Review of Final Sample for the Trabalho de Inquérito Agrícola (TIA), and Recommended Weighting Procedures

David J. Megill
Sampling Consultant
July 2002

1. Review of Final Sample for TIA

The final sample design for TIA is described in the report on "*Análise de Erros Padrão Simulados para os Resultados do Trabalho de Inquérito Agrícola 2002 Baseados numa Amostra de 80 Distritos para Moçambique*" (Megill, April 2002). A total of 80 districts are included in the sample, stratified by agro-ecological zones within province. The final sample consisted of 560 segments, with 8 small farms selected within each sample segment. In 18 strata the sample was selected in three stages, with districts selected at the first stage, segments at the second stage and farm households at the last stage; within each sample district seven segments were selected. In the case of the remaining 8 strata a two-stage sample design was used, with segments selected at the first stage; the sample segments in these strata were allocated to each district proportionally to the its number of households.

The 406 sample segments selected in the strata with a three-stage sample design for TIA were drawn from a subsample previously selected by Christopher Hill from all the primary sampling units (UPAs) in the sample for the Censo Agro-Pecuário (CAP); 20 of these sample segments did not have CAP data. However, the 154 sample segments selected in the 8 strata with a two-stage sample design were selected from the frame of UPAs with CAP data, since this sample was used for the CENVAR analysis of simulated standard errors. The reference report described a simple procedure which could be used to adjust this sample to also represent the part of the frame corresponding to UPAs without CAP data.

In preparing the final list of segments for the TIA listing operation, it appears that the final sample corresponds to the same sample segments used for the CENVAR analysis; that is, the sampling frame was not adjusted by adding the UPAs without CAP data to the frame. In this case a small bias may be introduced, so it is important to determine the number of UPAs which were excluded from the frame. The UPAs which are missing from the frame can be identified from a skip in the UPA serial numbers within the corresponding strata.

Fortunately, only two missing UPAs were identified out of a total of 592 UPAs in the frame for the 8 strata with a two-stage sample design (that is, about 0.034 percent), so the corresponding bias should be negligible. These two UPAs without CAP data are identified in Table 1.

Table 1. List of UPAs without CAP Data in Frame for 8 Strata with Two-Stage Sample, Missing from Original Frame

Province	Stratum	District Code	UPA Code from CAP
Nampula	0311	02	022
Tete	0502	07	117

In order to determine how the TIA sample was affected by omitting these two UPAs without CAP data from the frame, the procedures specified in the reference report for adding them to the frame and adjusting the sample were implemented in the spreadsheet TIA4SMPA.XLS. Once these two UPAs were added to the frame in their respective strata, only three sample UPAs were different in the new sample. The original and new sample UPA for each of these cases is specified in Table 2.

Table 2. Sample UPAs Changed in the Adjusted Sample Selection for TIA

Province	Stratum	Original Sample		Adjusted Sample	
		District Code	UPA Code from CAP	District Code	UPA Code from CAP
Nampula	0311	02	021	02	022
Tete	0502	07	115	07	116
Tete	0502	07	123	07	122

Since these changes would only have a very minor effect on the TIA results, the original sample UPAs can be used if the listing of households has already been conducted in these segments; they can be considered to be representative. However, if the listing has not been conducted yet, it would be possible to make this small correction to the sample.