,000
D: Activ not Cov in Cen & Survs	51,617	39,000	37,000
E: Remain. Priv Sector (C-D)*	136,111	127,160	138,534
Employ Survey Ests			
Totally enumerated		61,964	63,351
Sampled			
Small		41,419	50,543
Unknown		31,164	19,983
Other		228	103
F:TOTAL	134,500	134,775	133,980
Inf Sector Paid Emply (E-F)*	1,611	-7,615	4,554*

* Drought relief work also to be deducted - LFS 15,000

11.8 Household Data

The total number of households for LFS estimated at 317,909 compares reasonably with the HIES figure of 291,610. However the average household occupancy of 4.25 persons does not compare well with a HIES figure of 4.56 and a 1991 Census figure of 4.73. HIES is expected to be less than the census as some residents do not stay long enough to complete diaries. For LFS there is no reason and interviewers must have not probed for all residents. This will have some effect on household tabulations and no adjustment is possible for persons missing. The household tabulations are still very useful, and can be used to provide social and Labour Market Information(LMI). Most variables of interest have been summed to household records to make tabulations with Xtable relatively easy. See file documentation file.

11.9 Quarterly Estimates

The table below gives the results by quarter for some of the major economic variables using WEIGHT3, the quarterly weight.

Table 10 : 1995/6 Labour Force Survey - Persons 12 years & over
A: Major Economic Activity Rates by Quarter

	Qtr 1	Qtr 2	Qtr 3	Qtr 4
Total	950862	950972	950784	950783
Under-Employed	43800	26478	21370	22001
Not Under-Employ.	328985	309202	341413	288612
Employ	372785	335680	362783	310613
Under Employ. Rate (%)	11.7	7.9	5.9	7.1
Unemp	194570	201190	161208	171558
Economically Activ	567355	536870	523991	482171
Not Econ. Active	383507	414102	426793	468612
Unemployed A	108161	105764	81360	80440
Unemployed B	86409	95426	79848	91118
Econ. Active Excl U/P B	480946	441444	444143	391053
Unempl. Total Rate	34.3	37.5	30.8	35.6
Unemployment Rate A	22.5	24.0	18.3	20.6
Unemployment Rate B	15.2	17.8	15.2	18.9
Partic. Rate (Incl A & B)	59.7	56.5	55.1	50.7

B: Sector in Employment by Quarter

	Qtr 1	Qtr 2	Qtr 3	Qtr 4
Total	372785	335680	362783	310613
C-Govt	76815	73990	78975	71033
L-Govt	55691	39100	36802	27320
Parast	15892	15728	6574	9602
Pvt Trad. Agric	38497	43348	73302	43267
Informal	70499	50754	57983	49295
Other-Private etc.	115391	112760	109147	110096

The rates in Part A of the table generally seems reasonable with Qtr 3 showing the lowest unemployment. This is the second quarter of 1996 and people were likely to be busy in lands etc. after good rains earlier in the year. However the last line of part A showing the participation rate falling throughout the period is worrying. It suggest interviewers gradually became less probing about economic activities as the survey continued.

Part B is interesting as it generally shows movements by sector as expected (e.g. Local Govt drought relief declining and agriculture increasing strongly in the third quarter). However it also shows the dangers of small samples in the Parastatal sector where the third and fourth quarter figures are definitely too low.

It is recommended that the quarter data only be used for aggregates of at least 30,000.