Documentation for the Mesebetsi Labour Force Survey User Files

Table of Contents

Contents of the CD

The Mesebetsi user files CD contains cleaned and edited versions of the survey's three data files with supporting documentation. The data are in the form of unaggregated microdata; each record represents a person, and the analyst has full freedom to recombine and recode variables as needed. The data files contain all of the original (un-recoded) responses of respondents, as well as data tracking and sampling information and a number of recoded variables to make analysis easier.

As supporting documentation the CD contains copies of the questionnaires, codebooks, a large excerpt from the Interviewer Field Manual with detailed explanations of the questions and response categories, and an excerpt describing the fieldwork process.

To make the CD contents as accessible as possible for a wide variety of users, most files are available in two or more different formats. The data files may be read as SPSS 9.0 data sets (compatible with SPSS releases prior to 7.5 and convertible to other statistical programmes), as bBASE IV files, or as tab-delimited ASCII files. Analysts should use or convert from the SPSS version if possible, as variable labels, missing values and other variable formatting information is already encoded. If a data set is built from the bBASE or ASCII versions of the files, please consult the appropriate Codebook to add variable formatting information.

Whenever possible, supporting documentation has been made available as -.pdf files which may be read with Adobe Acrobat Reader, and as plain ASCII text with line breaks (-.txt). However, in the case of the questionnaires, the interviewer manual and this overview document, formatting was too complex to render as plain text.

All Acrobat (.pdf) files on this CD are fully searchable. You can usually find information about a particular variable by typing its name (e.g. "b18") in Acrobat's search dialogue box. The program will search the entire document, including tables, but can not search for text in picture diagrams. You can also search for specific words, numbers or strings, such as "occupation," "10780" or "household income."

For users who do not already have Adobe Acrobat, an installation file for Acrobat Reader 5.05 is available on this CD. After reading the licensing agreement, start Windows Explorer and double-click on the file ar505enu.exe. The installation program will guide you through later steps. Version 5.05 will generally work on machines running Windows 95 or more recent operating systems. For Mac users and those with earlier operating systems, other versions of the Acrobat Reader can be downloaded at no charge from www.adobe.com.

The CD is organised as follows: (Table 1)

Table 1: CD Contents by Directory

CD Root Directory

File name	Title	Description /Notes
Intro and documentation for the user files.pdf	Documentation for the Mesebetsi Labour Force Survey User Files	Overview and technical documentation of files (this document)
README.TXT	README file	Gives basic information and instructions for opening documents with Acrobat Reader

SPSS 9.0 versions of data files

File name	Title	Description /Notes
Admin.sav	Administrative and Sampling variables data file	Note: All SPSS files are compatible with SPSS releases
Household Roster.sav	Household Questionnaire data file	prior to 7.5.
RSI.sav	Randomly-selected individual (RSI) data file	

dBASE IV versions of data files

File name	Title	Description /Notes
Admin.dbf	Administrative and Sampling variables data file	File includes variable labels
Household Roster.dbf	Household Questionnaire data file	
RSI.dbf	Randomly-selected individual (RSI) data file	

Tab-delimited ASCII versions of data files

File name	Title	Description /Notes
Admin.dat	Administrative and Sampling variables data file	Can be read into spreadsheet or database programs. File includes
Household Roster.dat	Household Questionnaire data file	variable labels, but value labels, missing values and other
RSI.dat	Randomly-selected individual (RSI) data file	formatting must be taken from the file Codebooks

(continued on next page)

Questionnaires and Fieldwork Report

File name	Title	Description /Notes
Household Questionnaire.pdf	Household Questionnaire	
RSI Questionnaire.pdf	Randomly-selected individual (RSI) questionnaire	
Fieldwork report excerpt.pdf Fieldwork report excerpt.txt	Selections from MarkData fieldwork report	Provides background on sampling, listing and field procedures

Codebooks and Interviewer Manual

File name	Title	Description /Notes
Admin file Codebook.pdf Admin file Codebook.txt	Mesebetsi Sampling and Administrative Variables File: Codebook	Variable by variable listing of file contents; contains variable and value labels, formats, missing
Household Roster Codebook.pdf Household Roster Codebook.txt	Mesebetsi Household Roster File: Codebook	values and other information. All codebooks are available as
RSI Codebook.pdf RSI Codebook.txt	Mesebetsi RSI File: Codebook	Acrobat files and plaintext files.
Interviewer Manual question descriptions. pdf	Interviewer Manual Excerpt: Chapter 5, Instructions for completing the Ques-tionnaires	Provides detailed definitions of terms used in the questionnaires, response categories, procedures for asking the questions and recording answers.

To Install Acrobat Reader 5.0

File name	Title	Description /Notes
ar505enu.exe	Compressed installation program for Acrobat Reader 5.05. Compatible with Windows 95, 95, NT, 2000, Millenium, XP operating systems	Acrobat Reader is needed to read many files on the CD. If you do not have Acrobat reader, double-click on this file in Windows Explorer and follow installation instructions.

A brief introduction to the Mesebetsi Survey

The need for reliable and updated labour market statistics has become critical for policy-making and monitoring in the new South Africa. With information from the Population Census of 1996, as well as new revised data collection systems, we have come quite far. However, there are still large unexplored territories in our knowledge concerning the labour market. Furthermore, academics and politicians alike regard some of the existing statistics, such as those on unemployment and wages, as controversial. The aim of The Mesebetsi labour force survey is to help fill some of the existing knowledge gaps (Mesebetsi means "work" in Sesotho).

The project was initiated in consultation with the Department of Labour in 1998. The Department is however nor accountable for the opinions expressed in the report, which are the responsibility of the Fafo research team. The survey was conceived with the overall aim of shedding light on important aspects of the labour market relating to the mandate of the Department of Labour that are not normally addressed in surveys undertaken by Statistics South Africa and others.

The survey was conducted with approximately 10 000 household interviews nationally at the end of 1999 and the beginning of 2000. The core section of the interviews focuses on employment, demographics, and wages and covers all household members, i.e. about 50 000 individuals. We then randomly selected one person of working age in each household for more in-depth interviews. Among other things, we asked them about work status, time period in (or outside) the labour market as well as basic conditions of employment, labour relations, wages, occupation and organisation of work for those in employment.

Fafo would like to thank the Norwegian Development Agency NORAD and South Africa Department of Labour for funds provided for the project. The Norwegian embassy in Pretoria gave us substantial assistance. We would also like to thank the senior Department of Labour officials who initiated and provided backup for the project, and gave substantial comments to the analysis. Finally, we thank MarkData Pty Ltd. for conducting the fieldwork.

We conducted in-depth consultations about the questionnaire, definitions and fieldwork strategy with academics, social partners in NEDLAC and the Department of Labour, among others. We thank all the individuals and constituencies, as well as the reference group set up for the survey, who generously put aside their time in order to provide comments and suggestions to the project.

The picture that emerges from our survey is of a South African labour market that in some respects resembles those of our neighbouring countries and com- parable transitional countries to a greater degree than what has previously been acknowledged. There are large numbers of people in underemployment, in informal sector activities and in subsistence activities: A large majority of the working age population is in the labour force, whether employed, or unemployed and available for work.

The extent to which the challenges and tensions in our labour markets, and the expectations of the large numbers of people inside the labour force, are addressed as a basis for collective bargaining and regulation will shape our growth and reconciliation. Our challenge is to jointly formulate a strategy that addresses the expectations and hopes of the millions who look to the labour market and the social partners for solutions to their needs. On that basis, we hope the following report, data set and upcoming results will contribute with useful information and perspectives for new debates and consensus-making.
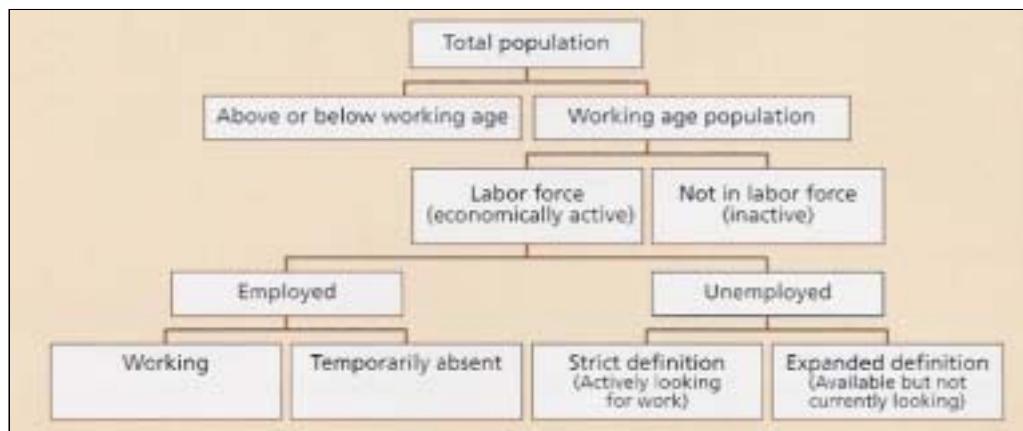
Labour Force Categories in the Mesebetsi

In this report we use the labour force categories developed by the International Labour Organisation (ILO, 1990). The working age population includes everyone age 15 or older, and in many analyses we focus on the pre-retirement age groups 15 to 65 years old. (An upper age limit of 65 was used to make findings more comparable to Statistics South Africa publications.)

A person is classified as employed if he or she worked for even one hour du- ring the seven days preceding the survey. Under the ILO system many kinds of subsistence work qualify as employment. In order to capture the full range of subsistence and informal employment the survey asked seven questions about different kinds of economic activity:

- Working for a wage, salary or payment in kind (such as food or housing) .Engaging in any kind of self-employment activity

- Doing unpaid work in any kind of family business

- Working on your own or the family's farm, garden, cattle post or kraal .Making anything from farm products or natural products, either for sale or for home use

- Catching or collecting fish or other food either for sale or family food .Doing construction or major repair work on your own house, business or farm.

People who engaged in any of these activities for at least one hour during the reference week were classified as employed. The survey also asked whether the person spent time fetching wood or water during the reference week .Although these activities do contribute to the household economy, they were not classified as employment unless done for other households for pay.

Figure 1: ILO Labour Force Classification

The **labour force**, or economically active population, includes everyone in the working-aged population who is either employed (working or temporarily absent from a job at the time of the survey), or unemployed. All other working- aged people are classified as **out of the labour force**; those are persons who do not want to work or are not available for work. Common examples are non- working retired people, full-time students, homemakers, the disabled, and

people of independent means who do not need to work. Others may refuse to work under the prevailing wages or working conditions.

The **unemployed** are those who did not work at all during the reference period but would be able to start working in the next four weeks if suitable work were available. For some purposes the unemployed are further divided into two groups. People who are available and who have taken active steps to find work in the past four weeks are said to be unemployed under the **strict definition**. This is the version South Africa's uses for its official unemployment rates. The **expanded definition** of unemployment includes those who are actively looking, as well as those who are available for work but have not taken definite steps to find work during the reference period. This latter group is sometimes called 'discouraged workers.'

For more information about Mesebetsi field procedures, please see Mark Data's *Field Report* in the Questionnaires and Field Report directory of this CD.

Other background information about the survey, and a discussion of its main findings, can be found in ***The Mesebetsi Labour Force survey Topline Report* (Liv Tørres, ed.) Oslo: Fafo Report 360, 2001.**

The report is available from Fafo offices in Norway and Johannesburg:

Institute of Applied Social Science
P.O. Box 2947 Tøyen
N-0608 Oslo, Norway
HTTP://www.fafo.no

Fafo South Africa:
PO Box 412 766
Craighall Park 2024, RSA
info@fafo.org.za
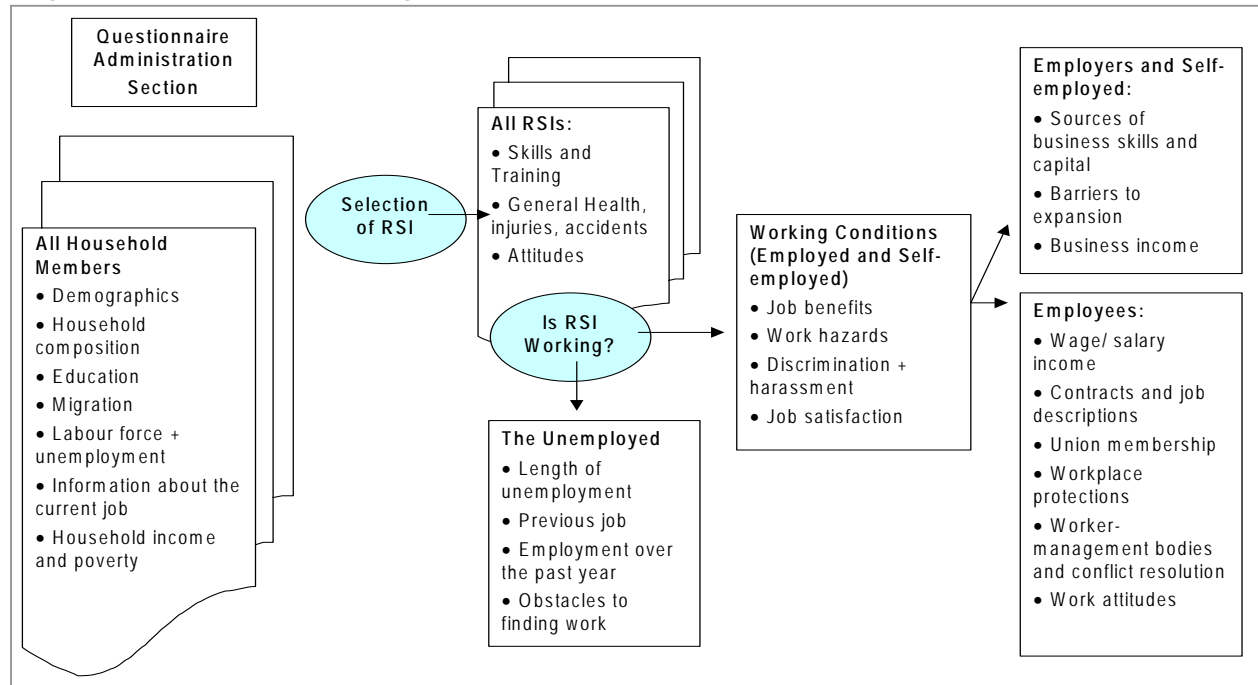
The Mesebetsi Questionnaires

The survey instrument was developed by Fafo over an 18-month period in consultation with the SA Department of Labour, Statistics South Africa, and many other governmental and non-governmental user groups. Its two main components are the Household Questionnaire, completed for each member of the sample household, and the RSI Questionnaire, answered by one randomly-selected individual (the **RSI**) in the household. The two questionnaires were bound into a single document for fieldwork.

- **The Cover Page**: Field workers used this form to identify and track the questionnaire, and to record visits to the household and later coding and editing

- **The Household Questionnaire:** In this part of the questionnaire, you ask information about all household members.

  o   On the **Flap page,** you will begin by listing all members of the household, and selecting the RSI, or randomly-selected individual.

  o   In Section A, **Demographic information**, you will ask questions on the age, marital status, education, migration and other characteristics for everyone in the household.

  o   Section B, **Employment and unemployment**, is asked only for household members aged 15 and above. It covers the different kinds of work that people have done in the last week, and some information about each person's working hours, occupations, and similar things. There is also a set of questions for people who are not working.

  o   Section C, **Questions for the Household as a whole**, contains questions about the main language used at home, type of dwelling, and household income.

- **The RSI Questionnaire** was asked only for the randomly-selected member of the household. It goes into more detail about the person's work or unemployment situation, and attitudes. All RSIs answered sections D through F, dealing with education skills and training, opinions about a number of work-related public issues, and person's health status. Beyond that point skip instructions directed interviewers to ask sections G through J as appropriate, depending on the respondent's current employment status and type of work in his or her main job. All section-skip instructions were keyed to specific answers given on the Household Questionnaire, to ensure that the population answering each section would be rigorously defined. The main sections were as follows:

  o   **Section D: Skills and training** (all RSIs).
  o   **Section E: Attitudes** (all RSIs).
  o   **Section F: Health** (all RSIs).
  o   **Section G: Unemployed or not Economically Active** (RSIs who are not working).
  o   **Section H: Working conditions** (RSIs who work for themselves <u>or</u> someone else).
  o   **Section I: Own-account workers and employers** (RSIs who work for themselves).

o **Section J: Employees only** (RSIs who work for someone else).

Figure 2 below illustrates the basic sequence of questioning during the interview, the main section-level skips employed, and the main topics covered. Table 5 (File and Questionnaire Match), which appears in a later section of this document, gives a more detailed breakdown of where the various groups of questionnaire items can be found on the user data files. For a variable-by-variable listing of file contents, consult the Admin, Roster and RSI codebooks.

Figure 2: Basic flow diagram of the Questionnaires

Overview of the User Files

The files currently in the data set are as follows:

SPSS (version 9.0) files: Mesebetsi Admin.sav, Mesebetsi Household Roster.sav, and Mesebetsi RSI.sav, with corresponding files in other formats. Files were created 30 May, 2000.

Table 2: Overview of the user files

File	Contents
Admin.sav Admin.dbf Admin.dat records = 50 949	• Contains all administrative and sampling items, along with Flap page data. • Covers all sampled units, including non-interviews and partial interviews, and all persons, including visitors. • Records are persons. To convert to a file in which each record represents one household, select if Rosnum=1. This file is not used for ordinary analyses.
Household Roster.sav (-.dbf / -.dat) records = 50 367 (10 815 households)	Contains code demographic and labour force information for all household members, (Flap page, Sections A and B) plus: • Variables relating to income, the dwelling and language, for the household as a whole, from Section C; • Created household-level variables; • Created person-level variables for each household member; Each record represents a person in the household. Visitors to the household are included, but only have Flap Page information. The file only contains complete interviews (ar2 interview status =1 or 2).
RSI.sav RSI.dbf RSI.dat records = 10 780	Extended questionnaire information for one randomly-selected household member age 18 and above. Contains: • All RSI questionnaire items • The RSI's data record from the Household Roster (Flap page, Sections A through C) • Household-level created variables from the Roster. • Person-level created variables derived from Roster items and the RSI questionnaire. This file only contains complete interviews (ar2 interview status =1 or 2). In 35 cases the RSI was not selected according to the correct procedure. These records were dropped from the RSI file, but the household data is valid and remains in the Roster.

Weights

All files contain statistical weights that are used to adjust for sampling design effects such as the deliberate over-sampling of people in less populous provinces. Analyses must be weighted with the correct weight to ensure accurate results.

There are four weights that may be used, depending on the aim of the analysis. These are **expweig2, expweig3, relweig2 and relweig3**. The *expansion weights* produce estimates weighted up to the South African population. The *relative weights* are created so that their mean is 1.0. Thus means and percentages will be the same as with the expansion weights, but Counts will equal the number of records in the file, not numbers in the population.

The numbering pertains to the sampling stage. Thus the relweig2 and expweig2 weights relate to households, and relweig3 and expweig3 relate to the RSI. Use the -2 weights to analyse household Roster data, and data in the RSI file when the unit of analysis is the household as a whole (examples: household size, households dwelling type). The -3 weights are appropriate for most analyses of the RSI as an individual.

The weights are present in the files as given below:

Table 3: Weights

File	Expweig2/ Relweig2	Expweig3/ Relweig3
Admin.sav (.dbf, .dat)	Present	Present
Household Roster.sav (.dbf, dat)	Present	Absent
RSI.sav (.dbf, .dat)	Present, but only meaningful for the household section (C) and household-level created variables	Present

Record identifiers

General identifiers

The household-level record identifier is **RECNUM**. All persons in a given household will have the same value for RECNUM. The person-level identifier is **ROSNUM**. As appropriate, files may usually be merged by those variables. All files are sorted by RECNUM and ROSNUM.

The variable **AC3** identifies the roster number of the person chosen as RSI in the household. The corresponding variable in the RSI file is **IC1**.

RSI identifier:

The Household Roster and Admin files also contain the variable **RSI**, which can be used to identify or select RSIs.

The RSI variable distinguishes RSIs from other household members, and also identifies RSI cases with missing or incomplete records. Coding is as follows:

> RSI= 0: Person was not chosen as RSI.

> RSI=1: Person is RSI, and completed the RSI questionnaire. This is the population of cases that appears in the RSI.sav file. Select for RSI=1 when selecting cases from the Household Roster or Admin file to be merged with the RSI file.

> RSI=8: The person was chosen as RSI, but the entire household was classified as a non-interview (31 cases). These cases appear in the Admin file only.

> RSI= 9: The RSI was chosen incorrectly (35 cases). The person's Household Roster information is valid, but the household is not represented in the RSI file.

Table 4: Identifiers

File	Household Identifier	Person identifier within household	RSI Identifier
Admin.sav	RECNUM	ROSNUM	RSI
Household Roster.sav	RECNUM	ROSNUM	RSI
RSI.sav	RECNUM	ROSNUM	(all records are RSI)

Coverage of the files in relation to the questionnaire

Table 5: File and questionnaire match

Section	Topic	Level	File
AI to AC	Administration and sampling information	Generally household level	Admin.sav (selected variables appear in Household Roster + RSI.sav).
Flap page	Household composition, age, sex	Individual HH members	Household Roster.sav / RSI.sav (RSI records only)
A	Demographics + education	Individual HH members	Household Roster.sav / RSI.sav (RSI recs only)
B	Labour force	Individual HH members	Household Roster.sav / RSI.sav (RSI recs only)
C	Housing + Income	Household	Household Roster.sav / RSI.sav
D	Skills & Training	All RSIs	RSI.sav
E	Attitudes	All RSIs	RSI.sav
F	Health	All RSIs	RSI.sav
G	Unemployment	Unemployed/ Out of LF RSIs	RSI.sav
H	Working Conditions	All working RSIs	RSI.sav
I	Own-account Workers & Employers	Self-employed RSIs	RSI.sav
J	Employees Only	RSIs working for someone else	RSI.sav

Variables added to the files:

Table 6   Variables created in Household Roster.sav and RSI.sav (In Alphabetical order)

Name	Contents	Comments	In files
allind	Combines 1-digit industry codes for employed + unemployed	Current job from b21ind1 / previous job from G4ind1; or never worked.	RSI only
allocc	Combines 1-digit occupation codes for employed + unemployed	Current job from b18occ1 / previous job from G3occ1; or never worked. B7 only cases are put in separate category.	RSI only
allstat	Combines 1-digit employment status codes for employed + unemployed (simplified)	Current job from b19 + b23 / previous job from G5; or never worked.	RSI only
b18occ1	Recode of present occupation	9 ISCO major groups	Roster+ RSI
b21ind1	Recode of present industry	15 ISIC major divisions	Roster+ RSI
B7only	Person's only activity during the reference week was fetching wood/ water 1= wood, water was only activity (age 15+) 0= All other cases (including children < age 15)	Useful for removing this marginally active group from analyses. E.g. select out when analysing RSI sections H to J, and in all analyses that rely on mainstream ILO definitions of economic activity. *See also:* "Coding for subsistence workers" at end of this document.	Roster+ RSI
citizen	S. African / non- S. African	Recode of A12	Roster+ RSI
EAtype5	5-way classification of EA based on 1996 Census (urban formal, rural traditional etc.)	Uninhabited and institutional EAs in 1996 are only coded urban or rural, as determined by field staff at time of survey	Admin, Roster+ RSI
educat	Highest education level completed	Recode of A4 into 5 categories	Roster+ RSI
expanded	Labour force status according to ILO expanded definition (available but not actively seeking work)	Working / unemployed / out of labour force	Roster+ RSI
g1dur	Duration of time since last worked	Unit is months; applies to RSIs not working during reference period. B7only cases not included.	RSI only
g3occ1	Recode of most recent occupation (unemployed)	9 ISCO major groups	RSI only

g4ind1	Recode of most recent industry (unemployed)	16 ISIC major divisions	RSI only
g7anywk	Months worked at all from Oct 1998 to Sept. 1999.	Sum of g7ptyr1 to g7ptyr3.	RSI only
g7ptyr1 to g7ptyr5	Measures of intermittent employment over a recent 12-month period	Number of months worked (etc.) from Oct 1998 to Sept. 1999; one variable for each G7 category.	RSI only
hhempl	Number of employed people in household	Based on lf_ilo (wood and water only is NOT employed)	Roster+ RSI
hhinc	Total HH income in past month	Based on C318	Roster+ RSI
hhkids	Number of people in household below age 15		Roster+ RSI
hhmaxed	Highest education of a HH member (excludes non-relatives)	Based on variable educat	Roster+ RSI
hhmigrs	Number of 5-year migrants age 15+ in household	Based on created variable Migrant	Roster+ RSI
hhold	Number of people in household age 65 or older		Roster+ RSI
hhoutlf	Number of HH members age 15+ not in Labour Force	Based on lf_ilo	Roster+ RSI
hhquin	Household income (quintile)	Based on C318	Roster+ RSI
hhsize	Number of permanent residents in household		Roster+ RSI
hhtype / hhtype2	Household type (based on A1)	hhtype= 9 categories; hhtype2= 5 categories	Roster+ RSI
hhunempl	Number of unemployed persons (expanded definition) in HH	Based on lf_ilo	Roster+ RSI
hrwage	Hourly wage or work earnings	Based on incrsi and b28	RSI only
incempl	RSI monthly income in the main job: only covers cases which were wage or salary employees	Based on J1_1 and J1_2. Time unit standardised to monthly.	RSI only
incquin	Work earnings quintiles	Based on incrsi	RSI only
incrsi	RSI monthly income in main job; includes both wage/salary and self-employed cases.	Covers all RSIs with work income. Time unit standardised to monthly.	RSI only
incself	RSI monthly income in the main job: only covers cases which were self-employed	Based on I8_1 and I8_2. Time unit standardised to monthly.	RSI only
inctype1 - inctype4	Unified income source variables, combining b16a-c and b36a-c	Must analyse using multiple response procedures.	Roster+ RSI
informal	current job is in informal sector	Classification based on	Roster+ RSI

		enterprise size, registration, accounts and UIF contributions. *For full definition, see* "Informal sector" at end of this document.	
lf_ilo	ILO labour force status: Employed/ Unemployed strict/ Unemployed expanded/ out of LF.	Wood/ water fetching (B7 only) is NOT counted as economically active.	Roster+ RSI
lf_iloww	ILO labour force status	Wood + water (B7) is counted as economic activity.	Roster+ RSI
LFP_EXP	Labour force participant according to ILO expanded definition	0= out of LF 1= L Force participant	Roster+ RSI
LFP_STR	Labour force participant according to ILO strict definition	0= out of LF 1= L Force participant	Roster+ RSI
migrant	Person moved province/ country or urban/ rural in past 5 years	Based on A9, A10 and present location information	Roster+ RSI
migrprov	Person moved province/ country in past 5 years	Based on A9, A10 and present location information	Roster+ RSI
migrurb	Person moved rural-urban or urban-rural in past 5 years	Based on A9, A10 and present location information	Roster+ RSI
pctmigrs	Percent of HH members age 15+ who are 5-year migrants	Based on hhsize and migrant	Roster+ RSI
province	Province		Roster+ RSI
RSI	RSI indicator	1= person chosen as RSI; 0= all others; 8,9= incomplete	Admin + Roster
RSIkid6	RSI has one or more own children age 6 or less in HH.	Often an important factor in LF participation for women.	Roster+ RSI
rurban	EA classified as urban / non-urban in 1996 Census	About 15 EAs were uninhabited in 1996; these were classified by MarkData during fieldwork.	Admin, Roster+ RSI
strict	Labour force status according to ILO strict definition (actively seeking)	Working / unemployed / out of labour force	Roster+ RSI
uelayer	All adults are classified in a continuum of employment, under- and unemployment statuses	cf Topline report fig 15.	Roster+ RSI
voced	Person has vocational diploma/ certificate or professional degree	Based on A4 and A5	Roster+ RSI

Other Notes for Data Users

Variable naming conventions

Most variables in the user files are named after the questionnaire item they represent. When the questionnaire item contains several sub-items under the same question number, the corresponding variable name will have a designation such as **b29b, c32_c, a11b4**, and so on. Flap Page variable names begin with '**fp**'. Generally the correspondence between variable names and the questionnaire is quite clear and straightforward.

Many of the created variables also contain information about the questionnaire variable they were derived from: e.g. **b18occ1** (single-digit ISCO coding of b18 occupation), **g1dur** (duration of unemployment, based on variables g1m and g1y), **B7only**, etc. Created variables which are not closely based on a single questionnaire item will generally have mnemonic names such as **lf_ilo, migrant, hrwage, rurban**. Whenever possible, created variables which refer to the household as a unit have been given names beginning with 'hh:' **hhsize, hhunempl, hhinc**, and so forth.

SPSS Measurement level

SPSS data files define the measurement level of each variable as scale, ordinal, or nominal. Although measurement level information will appear in the codebook, it is *not* accurate and should not be relied upon. It will have no effect on analyses. The measurement level of Mesebetsi variables is usually obvious from the question and coding categories, and users may re-define this parameter if they choose.

Multiple response variable groups

The Mesebetsi contains a number of multiple-response variable groups. These are derived from 'Check all that apply' and similar items that allow several answers to a question. Examples include **b13a, b13b, b13c** (actions taken to find work), **inctype1** to **inctype4**, and others. To get accurate results, variables of this type *must* be analysed using SPSS Multiple Response or similar procedures. Often the variable label will signal a group variable, by using wording like 'Actions taken to find work1', but that is not an infallible guide. Before doing an analysis, always check the original questionnaire wording to make sure you understand the question's wording and context.

Household members and visitors

Variables FPC and FPD of the household Roster file provide information about household membership. A **household member** is defined someone who has stayed with the household for at least 4 nights per week over the past four weeks (FPD=1), plus people who normally stay with the household but are currently in hospital or boarding school (FPD=2). All others (FPD=3) are not considered household members. The data files contain age and sex information for non-members staying with the household at the time of the survey, but are missing on all other variables. They were not eligible to be selected as RSIs.

The household Roster file contains 2432 people who are not household members, a number that is large enough to affect the age and sex distribution. They will be reported as system-missing on

most cross-tabulations, but should be selected out when running univariate frequencies on age and sex unless the analysis is specifically interested in non-household members.

Variable FPC asks if the person is a resident or visitor. There are 442 visitors in the file. Note that a visitor <u>is</u> counted as a household member if he or she meets the criteria above.

Informal Sector

The variable **informal** allows analysts to identify people who work in non-agricultural informal sector jobs as their main activity. In line with current ILO recommendations, the Mesebetsi's definition of the informal sector incorporates information about the size of the enterprise, formal registration as a business, record-keeping practices as its primary criteria. Specifically, a person is classified as informal sector if:

* The enterprise has less than 20 employees, and
* Is not registered or incorporated as a business, and
* Does not keep a set of accounts separate from the household budget.

Employees of small enterprises often had trouble answering the question about accounts. When this occurred, the enterprise was classified as informal if:

* It did not make unemployment insurance fund (UIF) payments, and/or
* It was located on a footpath, street, open space or market.

Keep in mind that informal sector work does not necessarily imply self-employment. As defined in the Mesebetsi, an informal sector worker may be a wage employee, an own-account worker, or an employer of others.

Different researchers may want to set different bounds on the scope of informal work. Accordingly, the variable is designed to capture other groups of interest as separate categories:

```
Value    Label
    0    Non-agric Formal sector
    1    Informal sector (non-agric) excluding domestic workers
    2    Domestic workers
    5    Agric, fishing & forestry workers
    6    Other subsistence work
  7 M    Wood + water only during ref week
  8 M    Other not working
  9 M    Unable to classify
```

The other categories are strightforward except for group 6, 'Other subsistence work'. These respondents are working on own farms for household consumption, collecting natural products, building their own houses, and similar activities. Agricultural, fishing and forestry workers may be working on their own account (with at least some production for the market), or they may be employees of formal sector farms.

Coding for Subsistence workers

The survey captured a good deal of economic subsistence activity which does not fit easily into the 1990 ISIC and ICSO-88 industry and occupation classifications.[1]

In asking about occupation and industry, the survey asked whether agricultural production was mostly for sale or home use. Many people working on their own farms or gardens answered that it was primarily for home use. Subsistence agriculture of this kind is defined as employment by the ILO criteria. Subsistence-oriented farmers and gardeners generally were assigned b18 occupation code 621 (subsistence agricultural & fishery workers) or b21 industry code 951 (own household subsistence activities) to distinguish them from more market-oriented informal agriculture.

Fetching wood or water for household use

was recorded as an economic activity in question B7, as it falls within the System of National Accounts production boundaries. However, people who <u>only</u> fetched wood and water during the reference week are not classified as being employed in the data set's main labour status variable (created variable LF_ILO). The file does contain an alternate measure of employment status which counts wood and water as employment (LF_ILOWW). This variable is not recommended as a general unemployment measure, as wood and water fetching is basically a household reproduction activity.

> However, analysts should keep in mind that most wood and water fetchers are members of rural subsistence households and may engage in gardening, caring for animals, collecting natural products and similar activities at other times, even if they did not do those things during the reference week.

> **We recommend that users remove wood and water fetchers from most analyses of employment. These cases can be easily identified using the variable B7only: B7only equals 1 if wood and water fetching for the household was the person's only activity during the reference week.**
>
> **Use the SPSS menus or syntax to select cases where B7only=0. That eliminates wood and water-only cases (n=2288 in the Roster/ 745 cases in the RSI file, unweighted). This applies to both the Roster and the RSI files.**

- If the person was collecting wood or water as a form of market-oriented self-employment or as a wage employee, he or she will *not* be included in the B7-only group. They will be recorded as employees or self-employed (B1=1 or B8=1), and B7only=0.

- Although we do not recommend that they be included in most analyses, information on occupation, industry and employment status information is recorded for wood and water fetchers in items b18 to b23. In most cases they are coded as follows:

  - b18 occupation = 621, Subsistence agricultural & fishery workers
  - b18occ1 = 9, Elementary occupations.

---

[1] References: United Nations Dept. of International Economic and Social Affairs, *International Standard Industrial Classification of all Economic Activities,* Statistical papers series M, No. 4, Rev. 3, New York: UN, 1990. International Labour Office, *ISCO-88: International Standard Classification of Occupations*, Geneva: ILO 1990.

- b21 industry = 951, Own household subsistence activities (new code, not in ISIC)
- b21ind1= 16, Own household subsistence activities (new code, not in ISIC)
- b19 employment status = 4, Helping out in family business with no pay
- b23 ownership status = 7, Private business or self-employment.

The codes are not always ideal, as the 1990 ISIC and ISCO-88 classifications do not make specific provision for this kind of activity. Regarding b23 and b19, wood and water fetchers are classified as working in a family enterprise without pay. Many subsistence households do sell or trade surplus at some times during the year, but as enterprises they are marginal

In a few cases the codes recorded for b21 industry and b18 occupation have no connection to wood and water gathering. These are not errors in the data. Rather, the person's usual work was reported as school teaching, mine labouring, farming etc., but the person did not work at these jobs during the reference week. They may or may not be listed as formally absent from work in variable b9, as they did not always have a definite job to return to.

In some non-rural cases a wood and water fetchers' occupation was recorded as 913 (domestic helper) because there were no agricultural subsistence activities. In such cases industry will be recorded as 951 (Own household subsistence). Domestic *employees* like maids are also coded as b18 occupation=913, but b21industry =950 (private household), b19 employment status=1 (employee), and b23ownership status=6 (private household). B23 category 6 is used only for private households with wage employees.