UNITED NATIONS SECRETARIAT   ESA/STAT/AC.79/1
Statistics Division   9 October 2000

                                              English only

_____

Gender issues in the measurement of paid and unpaid work

Expert Group Meeting on
Methods for Conducting Time-Use Surveys
23-27 October 2000
New York

Country Paper: SOUTH AFRICA
Time Use Survey 2000

COUNTRY REPORT: SOUTH AFRICA
TIME USE SURVEY 2000
Prepared by Yandiswa Mpetsheni and Debbie Budlender, Statistics SA
Expert Group Meeting on Methods for Conducting Time-use Surveys
New York 23-27 October, 2000

How data collection on time-use fit into the national statistical system

a.   Is this the first national time use survey for the country
The Time Use Survey (TUS) of 2000 is the first of its kind at a national level for South Africa.

b.   How time use survey is related to other Stats SA surveys
The time use survey was run separately, but has strong links with other surveys conducted by Statistics South Africa. Standard questions found in other Statistics South Africa questionnaires form part of the questions in the time use questionnaire. These include household demographic questions as well as questions on labour force participation, industry and occupation. These questions were included so as to allow comparison of TUS results with those obtained from other sources of information on labour force participation.

c.   Major factors that led to the decision to collect nationally representative data on time use
One important motivation for the study was the need for valuation of unpaid labour. Thus one part of the analysis of the TUS will involve an attempt to incorporate the data into satellite accounts. A second strong motivation was the hope that the TUS would provide more in-depth information on the informal economy and other under-counted forms of economic activity. A third important motivation is to inform government policy more broadly across a range of sectors.

d.   Plans in the next five years
The TUS is an expensive operation in terms of financial, human and other resources. Stats SA was able to undertake the current survey because we received generous assistance from the Norwegian Agency for Development Cooperation. For the future, we hope to attach a time use module on a five-yearly basis to our six-monthly labour force survey. The time use module will be one among a number of rotating modules which will be used to monitor issues which do not change dramatically from year to year.

2   Survey objectives

a.   Specific objectives of the TUS
- To measure and analyse the time spent from day to day by different individuals.
- To provide new information on the division of both paid and unpaid labour between women and men and other groupings
- To incorporate unpaid work in satellite accounts
- To gain more insight on the reproductive and leisure activities of household members

- To gain more understanding of produc tive activities such as subsistence work, casual work and work in the informal sector.

b.   Expected users

The expected users of the survey results are government departments, university departments as well as non-governmental organisations. All of these sectors are represented in the time use reference group, which was formed to be an advisory committee for the time use team.

The study will deepen policy makers' understanding of the economic and social well being of different groups, and how government and others can assist in meeting needs. Through the study, the general public will be able to quantify time spent on different activities.

c.   Main findings

The time use survey is being conducted in three tranches so as to capture seasonal variation. The firs t tranche was conducted in February 2000 and the second in June 2000. In October 2000 we will be conducting the last and final tranche of the survey. Full analysis of all three tranches will be undertaken early in 2001. The only results currently available are of the pilot survey that was conducted in November 1999.

3   Design specifications of survey and factors considered in design decisions

a.   Relating to survey objectives, design and resources

i.   TUS designed as an independent survey or a module in a core survey

The survey was designed as an independent survey because it was being done for the first time in the country. The survey is national and the sample design attempted to ensure adequate coverage of different types of enumeration areas in each of the nine provinces. Time use data is collected from two, randomly selected, individuals from each household, aged ten years of older. The survey was conducted in three tranches at different times of the year.

ii.   Why choose this particular survey design

The study was designed as an independent survey because it was an experimental pilot. The coverage is national because we want to know about, and compare, the situation of all South Africans. The survey is confined to people ten years and older as this age group was felt competent to answer the questions without special adaptation for children. The three-tranche approach was adopted so as to capture possible seasonal variations in activity.

iii.   Total survey costs

Total survey costs can only be given after fieldwork is complete, the data has been analysed and reports published.

b.   Method of data collection

Fieldworkers administered face-to-face questionnaires, which included a diary, to respondents. The questionnaire had basic demographic and employment questions common to Statistics South Africa questionnaires. The questionnaire was divided into five sections. The first section dealt with household information. This section was

administered to a responsible adult in each selected household. After completion of this section, the fieldworker used a selection grid to choose two individuals aged 10 years and older. The fieldworkers then interviewed the first selected person using sections 3 and 4 and the second selected person using sections 4 and 5. Sections 2 and 4 contain individual demographic questions and sections 3 and 5 are 24-hour diaries.

Respondents were asked about activities they were engaged in from 4am the day before the interview to 4am on the day of the interview. The diary had a fixed time interval of 30 minutes. Respondents could name up to three activities per time slot. The activities recorded were not prioritised.

d.   Collection of information on simultaneous activities
The survey collected information on whether activities were conducted simultaneously or not. Because the survey is not yet complete, information on the proportion of diaries reporting simultaneous activities is not available. Information on the five most common combinations of simultaneous activities is also not available. According to the pilot survey, listening to the radio, supervising people needing care, socialising with friends and family, cultural activity, watching television and eating were all commonly conducted simultaneously with other activities.

e. Collection of information on context variables
The context variables on which information was collected were location and purpose of travel. Location was collected in two variables. The first variable records where respondents were when they were conducting their activities. The second variable records whether respondents were inside or outside or, if travelling, the mode of transport.

f.   Time sample
All days of the week were represented. To try and get an even distribution of all days of the week, fieldworkers were given a minimum and maximum number of diaries to complete each day. Fieldworkers were also required to work on each weekend day and were to take their 'weekend' on Tuesday, Wednesday, Thursday or Friday. These rules attempted to ensure adequate information on the days which studies in other countries have found to show less 'normal' patterns. The actual spread of days will only be established when all time use tranches have been analysed. The second tranche included a public holiday which fell on a Friday. We did not give any special instructions about this day.

To cater for seasonal variations, the TUS was done in three tranches, the first one was in February, the second one in June and the last one in October.

g. Considerations in defining and sampling the reference population

i. Considerations in defining the reference population
As in all Stats SA surveys, the reference population was sampled from four strata which characterise the South African population. The strata comprise formal urban areas, informal urban areas, commercial farms and other rural areas. For the household section only responsible adults were interviewed while for the diary and demographic questionnaire respondents were 10 years and older.

ii.   Selection of sampled households

Dwelling units in all provinces were systematically sampled. In each primary sampling unit a total of 12 dwelling units were sampled. In cases where there were multiple households in a dwelling unit, all those households were interviewed. The database for sampling of households was the same as used in the Survey of activities of young people conducted during 1999. The database was constructed in May 1999 through listing of all the dwelling units in the selected primary sampling units. For each tranche of the TUS, the provincial coordinators were asked to update the listings in areas where significant changes were expcted. The main motivation for updates was the mobility of informal settlement dwellers.

iii.   Selection of household members

A selection grid was used to sample two people from among all household members who were 10 years and older. The selection grid allowed for random sampling.

iv.   Total sample size

The planned total sample size was 10 800 dwelling units. Each dwelling unit could contain more than one household. Within each household, two respondents were interviewed unless there was only one member aged 10 years and above. The exact number of diaries is not yet available and will be affected by multiple households, the number of members of eligible age, plus non-response and non-contact.

4   Activity classification

The trial UN classification of activities as amended by the expert group was used. This classification was adapted slightly to match local conditions but the basic structure was left unaltered. In particular, the first three categories between them account for virtually all work activities included in the System of National Accounts (SNA). An activity code index was developed in addition to the activity list. The activity list contained our coding schema in order of code, i.e numerical order under the ten international categories. The activity index contained activities listed in alphabetical order, with corresponding code alongside. The number of potential activities in a coding index is virtually infinite. After constructing a basic list from our first focus group and behind the glass tests, we added all the activities reported during the pilot which preceded our survey. To this basic list we added a few activities which we thought might cause particula r problems.

5   Assessment of design and implementation of field operations for time use data collection

i.   Procedures specific to time use survey

As in most surveys conducted by Statistics South Africa, fieldworkers administered the time use questionnaire rather than allowing respondents to complete the schedules themselves. This helps a lot in terms of standardisation as we are relying on trained fieldworkers rather than untrained respondents. The face-to-face approach also addresses the issue of illiteracy, as respondents are not completing the questionnaires themselves. Children who were interviewed were 10 years and older. Pre-tests suggested that at this age we did not require any special method of collecting data from them.

Respondents, even those who were illiterate, seemed to be attuned to the concept of "clock-time". What made this easier is the 30-minute interval that was used instead of the 10-minute or 15-minute interval used in developed countries.

ii.   Incentives
None of our respondents are given monetary or in-kind incentives to cooperate. Nevertheless, our response rate is very high by international standards. For the first time use tranche the response rate was 84%.

iii.   Average length of an interview
The average length of the interview was less than an hour but the exact minutes can only be provided after data analysis is completed.

iv.   Callbacks
If respondents were not available, fieldworkers were asked to make three callbacks for the household part of the questionnaire, three visits for the demographic and diary part for the first person and three visits for the demographic and diary part for the second person. A total of nine visits were thus possible per questionnaire. If after all the callbacks a fieldworker still could not find a particular respondent, that would be regarded as a non-contact.

v.   Substitution of households and respondents
Substitution of households or respondents was not allowed, as that would have compromised the representivity of the sample.

b.   Recruitment and training of field staff
Fieldworkers were recruited in all our 9 provinces. The number of fieldworkers per province was dependent on the number of primary sampling units sampled in that province. The minimum qualification required of fieldworkers was matric/grade 12. All fieldworkers wrote a competency exercise and only those with high marks were selected. An extra fieldworker was sent from each province for the first training so that we would have an extra person to be a substitute if the need arose.

All fieldstaff, fieldworkers and provincial time use coordinators were trained at head office for the first tranche of the time use survey. The motivation for the head office training was that this was the first time such a survey was being done in South Africa and there was thus no experience. We therefore wanted everyone to have a common understanding of time use concepts.

The trainers were Stats SA staff and a visiting consultant from Norway, Gustav Haraldsen. Trainees were divided into five groups according to provinces. In all the five training rooms the questionnaire was translated into local languages used in particular provinces. We did this as we did not want the to lose the essence of the questions when they were translated into the different local languages during the interview. The normal practice in Statistics South Africa is to have English questions on the questionnaires and fieldworkers are expected to translate in the field when interviewing someone whose language of choice is not English. The responses are recorded in English.

At the end of fieldworker training supervisors were chosen and trained. All fieldworkers worked in teams of three or four under a supervisor. Fieldwork in the first tranche was from 7-29 February 2000. Fieldwork was hampered in several of the provinces by exceptionally heavy floods.

For the second tranche, the training at head office was for provincial coordinators and for the few fieldworkers who were replacing those who for some reason could not be involved in the second tranche. This training lasted two days. A further two-day refresher training was conducted in each province for both new and old fieldworkers. Fieldwork for the second tranche was from 5-25 June 2000.

c.   Field coding and field editing of time use activities
Fieldworkers did their own coding of both activity and location. Location was coded during the interview. The activity coding was done in the evening of the interview while details were still fresh in the fieldworker's memory. Some fieldworkers did the coding in teams after each day's work while others did it individually on the evening of the interview. Coding was then checked first by the supervisors and subsequently by the provincial coordinators.

6 Assessment of design and implementation of data processing of time-use data

a.   Organization and staffing of data processing, including:

i.   Development of special computer programmes
As with other surveys, Stats SA staff wrote the programs to capture the data. These tailor-made programs are written in Visual Basic. The data is captured into SyBase and converted to SAS format. We wrote the editing and cleaning programmes in SAS and will also use SAS for analysis of the data. Most of the programming for analysis will be done by gender unit staff.

b.   Coding, editing and imputation

ii.   Common errors in coding
The common errors reported here were those that were picked up in data from the first tranche. Each of these errors was addressed in the training for the second tranche, with the hope that error rates would decrease.

Many of the identified errors occurred in the household and demographic sections of the questionnaire rather than the diary. The work-related questions, in particular, posed problems. For example, in the household questionnaire the respondent would give, as the main source of income, a source not mentioned among 'all sources'. In the demographic questionnaire, a respondent would mention having done some work in the past seven days, but later say they had never worked. This was particularly common when the work was subsistence activity. There were also some confusion in how to answer questions in respect of paid domestic work. And there were some people who said they had their own business but gave the source of income as wage or salary. There were sometimes other inconsistencies between responses to work in past seven days, occupation, industry and status in employment.

A common problem in respect of the diaries was the inclusion of 'instantaneous' happenings, such as 'leaving', 'arriving' and 'waking up' as activities. We stressed in the training that these were not valid activities, but they were nevertheless included in the diaries – often without any other activity in the timeslot. The discussion below on imputation describes how we dealt with this.

Another common problem was in coding of relaxing. Our rule was that the code for relaxing was only to be used when there was no other simultaneous activity in a timeslot. Where there was another activity, relaxing was to be coded as 000. The rule was not observed by all fieldworkers and the data had to be corrected. The code was changed to 000 and the simultaneity indicator of the other activity was set to 'No'.

iii.   Common types of edits
The most common edits checked (a) for skips and (b) for inconsistencies between responses to different questions. We wrote cleaning programs for each of the files (see below) which identified records with possible problems. Staff worked manually from these listings to check the captured data against the original questionnaires. They then corrected the SyBase data manually where appropriate.

A second important edit checked for blank activity fields. Here again staff worked manually, correcting what they could.

The original file contained asterisks for all fields which were either 'not applicable' or had missing data. A second round of automatic editing replaced the 'not applicable' fields with ampersand signs.

iv.   Common imputations
Imputation in respect of 'instantaneous' activities such as arriving, leaving and waking up were done manually. For 'arriving' and 'leaving' we used the relevant travel code of the previous or following timeslot, but inserted a 1 in the final digit so as to be able to distinguish the imputations during analysis. Where 'waking up' was the only activity recorded for a half-hour, we used the code 011 to show that it was related to sleeping, but not exactly this. Again, this will allow researchers to make their own decisions during analysis. All of these edits were done manually.

A third round of automatic editing imputed values for the few time slots in which there was no valid activity. The two rules adopted were:
(a) We allocated the same code as for the previous time-slot for the person where time-slots in the range 2-47 were empty.
(b) If the first time-slot for a person was empty, we allocated the sleeping code.
We distinguished these imputed values from original data by having a '7' as the third digit.

We have done few, if any, other imputations. As this is the first national time-use study in the country, we want to make the data available in a relatively unpolished form. We are hoping that by studying the inconsistencies and other problems, we and others will be able to suggest improvements in methodology for further studies.

d. Time-use data file
The data from the questionnaire was captured in four files:

- A control file (GEN) records the final result codes for the household as a whole and each of the selected individuals. The result codes indicate whether the interview was completed successfully and, if not, the reason for failure. There is one record for each household, but several result codes within each record.
- A household file (HHOLD) records most of the responses to the first section of the questionnaire. There is one record for each household.
- A members file (PERSONS) which contains person number (in order of age), sex, gender, and population group of all household members. There is one record for each member, giving multiple records for each household.
- A demographic file (PERSON) contains the responses to the second and fourth sections of the questionnaire and a few of the responses to the third and fifth sections. It thus contains. the demographic details of the two selected respondents in each household and a few overall questions relating to the diary day. There is one record for each selected person giving one or two records for each household.
- A diary file (DIARY) records how each of the selected respondents spent each half hour of the preceding day. There are 48 records for each selected person – one for each half hour of the preceding day. There are thus 48 or 96 records per household. Within each record, there are fields allowing for up to three activities.

We have created a fifth file from the original DIARY file. This file contains one record for each activity, giving a possible maximum (never reached!) of 144 activities for any one selected individual. Each record contains the person identification, activity code, location codes, simultaneity indicator, and time slot. The file also contains two different measures of minutes. The first measure (which we can call 'proportional time') divides the total time taken for an activity between the simultaneous activities occurring in that period. The second measure (which we can call 'full time') allocates the total time taken for an activity even if other activities occurred simultaneously. So, for example, if two simultaneous activities are recorded in a given half hour, the first measure will allocate 15 minutes to each and the second measure will allocate 30 minutes each.

7 Tabulation plan for time-use data

a. Development of the tabulation plan
In our first work on a tabulation plan we concentrated on finding issues that would be revealed by the time-use survey and would not have been available from other surveys. We looked, in particular, at how we would combine demographic and household data with the analysis of activities. The first draft of the plan was drawn up by the gender unit. A meeting was arranged for reference group members both inside and outside Stats SA, and the tabulation plan sent to them beforehand. The meeting was attended primarily by Stats SA staff, but we received comments and suggestions from several external reference group members.

The more standard tabulation plans were informed by discussions with our Norwegian technical advisers who suggested how we could achieve tabulations which are comparable with other countries. Our proposed treatment of simultaneous activities is informed by the work undertaken by the US Bureau of Labor Statistics in preparation for their time-use study.

The first part of the standard tabulations will use proportional time (see above) to describe how the 'average' South African, and the 'average' in various groups, spends the 24 hours of their day. We will present:

(a) the percentage of people in each sub-group recording each of the activities/groupings;

(b) the average time spent on each activity/grouping both in terms of absolute minutes and in terms of percentages of the 24-hour day for all people in the population or sub-grouping; and

(c) the average time spent on each activity/grouping in terms of absolute minutes and percentages of the 24-hour day for all people in the population or sub-grouping who engaged in that activity.

The tabulations will be by individual 3-digit category of activity, the ten 1-digit categories, and the three SNA groupings. Group statistics will be provided by population group, sex, age group, economic status, province and rural/urban.

The second part of the standard tabulations will use full time (see above) in recording activities which are commonly done simultaneous with other activities. These tables will not cover all activities and will not sum to 24 hours. They will, instead, record:

(a) how many people engaged in the activity;

(b) how long all members of the population or sub-group spent on that activity on average; and

(c) how long, on average, those who did the activity in each group spent on the activity in a day.

b. Basic summary statistics generated

The summary statistics will be generated according to the ten categories and three SNA groupings, dissagregated by population group, sex, age group, economic status, province, and rural/urban.

c. Distinctions between paid and unpaid work

The basic tabulations will be done both according to the ten categories of the trial classification, and the three SNA groupings. These do not give an exact match with paid and unpaid work in that SNA work can include some unpaid activities. In early testing we included a paid/unpaid context variable, but later dropped this whe n the trial classification seemed to have a close enough match for our purposes.

d. Estimation procedures

We used the sample frame which had been constructed for the Survey of activities of young people (SAYP) for the time use survey. For sampling purposes, urban areas were stratified into formal and informal settlements and non-urban areas into commercial farming and other (largely traditional) rural areas. The SAYP sample frame contained a disproportionately high number of urban informal and commercial farming areas for reasons related to the expected prevalence of child labour in different areas. The results of the TUS will, however, be weighted back to the relevant proportion of the population. Data from the two selected individuals in each of the households will be weighted so as to reflect the proportions in the population in terms of population group, sex and age group. Where sample size dictates, these stratification groups may need to be combined. The final results will be weighted so as to yield the total of sampled individuals rather than the total population of the

country aged 10 years and above. The proportions of the different sub-groups will, however, reflect the proportions in the total population of the country.

e. Titles of major tables
Not available as yet.

8 Dissemination and use of time-use data

a. Provision of results to expected users in the country and internationally
The data from the full survey is not yet available. The uncleaned and unweighted data from the first tranche were used in preparing a paper for an ILO expert group meeting on place of work. Gender unit staff have also already included information about the time-use survey in workshops and presentations. The fact that Stats SA is conducting a time use survey is relatively well known as it has been reported in several government and non-government reports.

b. Ways in which survey results will be made available
In terms of publications, we are planning at least two by mid-2001. The first will consist primarily of standard tables, with a short introduction and summary. The second will be more discursive and include graphs and discussion on a range of topics. Further publications will focus on specific issues. We are planning a formal launch of the first publications to which stakeholders from both inside and outside government will be invited. We will also offer briefings to groups such as the Parliamentary Committee on the Quality of Life & Status of Women, the Commission on Gender Equality and others.

The data will be made available to the public in the form of flat ASCII files with accompanying metadata. Prices will differ for private companies, non-profit organisations, and education-related users.

c. *Use of time-use data to improve measurement of paid and unpaid work*
During 2001 the gender unit will work together with the national accounts division of Stats SA and Norwegian advisors in an attempt to construct parallel accounts.