

Survey structure and sample design: consultant's recommendations

Chris Scott, World Bank consultant. October 1989

1. BACKGROUND

The Ghana Living Standards Survey (GLSS) is a multisubject household survey carried out by the Ghana Statistical Service (GSS) in the context of the World Bank's Social Dimensions of Adjustment (SDA) programme. The survey began in September 1987 and the 2nd year's operations were completed in September 1989. The survey design closely followed the SDA intensive survey design recommended at that period, which was itself largely based on the World Bank's Living Standards Measurement Survey project.

It was agreed between the GSS and the World Bank late in 1988 that the 3rd year of the project should depart substantially from the first 2 years by giving much fuller attention to household expenditure and consumption with some reduction in the other objectives and employing a larger, and more widespread, sample. In certain respects this was intended as a return to an earlier survey plan developed with the assistance of ECA and UN advisers under the African Household Survey Programme. A joint ECA - World Bank mission visited Ghana in July-August 1989 at which time a project proposal was drawn up for the forthcoming Consumption and Expenditure Survey. (It may be noted, however, that if this survey were to be fitted into the GLSS programme without a break in the timing of the programme, such a proposal would have to have been completed very early in 1989.)

The present consultant was asked to work with the GSS to prepare a detailed technical plan in line with the funding proposal in October 1989. This report is the result. While the consultant takes responsibility for the report's recommendations it should be stressed that they have been prepared in close collaboration with the relevant workers at the GLSS; much of the key work in this report is theirs as well as several of the proposed solutions to methodological problems,

Section 2 of this report describes the proposed survey structure, that is, the way in which the interviewers' visits are distributed in space and in time among the sample households. Section 3 gives the proposed sample design and Section 4 the operational timetable,

2. SURVEY STRUCTURE

The basic constraints are those enumerated in the project document. However, some of these are more flexible than others. We have assumed that the over-riding constraint is the use of no more than 10 vehicles for use by the survey field teams. Within this framework the document proposes the use of ten 7-person teams constituted thus:

- 1 supervisor
- 3 interviewers
- 1 anthropometrist
- 1 driver
- 1 data entry operator

It is also assumed in the document that 10 interviews will be carried out with each household, spread over 1 month. The resulting total sample size is 4320 households. The present report suggests some modifications to these parameters but without exceeding the total budget or the limit of 10 cars.

The following objectives are cited in the project document for the survey:

- 1) To provide information on patterns of household consumption and expenditure at a greater level of disaggregation than is currently possible.
- 2) To serve as a basis for the construction of the consumer price index.
- 3) For updating the national accounts.
- 4) In combination with the GLSS, to give a data base for national and regional planning and for the estimation of consumption as a proportion of household production.

Objective (2) hardly seems relevant here: the first 2 years of the GLSS are fully adequate for this purpose. (See Appendix 1.)

Translating these requirements into operational terms, it appears that the over-riding objective is to obtain high quality data on household consumption and expenditure with a larger and more widespread sample than that of the GLSS. Consumption here refers to the monetary value of home-produced food consumed during the reference period. As far as the GLSS design (1st 2 years) is concerned the quality of these data suffers from a number of limitations. The present design attempts to overcome these by various methods now to be discussed.

2.1 The recall problem

The GLSS recall period for consumption and expenditure on most items is 2 weeks. A special experimental study carried out in Ghana* appears to indicate a memory loss averaging over 20% for such data, at least for frequently purchased items. Even a 3-day recall period appears to yield nearly 15% under-reporting, while a 2-day period yields about 5% under-reporting. (These figures are relative to the 1-day recall period as a standard; of course it is possible that the standard is itself subject to mis-reporting.) These findings present a problem for the survey designer. To arrange interviews every day with each household appears excessively expensive (although this has been done in many - indeed most - African surveys). However there is some hope that the problem can be overcome through the use of the account-book, or diary, method: here the household itself records its own expenditures every day and the interviewer passes by from time to time to check that the diary is being used and to transfer the data into a questionnaire. Obviously this requires the presence of a literate person in the household. The method has been widely used in African urban surveys, often using the services of a school child in the household.

We therefore propose as a compromise solution:

Rural sector. Interviewer's visits at 2-day intervals

Urban sector. Diary method. Interviewer's visits at 3-day intervals.

2.2 Reference period problem

The reference period for each household is the total period to which the consumption/expenditure reporting relates. If there is only one interview, the reference period is equal to the recall period, otherwise the reference period is generally the sum of the recall periods for each household.

* Scott, C. & Emenuegbo, B. (1989). "Effect of recall duration on reporting of household expenditures: an experimental study in Ghana". Forthcoming. World Bank, SDA Special Studies series.

In the urban sector a high proportion of households receive a monthly wage or salary. Pay-day may have a strong effect on expenditures and some effect on consumption. If the reference period is less than 1 month, it will be a matter of chance whether pay-day falls in the reference period. If it does, the household will appear richer than it really is; if it does not, it will appear poorer than it is. Either way such households will tend to be misclassified as to their level of living. Thus, in the urban sector we need a 1-month reference period.

In the rural sector monthly wages are rare. Here the periodicity is more often weekly, due to the weekly market.

In the light of these requirements we recommend the following:

Rural sector. 8 interviewer visits at 2-day intervals.

Urban sector. 11 interviewer visits at 3-day intervals supported by diary of expenditures and consumption.

2.3 Obtaining complete time coverage

For high quality data, the survey needs to cover practically the whole year. This presents a problem in relation to field workers' leave. If the team works every day of the month, how do they get their week-ends off? If they work every month of the year, how do they get their annual leave? These are two separate problems. We suggest the following solutions:

Within the month. The team has 2 regular interviewers and 1 supplementary interviewer. The regular schedule of visits by each of the regular interviewers is taken over by the supplementary interviewer for 2 days in the 16 days (rural) or for 6 days in the 33 days (urban). On these days the regular interviewer takes a rest-day, equivalent to a week-end.

Within the year. The year is divided into 11 cycles of 33 days. 11 teams are recruited. Each team works for 10 cycles and takes 1 cycle off as leave. Each team takes a different cycle as leave so that there are 10 teams at work at all times.

The use of a cyclic period different from 1 month has the advantage of staggering the position of pay-day, or market day, from cycle to cycle, thus avoiding interactions with possible bias cycles, such as start-up bias, recall bias, fatigue, etc.

2.4 Interviewer workloads

The repeat visits to each household during the month make it possible to spread the household's total interviewing workload over a number of interviews, thus reducing the length of each interview. Obviously this does not apply to the main consumption/expenditure questionnaires for frequent items: these must have short recall periods and long reference periods, and so will need to be repeated at every visit. The remaining questionnaires can be spread over many interviews, each being used only once in each household. The questionnaire on less frequent expenditures, in particular, will be used at the last visit, with a recall period equal to the reference period - 15 days in the rural sector and 30 in the urban.

We assume that each interviewer can do 5 interviews per day, whether urban or rural. Thus a workload for the 16 days in a rural cluster is 10 households, while a workload for the 33 days in an urban cluster is 15 households. The rural workload divides into 2 batches of 5 households. Batch 1 is interviewed on days 1, 3, 5, 7, 9, 11, 13, 15 while batch 2 is interviewed on days 2, 4, 6, 8, 10, 12, 14, 16. (These day numbers are not actual calendar dates, since the cyclic period of 33 days will lead to workloads in different clusters starting on

different dates.) The supplementary interviewer will take over the work on days 7, 8, 15 and 16 (or on different days if more convenient, at the discretion of the supervisor). By having the supplementary interviewer take over days 15 and 16, in particular, we give the regular interviewer time to move to another cluster for the next workload. Turning to the urban workload of 15 households, this divides into 3 batches of 5 households, which are interviewed thus:

Batch 1. Days 1, 4, 7, ... 25, 28, 31

Batch 2. Days 2, 5, 8, ... 26, 29, 32

Batch 3. Days 3, 6, 9, ... 27, 30, 33

Here the supplementary interviewer might take over days 7, 8, 9 and 31, 32, 33. Again, the arrangements can be flexible at the discretion of the supervisor but the inclusion of the last 3 days is designed to give the regular interviewer time to move to another cluster.

2.5 Spread of the interviewers among clusters, and total sample size

A preliminary look at the sample size resulting from the above leads to the following:

Rural. 10 households x 22 cycles x 2 interviewers per rural team

Urban. 15 households x 11 cycles x 2 interviewers per urban team.

Allowing 7 rural and 3 urban teams at work in each cycle this leads to a total sample of:

$$\begin{array}{r} 10 \times 22 \times 2 \times 7 = 3080 \text{ rural households} \\ 15 \times 11 \times 2 \times 3 = \quad 990 \text{ urban households} \\ \hline 4070 \end{array}$$

If the 2 interviewers of one team were to work in the same cluster this would lead to 154 rural and 33 urban clusters, with a "take" of 20 households in each rural cluster and 30 in each urban cluster. These figures are unsatisfactory: the takes within clusters are inefficiently high and the total number of urban clusters is much too small for analytic purposes. As a first modification we suggest having each interviewer work in a separate cluster.* This will halve the above cluster-takes and double the number of clusters. However the number of urban clusters, now 66, is still uncomfortably small and we therefore suggest one more modification: in the urban sector 3 regular interviewers will be used in each team instead of 2. This leads to the following sample size:

Rural. 3080 households in 308 clusters

Urban. 1485 households in 99 clusters

4565 407

The organizational implications of these decisions should not be overlooked. In the rural sector one supervisor will have to look after 2 regular interviewers who will (usually) be sited in two different locations. Moreover the supplementary interviewer will have to work in both locations. Since the number of rural clusters is more than twice as high as in the GLSS, the clusters will be rather closer together than in past years. We do not believe that an additional sampling stage is needed to bring them still closer. However, greater care may be needed in working out the teams' itineraries to ensure that two workloads allocated to the same team are always as close together as possible within the

* As will be explained in Section 3, owing to the special sampling method used there will be occasional cases of two workloads in one cluster. Where this happens, two interviewers will still be working together in the same cluster. For simplicity we ignore this exceptional case in the present discussion.

selected sample. In particular, where two workloads fall in the same cluster they should always be paired together and allocated to the same team on the same dates. In the urban sector there will be 3 interviewers per team. Itineraries should be arranged so that as far as possible the team members operate in the same town at any given moment, or at most in 2 nearby towns.

To allow for the additional team member in each urban team within the planned budget, and in line with the agreed principle of simplifying that part of the survey not concerned with consumption/expenditure, we suggest that the anthropometry module be dropped altogether for this survey, thus reducing all teams by 1 person.

2.6 Longitudinal sample

In an expenditure survey in which each household is covered for one month and the sample as a whole is spread evenly over the year, we can estimate mean annual expenditure per household from the sum of the mean monthly expenditures. Unfortunately this simple procedure does not work if the objective is to estimate the spread of annual expenditure among households. The standard deviation of households' annual expenditure is not 12 times the standard deviation of their monthly expenditure, even if the latter is the same in all months. The correct multiplier will be less than 12 and depends on the correlation between different months for the same households.* The required estimate depends on knowing the average value of this correlation and this in turn requires a longitudinal element in the data collection.

It is recommended in the project document that information on the level of correlation be obtained by training locally recruited persons in a strictly limited task: that of recording daily expenditures over a cycle (16 days rural or 33 days urban) for 3 sample households. The work will start immediately after the departure of the regular interviewing team and continue until the end of the year. The sample will rotate each cycle, using always 3 households from the initial sample in the cluster considered. In the project document it is assumed that this procedure will be applied in every sample cluster. However, with the increase in the number of planned clusters this now appears an excessively large sample which it would be difficult to control. Thus it is now proposed to limit this study to the sample relating to one interviewer only in each team.

An example may make this clearer. Suppose one of the teams is working on the 14th rural (16-day) cycle in an area. One of the regular interviewers of the team has been designated to train a local person by bringing him along to observe his expenditure interviews. When the team leaves the area at the end of the cycle, the local interviewer begins work, using a specially prepared booklet in which all the relevant dates and respondents have been entered in advance. For the following cycle (the 15th cycle) he makes daily visits to households 1, 2 and 3 of the sample of 10 used by the regular interviewer who trained him. In the 16th cycle he switches to households 4, 5 and 6, and in the 17th cycle to households 7, 8 and 9. In the 18th cycle he returns to numbers 1, 2 and 3, and so on until the end of the last cycle (cycle 22). Note that household 10 is not used. In the urban sector all of the 15 households can be used because 15 is divisible by 3.

The only purpose of this exercise is to obtain some fairly rough estimates of the correlation between months for the same household. A substantial drop-out rate must be anticipated but should not seriously bias the estimates.

2.7 Data editing and feedback

In the GLSS, interviewers made two visits to each household at an interval of 2 weeks. In the intervening period the data from the first interview were

* For details see Scott, C. (1989). "Estimating the distribution of households' annual expenditure from cross-sectional monthly data." World Bank, SDA Special Studies.

transferred to the data entry operator who entered them using an edit program. Errors found were listed and the interviewer had time to pick up the list before returning for the second visit to the household, where he carried out any necessary checks resulting from the computer edit. This facility for field correction was thus available only for that part of the questionnaire used in the first visit, which incidentally did not include the consumption/expenditure modules.

In the consumption/expenditure survey it would be possible to check in the same way on the results of the first few sections of the questionnaire in time for the interviewer to have the list of errors in hand before his last visit to the household. One disadvantage is that the household roster (Section 1) would have to be copied out, since it is needed both for the data entry program and by the interviewer working on the last few sections.

Any arrangements of this kind should be worked out later when the contents of the different sections of the questionnaire have been settled. We therefore abstain from any recommendations on the matter in this report.

3. SAMPLE DESIGN

The guiding principle of the sample design proposed below is to continue with the methodology of the first 2 years of the GLSS with minimal modification. This methodology is already familiar to the staff working on GLSS and has been described in some detail in an earlier paper.* It will be briefly summarized here.

A master sample of 300 EAs was selected using stratified systematic sampling with probability proportional to the number of census households. Eight replicates of 100 EAs each were constituted by systematic allocation on the basis of the order used in the initial EA sample selection.

Replicates 1 and 5 were used in the 1st year and replicates 1 and 6 in the 2nd year. Thus there remain 5 replicates (2, 3, 4, 7 and 8) which have not yet been interviewed, making 500 EAs. Replicate 2 has already been household-listed (in July-August 1989) in anticipation of the 3rd year.

The design used in the first 2 years was devised to provide a self-weighting household sample together with fixed interviewer workloads. The EAs in the newly selected replicates were household-listed shortly before the survey. The numbers of households listed were used, in conjunction with the numbers found in the census, to allocate 200 workloads of exactly 16 households among the 200 selected EAs in such a way as to yield a self-weighting household sample. This led to some EAs getting no workloads and some getting 2, or in one case 3.

In attempting to apply this method to the present survey two problems arise.

First, the required number of EAs is 407 (see Section 2.5). If we decide to avoid EAs used before, the number of EAs available is 500. So, clearly, we have to draw on all of the 5 available replicates. This situation suggests that we may best ignore altogether the existing division into replicates and simply select the EA sample directly with equal probability from among the available 500. On the other hand it seems desirable to take advantage of the listing work already done (Replicate 2, in July-August 1989). Thus we propose selecting the whole of Replicate 2 and a further 307 from among the 400 EAs of replicates

* Scott, C. & Amenuvogbe, B. (1986). "Ghana Living Standards Survey: Recommended sample design." Working Paper, Ghana Statistical Service, Accra.

Second, the use of different sampling parameters in the urban and rural sectors, described in Sections 2.2, 2.3 and 2.4, combined with the need to avoid splitting any team between the rural and urban procedures, leads to a slight anomaly which requires adjustment. We have already seen in Section 2.5 that the decision to use 7 teams in the rural and 3 in the urban leads to a sample breakdown of 3080 rural and 1495 urban households - respectively 67.5% and 32.5%. By good fortune these figures agree quite closely with the percentage rural/urban in the whole population. Our best estimate of the latter, in terms of household numbers, is 65% and 35%. If these two were the same we could use a single overall sampling fraction throughout, implying self-weighting, and be sure of attaining the above distribution. The agreement is so close that we propose to make an ad hoc adjustment based on the following reasoning.

The first set of percentages (67.5 v. 32.5) relates to the mode of field organization which is different for the "rural" and "urban" sectors. The second set (65 v. 35) relates to the census classification. However, there is no reason why the criterion for organizing field work needs to be the same as the criterion adopted for census reporting. We propose therefore to reclassify 2 1/2% of the master sample EAs from "urban" to "rural" for the purpose of field organization, i.e. to make visits at 2-day intervals over a 16 day period in these areas instead of 3-day intervals over a 33-day period. Obviously we are free to reorganize the field work in such a manner without any implications for sampling or substantive analysis.

For the purpose of this adjustment we have reclassified 13 EAs from urban to rural before subselection for the consumption expenditure survey. The EAs reclassified are those from the smallest towns in the master sample.

After this adjustment we have the following parameters:

	No. of EAs in master sample replicates		No. of EAs to be selected
	Replicate 2	Replicates 3,4,7,8	
"Rural"	67	269	308
"Urban"	33	131	99

The next step was to select the EA subsample from the master sample, using the numbers in the right hand column and taking all of Replicate 2 and the required remainder from the other replicates pooled. Selection was made by systematic sampling within the master sample with equal probability in each sector.

The remaining steps of sampling are exactly as in previous years (see Scott & Anonvogbe 1986, op. cit.) except that the parameters are different. After household listing in these EAs, 407 workloads will be allocated among the 407 EAs, using the same method as in previous years. This will lead to a number of cases (perhaps about 50) where there is no workload in the EA and a similar number where there are 2 workloads or possibly 3.

Note that, in selecting households for each workload, an allowance should be made for a short list of replacements. Each rural workload will be selected with 15 households and each urban workload with 20, these figures including 5 spares in each case. The spares should be identified in the list of 15 (or 20) by systematic subsampling at intervals 3 (or 4) from a random start. The spares should be typed up on a separate sheet, which should be made available to the supervisor but not the interviewers in each team.

* The listing in Replicate 2 will be older by 8 or 9 months than in the other 4 replicates. We preferred to save listing costs and time by making full use of Replicate 2 nevertheless, despite the small additional bias. See Appendix 2 for data on list turnover.

4. TIMETABLE OF OPERATIONS

In drawing up the survey timetable one feature still remains to be determined at the time of writing: the SDA Unit at the World Bank has requested the Ghana Statistical Service to undertake, under contract to the World Bank, a pilot test of the new SDA questionnaire proposed for the "Priority Survey". We refer to this below as the "PS Pretest". This work would involve 12 interviewers working in 6 teams with 5 cars (2 teams in Accra would share one car), for 4 weeks including 1 week of training.

The simplest way of fitting this pretest into the programme of activities for the preparatory phases of the Ghana Household Consumption and Expenditure Survey would be to carry out the PS Pretest in January 1990 before any significant field activities have begun. However this would lead to a 4-week delay in the start of the Consumption/Expenditure survey, which would go into the field on 28 May 1990. A second option, which seems just manageable, would be to squeeze the PS Pretest into the period when the final questionnaires are being printed for the Consumption/Expenditure survey. In this option the listing operation for the main survey overlaps in the field with the PS Pretest. This would represent a tight squeeze from three different angles: time, personnel and cars, but is theoretically possible if nothing goes wrong. In this option field work starts on 30 April 1990.

The timetables for both options are presented overleaf. They have been worked out only up to the start of the main field work.

The choice between these two options appears to be a matter for negotiation between the Ghana Statistical Service and the World Bank.

TIMETABLE OF OPERATIONS UP TO START OF FIELD WORK

Finalise survey design and survey design	End October 1989
Complete revision of questionnaires	End November
Printing of pilot questionnaires	1 - 31 December
Revise manuals	1 - 15 December

	OPTION 1	OPTION 2
P.S. Pretest	8 Jan - 2 Feb 1990	5 - 31 Mar 1990
Training for pilot (Cons/Exp Survey)	5 - 16 Feb	8 - 20 Jan
Field work for pilot	19 Feb - 16 Mar	22 Jan - 16 Feb
Finalise questionnaires and manuals	19 - 30 Mar	19 Feb - 2 Mar
Questionnaires and manuals to printer	30 Mar	2 Mar
Questionnaires and manuals received from printer	4 May	6 April
Training of field workers	14 - 25 May	16 - 27 Apr
Send-off for field teams	26 May	28 Apr.
Start of field work	28 May	30 Apr
Sampling and related activities:		
Selection of EAs	31 Oct 1989	31 October 1989
Listing of households	27 Mar - 11 May	19 Feb - 13 Apr
Selection of workloads	14 - 18 May	16 Apr - 20 Apr
Selection of households	21 May - 15 June	23 Apr - 19 May

APPENDIX 1. Note on weights for a consumer price index

The provision of weights for a consumer price index is often mentioned as an objective of a household budget survey, and has indeed been cited among the objectives of the present survey. The regular publication of price index numbers in the GSS Statistical News Letter provides a convenient opportunity for a simple analysis of the effect on the index of errors in the weighting. The data for the following analysis come from the News Letter of 14 August 1989.

Separate weights are cited for 9 categories of expenditure. They are as follows:

Food	49.2
Beverages and tobacco	6.2
Clothing and footwear	19.2
Gross rent, fuel and power	6.8
Furniture, furnishings, h'hold equipment & operation	5.1
Medical care and health	1.8
Transport and communications	4.3
Recreation, education & cultural services	5.5
Miscellaneous goods and services	1.9
	<hr/> 100.0

To examine the effect of errors in the weights let us adopt a set of weights which are as wrong as possible and see what happens to movements of the index.

The purpose of the weights is to differentiate the effects of different categories. Let us perversely make all 9 weights equal. One might reasonably say that nothing could be more erroneous than this.

Using such equal weights and comparing the combined index for January 1989 with that for January 1988, we obtain an increase over the year of:

- 26.03% with the correct weighting
- 27.98% with the "worst possible" (equal) weighting.

Most of this difference is due to the transport category, whose movement is much more irregular than the others. If we exclude this category the difference practically vanishes. We obtain increases over the year of:

- 25.33% with the correct weighting
- 25.53% with the worst possible weighting.

Could these results be due to special peculiarities of the year 1988/89? One way of answering this would be to look at the movement over the complete period covered by the index, namely the 12 years from 1977 to June 1989. With the correct weighting the average annual increase works out at 48%: with the worst possible (equal) weighting it is 50%.

These results strikingly illustrate the conclusion that movements of the consumer price index are very little influenced by the precise weights adopted, especially in times of high inflation. In particular there is no justification for mounting an expensive large-scale survey if the sole purpose is to estimate the consumer price index weights. A very rough survey on a hundred or so households will be fully adequate for this purpose.