STATISTICS DENMARK

Statistics Norway
Statistisk sentralbyrå

Statistiska centralbyrån
Statistics Sweden

MZ:2010:07

IOF 2008/09

Summary report from a series of missions to the National Statistical Institute of Mozambique, Maputo Mozambique

01 June 2009 -28 May 2010

within the frame work of the

AGREEMENT ON CONSULTING ON INSTITUTIONAL CAPACITY BUILDING, ECONOMIC STATISTICS AND RELATED AREAS

between

INE and Scanstat

Triebkorn, Megill, Mathiassen, Kiøsterud, Otto

Instituto Nacional de Estatística

Ref: Contract DARH/2008 /004

Lars Carlsson
lars.carlsson@scb.se

Table of contents

This summary report includes the following mission reports:

Executive summary in English

In the SEN Strategic Plan 2008 - 2012 one of the priority activities is the execution of the Household Budget Survey 2008/09, which has as some of its main objectives to gather data making it possible to evaluate the Action Plan of Reduction of Absolute Poverty (PARPA II) and monitor the fulfillment of the Millennium Goals.

The National Institute of Statistics has a lot of experience in conducting household surveys, but there are still weaknesses in database cleaning and in analysis of the information, taking into account the multi sectoral data that this survey include. For this reason a combined advisory mission to support INE is necessary. The following expertise is needed:

1. **A specialist in household surveys**, especially in the area of household budget and expenditure;

2. **A specialist in sampling for household surveys**, with experience in household budget surveys;

3. **A specialist in poverty analysis**, using income and expenditure information from the households.

The experts were offered from of the Scanstat Consortium and their respective mission reports can be read below in reverse chronological order.

Resumo executivo em Português

No Plano Estratégico do SEN 2008 - 2012 uma das actividades prioritárias é a execução Inquérito sobre Orçamentos Familiar 2008/09, que tem como alguns dos seus principais objectivos a recolha de dados que permitam avaliar o Plano de Acção de Redução da Pobreza Absoluta (PARPA II), e acompanhar o cumprimento das Metas do Milénio.

O Instituto Nacional de Estatística tem muita experiência na condução de inquéritos aos agregados familiares, mas ainda há deficiências na limpeza do banco de dados e na análise das informações, tendo em conta os dados multi setoriais que este inquerito inclui. Por esta razão, uma missão combinada de assessoramento para apoiar INE era necessário. As seguintes especialistas eram necessárias:

1. **Um especialista em inquéritos aos agregados familiares**, especialmente na área de orçamento doméstico e despesas;

2. **Um especialista em amostragem por inquéritos aos agregados familiares**, com experiência em inquéritos aos agregados familiares;

3. **Um especialista em análise de pobreza**, usando rendimentos e despesas a partir de informações dos agregados familiares.

As especialistas foram oferecidas do Consórcio Scanstat e os seus respectivos relatórios de missão pode ser lido abaixo, em ordem cronológica inversa.

STATISTICS DENMARK

Statistics Norway
Statistisk sentralbyrå

Statistiska centralbyrån
Statistics Sweden

Mission Report

for a short-term mission as survey data processing specialist

From 22 April to 28 May 2010

James Otto

Instituto Nacional de Estatística

Ref: Contract DARH/2008 /004

James Otto

USA: *212 Third Street, Apt. 1d*
Troy, New York 12180-4650
E-Mail: *bwJim00@Yahoo.com*
Telephone: *518 271-8060*

Mozambique: *Hotel Terminus, Maputo*

Table of contents

I.1 EXECUTIVE SUMMARY

The National Institute of Statistics (INE) conducted its third Household Budget Survey (IOF2008/9) from September 2008 to September 2009. The sample of 10848 households in 1060 enumeration areas (EA) represented the nation, rural and urban zones as well as the provincial level. In each urban EA 12 households were selected to be interviewed, and 9 households were selected in each of the rural EAs.

Data processing was a major problem for this survey because of the data processing activities of the 2007 Population Census and the lack of professional data processing staff in the Survey and Censuses Department of INE. The initial data entry system had many problems and was replaced by a new system in June 2009. The new system included double data entry with verification. Errors detected in verification were corrected by reference to the questionnaires. However, there was no computer system for data validation following data entry because of the previously mentioned lack of personnel. Data entry for the entire survey was completed in October, 2009. Manual validation of the entered data continued until the World Bank provided a data analyst in November, 2009 who was able to assist with the automated validation, correction and analysis of the survey data.

The purpose of the data processing mission was to adapt the data processing system that had been developed for the 2002-3 IAF survey to process the 2008-9 IOF survey. Because of the limited time available for the mission (4 weeks, extended to 5 weeks) the work concentrated on aggregating the income and expenditure data in order to prepare nominal estimates of average household expenditure and produce tabulations that could be used for comparison with the 2002-3 survey data.

The 2002-3 system was adapted incrementally for the 2008-9 data. The data were converted to a Microsoft Access database with the same general structure used for the previous survey. The initial estimate of average daily household expenditure was prepared after one week. The system was extended over the next three weeks. The updated results, including comparison with the 2002-3 results, were presented to INE, MPD and donor organizations for discussion during the last week of the mission.

I.2 Terms of reference:

The Consultant will be responsible for the following activities:

1. Check the specifications for the integrated validation of all Income and Expenditure questionnaires including the household characteristics according the methodology used on IAF 2002/3;

2. Check the tabulation plan for all IAF2002/3 questionnaires;

3. Check the specifications for the aggregation of the income and expenditure data. This will include:

   Methods for the imputation of values for all transactions without values;
   Calculation of spatial and temporal weights to allow annual
   National estimates of household income and expenditure;
   Definition of household income and expenditure categories as multiple levels of detail;

4. Develop programs to implement the specifications defined in 1-3 according the methodology used on IAF 2002/3;

5. Develop the program of data archive for the IOF 2008/9.

I.3 Activities during the mission:

1. Created a set of integrated Microsoft Access databases for all questionnaire and derived data from the IOF2008-9 survey. See database summary document for details.

2. Modified the programs used to process the data from the IAF2002-3 survey to process the data from the 2008-9 survey. This included programs:

   to aggregate the expenditure and income data

   to assemble a summary of price data from the prices recorded in the community questionnaire and the observed transactions in the daily expenditure questionnaire, classified by product, province and urban/rural

   to validate prices recorded in daily purchases and estimated prices of own consumptions in the daily expenditure questionnaire

   to create the variables used to disaggregate results in the income and expenditure tables

   to produce the principal expenditure and income tables and to estimate sampling errors for the expenditure and income aggregates

   to compare the results of the 2008-9 survey with the results of the 2002-3 survey

3. Developed program modules to process the household data. The modules were tested in a program that prints a list of all the data in the household questionnaire.

4. Prepared preliminary documentation of the survey databases and the programs used to process the survey data.

5. Installed the survey data processing system on computers used by the INE data processing staff.

6. Trained the INE data processing staff to use the data processing system.

The following tasks were not done because of lack of time:

1. Development of the validation program for the household questionnaire.

2. Development of tabulation programs for the household questionnaire.

3. Creation of a complete data archive for the survey.

Notes:

1. Validation of the income and expenditure data was done by Carlo Azzarri prior to my arrival.

2. Imputation of the value of own consumption was also done by Carlo Azzarri of the World Bank prior to my arrival.

3. The calculation of household weights was done by INE's sampling specialist Carlos Creva.

4. There were problems with the community questionnaires. The initial data file had information for only 497 of the 624 rural enumeration areas. Data for an additional 112 enumeration areas were either found in files that had been misplaced or were re-entered from the questionnaire. Questionnaires for 25 enumeration areas were not found.

   Many duplicate price observations were detected when the recovered/re-entered data were processed.
   The questionnaire does not provide for multiple observations of the same product in a market or for multiple markets in an enumeration area. To permit further processing, all exact duplicates were eliminated and only the first observation was retained for any remaining duplicates. The price information in the community questionnaires is very useful and these problems need to be resolved.

   The price data in the survey community questionnaire database should be compared to the data in the questionnaires. Any errors detected will need to be corrected. If duplicates are to be allowed there needs to be a way to uniquely identify every observation. The World Bank plans to use the community questionnaire price data as part of the poverty assessment and they should be involved with the error detection and correction.

   The validation of prices in the daily expenditure questionnaire is a continuing activity. Data for one province have been corrected. This work will continue until all provinces have been corrected. The income and expenditure summary tables will be prepared periodically to monitor the effect of the corrections (expected to be small).

5. The aggregation of values for durable goods and for imputed rents poses methodological problems that have been resolved by using regression models. This was done as part of the poverty assessments in 1996 and 2002. In 2002 after futile attempts to incorporate the regression models in the survey data processing system, we decided to use the data from the MPD regression models. MPD provided the estimates for the use value of durable goods and the imputed value of rent for each household. This data was then used in system aggregation procedure. There is a general description of the need for regression models to estimate this data in the document "Components of household expenditure aggregates".
   A summary of the rent information recorded in the questionnaire was prepared for use by the INE National Accounts department.

I.4 RECOMMENDATIONS

The IOF 2008 demonstrated that INE's capacity just to process surveys is limited by lack of qualified personal. There is only one programmer in the Census and Survey Department and this limits activities to one survey and/or census at a time. This mission has successfully produced provisional results for IOF 2008, however there is much more work to do. Ideally this work should be used to develop INE's data processing capacity. INE data processing staff should do the remaining work with supervision and training provided by technical assistance. The assistance can be local, remote or a combination of the two. What is most important is that INE do the work.

The tasks to be done are:

Household questionnaire validation and correction
Tourism questionnaire validation and correction
Community questionnaire validation and correction

Tabulation and analysis of the household, tourism and community data

Produce final income and expenditure results adjusted with price deflators
Produce additional macro and/or micro level results for INE departments

Creation of the survey archive

APPENDIX I.1. Persons met

INE
Dr. João Loureiro, INE President
Manuel da Costa Gaspar, Vice-President
Arão Balate, Director of surveys and censuses
Fátima Zacarias, Director of demographic and vital statistics
Cristóvão Muahio, Head of Department
Carlos Creva, Sampling specialist
Eugenio Matavel, Data processing supervisor
Ramiro Mozinho, Data cleaning supervisor
Firmino Guiliche, National accounts
Monica Magaua, National accounts
Ernesto Samo, National accounts

Lars Carlsson, Scanstat resident advisor

MPD
Channing Arndt, Poverty assessment consultant
Sam Jones, Poverty assessment consultant

World Bank
Antonio Nucifora, World Bank Resident Mission
Carlo Azzarri, Data analysis consultant

Consultants
Eduardo Arriaga, Census data analysis consultant

APPENDIX I.2. Related files

1. List of files installed on INE computers
2. IOF2008 database summary
3. 2008 household aggregation procedure
4. Selected queries used in 2008 household aggregation procedure
5. IAF2002 program specifications
6. IOF 2008 results
7. Tables comparing IOF 2008 and IAF 2002 results
8. Comparison of 2008 nominal consumption estimates from INE and MPD
9. Household questionnaire validation specifications prepared by Erwin Triebkorn
10. IOF 2008 tabulation plan prepared by INE
11. Components of household expenditure aggregates
12. System: VBA program modules for IOF2008
13. IOF2008classes – VBA class modules for IOF2008
14. QAFclasses – VBA class modules for 2008 hh questionnaire

IOF 2008 results:

IECore: Moçambique Estimativas Básicas de Despesas e Receitas (2008-9)
IECoreProv: Moçambique Estimativas Básicas de Despesas e Receitas (2008-9) by Province

D.1 - Media das despesas diarias dos agregados familiares
D.2 - Estrutura percentual da despesa diaria dos agregados familiares (MICRO)
D.3 - Media das despesas diarias per capita dos agregados familiares (MICRO)
D.4 - Media des despsesas diarias per adulto equivalent dos agregados fam. (MICRO)

D3.1 - Media das despesas diarias dos agregados familiares
D4.1 - Media das despesas diarias dos agregados familiares

DM.1 - Total des despesas diarias dos agregados familiares (em milhãoes de Meticais)
DM.2 - Estrutura percentual da despsesa diaria dos agregados familiares (MACRO)
DM.3 - Total des despesas diarias per capita dos agreg. fam. (em milhãoes de Meticais)

AggrProd - Total des despesas e receitas par produto

Tables comparing IOF 2008 and IAF2002 results:

National 2008/2002: media des despesas e receitas dos agregados familiares
Rural/Urban 2008/2002: media des despesas e receitas dos agregados familiares
Provincial 2008/2002: media des despesas e receitas dos agregados familiares

STATISTICS DENMARK

Statistisk sentralbyrå
Statistics Norway

Statistiska centralbyrån
Statistics Sweden

Mission Report

for a short-term mission as specialist in household surveys, especially in the area of household budget and expenditure

***From 2 November to 20 December* 2009**

Erwin Triebkorn

Instituto Nacional de Estatística

Ref: Contract DARH/2008 /004

Erwin Triebkorn

Erwin.t@gmx.de

Karl Liebknecht Str.14
D-15370 Petershagen

49-33439-65637

Table of contents

II.1 EXECUTIVE SUMMARY

The National Institute of Statistics (INE) conducted its third Household Budget Survey (IOF2008/9) from September 2008 to September 2009. The sample of 10848 households in 1060 enumeration areas (EA) represented the nation, rural and urban zones as well as the provincial level. In each urban EA 12 households were selected to be interviewed, and 9 households were selected in each of the rural EAs.

The present mission of the household survey specialist is a continuation of a previous one in June - July 2009, which focused on cleaning and analyzing the database for the first 6 months of the survey. Due to serious problems with the data processing management that work could not be finished in time. However, the experience in processing the IOF data for the first six months has helped to improve the final data processing.

Director Arão Balate defined cleaning the data base of IOF 2008/9, in particular, the data of the second 6 months as the main objective of the present mission.

When the mission began, the data entry for the 12 months of the survey had been completed. A working group headed by Ramiro Mozinho fom INE, was comparing the data in the file with the questionnaires. The household survey specialist gave his constant assistance to the consistency checking. A first attempt was made to aggregate daily expenditures per capita including the various components of consumption. This work has received valuable support from Ellen M. Payongayong but due to other engagements her support was limited in time.

On November 20 Carlo Azzarri from the World Bank started to support the processing of the IOF data, developing programs for consistency checks of the various modules of the survey. On December 1 the sampling expert David Megill started to work on the sample weights and calculation of sampling errors.

Thanks to the joint efforts of the IOF team the scrutiny and validation of the 12 months data for most of the modules has been finished and the sample weights calculated. The very first aggregated data on the total household expenditure are available. The analysis of these preliminary data shows that one can expect useful results. The final report on the IOF results is planned by INE to be completed in February 2010.

II.2 INTRODUCTION

The National Institute of Statistics (INE) conducted its third Household Budget Survey (IOF2008/9) from September 2008 to September 2009. The main objective of this household survey was to collect data for updating and improving the information on the living conditions and poverty in Mozambique, especially on the changes that have taken place since the previous survey in 2002/3.

A sample of 10848 households in 1060 enumeration areas was selected to ensure representative results for the nation, rural and urban zones as well as at the provincial level. The mission report

by David Megill, specialist in sampling for household surveys, has a detailed description of the sampling methodology. Therefore reference can be made to that report on these issues.

The interview period for the household lasted one week. During this time each household was visited by the enumerator at least three times to collect besides general characteristics of the household, data on daily expenses, consumption of products produced by the household, possession and purchase of durable goods, expenditures on education and health, housing conditions and others. The data collection in the field finished in September 2009 with a total of 10766 households with completed interviews. A total of 1079 AF had to be substituted by households from a preselected list of random reserve households. That gives a replacement rate of 9.9 percent. The main reason of the substitution was absence of the originally selected households.

Unlike the previous surveys (IAF1996/7 and IAF2002/3) which were conducted with permanent technical assistance from abroad, IOF2008/9 was prepared and conducted by INE without such assistance. Although, there can be found some weakness in the design of the questionnaires and in the manuals, INE has proofed that it is able to conduct such comprehensive and complicated survey by its own staff.

Serious problems arose because of the limited capacity in data processing. Actually, it is of vital importance, both from the viewpoint of quality of data and of timely generation of the survey output that preliminary processing operations initiate at the very beginning of the survey. But the experienced processing staff of INE was engaged in the processing of the Census of Population. A pilot survey was conducted to check the survey instruments and the field organisation but the data were not processed.

Originally it was planned to process the data of first 6 months and produce a preliminary report. The specialist in household surveys assisted in these activities during his previous mission in June – July 2009. Due to limited capacity in programming and data processing the data cleaning could not be finished at that time and was continued by INE staff after that mission.

II.3 ACTIVITIES DURING THE MISSION

During a meeting with Director Arão Balate and the Chief of Department Christovão Muhaio on November 3 the household survey specialist was informed about the actual status of the work. At that meeting Arão Balate defined cleaning the data base of IOF 2008/9, in particular, the data of the second 6 months as the main objective of the present mission.

At the beginning of the mission the data entry for the 12 months had been finished. The entry of the data from the second semester used a new application developed by Ellen Payongayong and ran much better than the first one. The data were entered twice in a dependent matter. During the second round the entries were automatically compared to the first one and in case of not matching were rejected for validation. Although, compared to the first 6 months, the data quality had improved there were still many errors and missing values. Most probably the data entry staff had forced the entry. The data cleaning was still affected by constraints in programming. The cleaning was not systematic because there was no comprehensive program. The data cleaning focused mainly on the demographic indicators, the household expenditures and other components of consumption. The main problem was linked with checking the quantities of products bought or produced and consumed by the household. A large part of products is sold in local units of measurement instead of standard units, i.e. kg or litres. The list

of local units of measurement contains more than 50 categories. The most common units are can, heap, sack and cup, but pot, basket and even a handful are also used. The enumerator had to transform the quantities measured in local units to standard units. Although the enumerators should have used a balance many questionable data have been found.

A price list was elaborated on the basis of the prices collected by the supervisor in the local markets. For products not included in the community questionnaire the price was found using the data on purchases effected by the household. To avoid the impact of extreme values the median price instead the average was applied. In that way the price was calculated according to province, zone, urban and rural area and for the country as the whole. The number of observations had to be 10 or more. The price list was used for checking the daily expenditures, consumption of products produced and consumed by the household as well as products received in kind as a payment for work. It should be emphasized that staff from the Price Department and National Accounts Department took part in this work.

Ellen Payongayong assisted in this work but due to her other commitments she could not spend sufficient time in supporting IOF. The situation has improved, after Carlo Azzarr had joined the team. He is an expert in data processing on a short term mission financed by the World Bank. The data processing experts suggested restricting the validation to ensure conformity of the data base with those in the questionnaires. They recommended correcting data by computer based on general rules instead of individual decisions. That procedure speeded up the validation process. By the end of the mission the validation of the main components of the survey had been finished. That means, the data base is in conformity with the data in the questionnaires and necessary corrections can be done by computer without going back to the questionnaires. However, before these corrections can be done, the following cases need more studying and taking decision:

There are many missing data on rents paid by households and rental values of owner-occupied or rent-free housing enjoyed by households. These rents are to be estimated by data collected on rents paid for houses with similar facilities. But in Mozambique the renting of houses is not a common practise. The provisional data of IOF 2008/9 show that about 10% of the households in urban areas and less than 1% in rural areas live in rented houses. The limited number of households living in rented dwellings does not permit the estimation of rent on the local level. More investigation is needed to find a suitable solution.

A considerable number of households were found which declared no expenditures or an insignificant amount and nothing in own consumption was found. This seems to be biased. But sometimes a note was found in the questionnaire that the household had bought products in bulk before the reference period or received help from the neighbours, without providing any details on type and amount of this help. It remains to be decided how to treat these cases.

The danger of double counting occurred when the enumerator put the same items into the daily as well as monthly expenditures. The questionnaire on monthly expenditures (DM) was filled out during the second visit, i.e. 3 days after the first visit. The calculation of the daily purchases has to consider this.

The validation and correction process has to be continued. The working group headed by Christovão Muhaio carries on with its activities. In January 2010 Carlo Azzarro will continue giving his support.

A first attempt was made to aggregate the consumption at the household level including the following components:

- Daily expenditures,
- Consumption of products produced by the household,
- Income in kind from work,
- Value of bought durable goods,
- Monthly expenditures,
- Expenditures on education,
- Expenditures on health and other parts of monthly expenditures.

These data had been collected using different questionnaires and for different reference periods. They had to be transferred into a unique period, i.e. per day or month.

From the first very preliminary data from the IOF2008/9 we found that the monthly consumption expenditure per person at current prices amounts to 684 Metical. Large differences were found between the level of expenditure in urban areas (1019 Metical) and rural areas (536 Metical). In comparison with the IAF 2002/3 the expenditure per capita at current prices has more than duplicated. But taking into account that in the same time the prices increased by 188 percent, we get a real increase of expenditure per capita of about 10 percent.

It seems that the IOF2008/9 delivered a good basis for a deeper going analysis on the development of living conditions of the population in Mozambique, but more work is necessary to get comparable data with those from the previous survey.

II.4 RECOMMENDATIONS

1. The Government of Mozambique's Action Plan for the Reduction of Absolute Poverty for 2006-09 (PARPA II) aimed at reducing the incidence of poverty from 54 percent in 2003 to 45 percent in 2009. Therefore the main objective of IOF2008/9 was to find out how this aim has been achieved. This requires absolute comparativeness of the results from the two surveys. The processing of the final data of IOF2008/9 should follow the same procedure as it was applied for the previous survey. There exists a complete documentation of the methodology applied in 2002/3. If for some reason that will not be accepted than the IAF2002/3 data should be processed again following the new procedures.

2. It is very important, both from the viewpoint of quality of data and of timely generation of the survey output, to develop the preliminary processing system soon after the survey instruments, i.e. the questionnaires and manuals, been adopted.
The consultant in data processing, James Otto, wrote in his final report after the IAF 2002/3:" INE has developed its capacity to collect and process survey data efficiently. However, the capacity to design and develop survey data processing systems and to analyze and publish survey results needs to be considerably strengthened. Because of heavy pressure for fast survey results, technical assistance for these activities has been biased towards producing results rather than building capacity." It is of vital importance that INE takes measures to use better the technical assistance in data processing for capacity building.

APPENDIX II.1. Persons met

INE
Gaspar, Manuel da Costa, Vice-President
Balate, Arão, Director of surveys and censuses
Zacarias, Fátima, Director of demographic and vital statistics
Muahio, Cristóvão, Head of Department
Matavel, Eugenio, data processing supervisor
Mozinho, Ramiro, data cleaning supervisor

Carlsson, Lars, Scanstat resident advisor

Consultants
Megill, David, Consultant in sampling
Payongayong, Ellen, Consultant in data processing
Azzarr, Carlo, Consultant in data processing

APPENDIX II.2 Terms of Reference

September 22, 2009

TERMS OF REFERENCE

for two short-term missions on the

The Household Budget Survey 2008/9

26 October – 20 December, 2009

within the

AGREEMENT ON CONSULTING IN INSTITUTIONAL CAPACITY BUILDING, ECONOMIC STATISTICS AND RELATED AREAS
between INE and Scanstat.

Consultants: 1. Erwin Triebkorn, 2. David Megill

Counterparts: The consultants will further work in close coordination with the staff of the Family Budget Survey (IOF 2008/09) at INE, headed by INE Vice President Manuel Gaspar, Director Arão Balate and Director Fatima Zacarias.

Background
In the Strategic Plan 2008 - 2012 one of the priority activities is the execution of the Household Budget Survey 2008/09, which has as some of its main objectives to gather data making it possible to evaluate the Action Plan of Reduction of Absolute Poverty (PARPA II) and monitor the fulfillment of the Millennium Goals.

Main reasons for the Missions
During the period 2 June to August 5 the National Institute of Statistics had three short term experts from Scanstat here to assist in the development of an intermediate report built on data from the first 6 months of the HBS. INE is now requesting that the experts return again to assist in database cleaning and in basic analysis of the full 12-month information, taking into account the multi sectoral data that this survey include. For this reason a combined advisory mission to support INE is necessary. The following expertise is needed already this year:

4. **A specialist in household surveys**, especially in the area of household budget and expenditure;

5. **A specialist in sampling for household surveys**, with experience in household budget surveys;

Taking into account the multi sectoral data that this survey include a third expert, a **specialist in poverty analysis,** is needed next year when data is fully available and the INE staff has possibility to participate in training activities using the income and expenditure information from the households.

Beneficiaries of the mission
The mission will benefit the National Statistical System in general and in particular the National Statistical Institute which will get further technically trained by this mission.

Objectives
The overall objective of the mission is to support INE making a HBS report using data gathered during the full twelve months of data collection in the field. The experts are supposed to contribute in the following way:

1. The specialist in household surveys, especially in the area of household budget and expenditure;
    a. Support INE in the verification of the consistency check program;
    b. Support the INE in the final cleaning of the database;
    c. Support INE in the implementation of the tabulation plan;
    d. Support INE in preparing the draft report;
    e. Support the INE in the documentation of all procedures described above.
    f. Support the INE with the consultants experience in this type of surveys and to perform other related activities;
    g. Duration of the mission; 8 weeks (tentatively October 26 to December 20, 2009).

2. The specialist in sampling for household surveys with experience in household budget surveys will continue the work done during the previous period;
    a. Support INE in calculating the weights for all variables of the survey
    b. Support INE in expanding the results to different levels in all aspects;
    c. Support INE with the master sample design
    d. Support the INE with calculation of the final IOF weights
    e. Support INE with calculation of sampling errors for selected IOF indicators,
    f. Support INE in the documentation of the procedures described previously;
    g. Support INE with the consultants experience in this type of surveys and to perform other related activities;
    h. Duration of the mission; 2 weeks (tentatively November 30 to December 13, 2009)

The specialist in poverty analysis, using income and expenditure information from the households will have the ToR written separately later this year as the mission not can be made before February or March 2010.

Expected results
1. Completion of the consistency check;
2. A clean database containing the full twelve months;
3. Tables generated on the basis of the full twelve months;
4. A draft preliminary report.

Activities
The detailed agenda of the missions will be specified in the first day together with the counterparts.

Tasks to be done by INE to facilitate the mission

1. Elaborate ToR for the mission
2. Prepare and supply the consultants with necessary documents and information, such as mission reports, strategies, plans etc.
3. Supply good working conditions for the consultants.

Source of Funding

Project: MPD – 2008 – 0006 – Inquérito Sobre Orçamento Familiar – IOF

PAAO09 – 1.4.6 Inquérito ao Orçamento Familiar - IOF 2008/2009

Timing of the mission

See above

Place

The premises of the National Institute of Statistics in Maputo

Language

Portuguese

Report

The consultants will prepare a draft report (to be added to the one made during the 6-months phase) in Portuguese with findings and recommendations which must be agreed with the counterparts of the INE. The final draft must be submitted for final comments by INE, within a week after the completion of its mission. The structure of the report will be agreed with the counterparts and most of the content will be the recommendations.

Statistics Denmark as Scanstat Lead will publish the final version on www.dst.dk/mozambique within 3+ weeks of the end of the mission.

These Terms of Reference were prepared by Arão Balate INE/DCI

Day / / .....

Approved by Luis Mungamba, Contract Manager for the INE – Scanstat Contract

Day / / .....

STATISTICS DENMARK

Statistisk sentralbyrå
Statistics Norway

Statistiska centralbyrån
Statistics Sweden

Mission Report

for a short-term mission of the specialist in sampling for household surveys

From 30 November to 19 December 2009

David J. Megill

Instituto Nacional de Estatística

.

Ref: Contract DARH/2008 /004

David J. Megill
davidmegill@yahoo.com

1504 Kenwood Ave.
Alexandria, VA 22302

1-703-824-0292

Table of contents

III.1 INTRODUCTION AND TERMS OF REFERENCE

The *Instituto Nacional de Estatística* (INE) conducted the *Inquérito sobre Orçamento Familiar* (IOF) 2008/09, or Household Budget Survey, in a nationally-representative sample of 10,848 households in 1,060 sample census enumeration areas (EAs) over the 12-month period from September 2008 to August 2009. A new listing of households was conducted in each of the sample EAs. In each sample urban EA 12 households were selected to be interviewed, and 9 households were selected in each of the sample rural EAs. The sample EAs for IOF 2008/09 are divided into four nationally-representative subsamples, and one subsample was assigned to each quarter. The sampling consultant previously participated in a team that worked with the IOF data for the first 6 months of data in June/July 2009, and he calculated the weights for the corresponding sample at that time.

Now that the 12 months of IOF data are available, the IOF team is completing the final phase of data editing. The sampling consultant worked with Carlos Creva, Statistician, and other INE staff on the calculation of weights for the full IOF sample.

The following are the terms of reference for the sampling consultant:

a. Support INE in calculating the weights for all variables of the survey;
b. Support INE in expanding the results to different levels in all aspects;
c. Support INE with the calculation of the final IOF weights;
d. Support INE with the calculation of sampling errors for selected IOF indicators;
e. Support INE in the documentation of the procedures described previously;
f. Support INE with the consultant's experience in this type of surveys and to perform other related activities

Although the master sample design was originally included in the terms of reference, INE separated this activity into a different contract funded by the UNFPA given the large level of effort over an extended period of time that would be required. The calculation of sampling errors for selected IOF indicators can only be accomplished once the IOF data edits are complete and the survey data file is considered final. Therefore this activity may be followed up early next year under a separate contract if necessary. The main activity of this mission, the calculation of the IOF weights, was only completed by the end of the second week, given that it took time to compile the listing summary information for all the sample segments needed for the calculation of the weights.

The sampling consultant had previously worked with Carlos Creva, INE Statistician, in developing the sample design for IOF 2008/09. The sampling methodology is documented in the report "Sampling Recommendations for the 2008/09 Mozambique *Inquérito Sobre Orçamento Familiar*" (Megill, March 2008). This report also describes the procedures for calculating the weights.

The mid-term mission report by Erwin Triebkorn, the Scanstat specialist in household surveys, has a detailed description of the status of the processing and editing of the IOF 2008/09 data. Therefore reference can be made to that report on these issues.

The sampling consultant again worked closely with Carlos Creva, INE Statistician, in implementing the weighting procedures for the IOF 2009. He also collaborated with other INE staff, his Scanstat consultant colleague, Irwin Triebkorn, Carlo Azzarri of the World Bank and Ellen Payongayong, Resident Advisor with Michigan State University, during this mission. He would like to thank Dr. João Loureiro, INE President, Manuel Gaspar, INE Vice-President, Arão Balate, Director, *Direcção de Censos e Inquéritos*, and Cristóvão Muahio, Chief, *Departamento de Metodologia e Amostragem* (DMA), as well as his Scanstat colleagues, for this collaboration.

III.2 ACTIVITIES DURING THE MISSION

The sample design and estimation procedures for the IOF 2008/09 are described in the report on "Sampling Recommendations for the 2008/09 Mozambique *Inquérito sobre o Orçamento Familiar*" (Megill, March 2008). The first activity during this mission was to obtain all the information needed for calculating the weights for the full 12 months of IOF 2008/09 data. Creva had compiled all the information on the geographic codes and the number of households for each IOF sample EA from the sampling frame into a spreadsheet that was used for calculating the weights. The sampling consultant used this spreadsheet to compile the information on the total number of households listed in each sample EA for IOF, the number of sample households with completed questionnaires by EA, and other information required for the calculation of the weights. The final distribution of the selected EAs and households with completed questionnaires for the full IOF sample are presented in Table 1. The SPSS file with the household sample control information from the IOF data was used to tabulate the information in Table 1. The number of sample EAs in each stratum matches the original sample distribution exactly. The IOF enumeration procedures included the substitution of non-interview sample households from a preselected list of random reserve households. Therefore most of the urban sample EAs have 12 completed household questionnaires and most of the rural EAs have 9 completed questionnaires.

Table III.1 Distribution of Sample EAs and Households with Completed Questionnaires for IOF 2008/09 by Province, Urban and Rural Stratum

	Total		Urban		Rural	
Province	No. Sample EAs	No. Sample House-holds	No. Sample EAs	No. Sample House-holds	No. Sample EAs	No. Sample House-holds
Niassa	80	812	32	384	48	428
Cabo Delgado	80	780	20	240	60	540
Nampula	160	1,575	48	570	112	1,005
Zambézia	160	1,523	28	336	132	1,187
Tete	80	764	16	191	64	573
Manica	80	803	28	335	52	468
Sofala	80	847	44	524	36	323
Inhambane	80	792	28	330	52	462
Gaza	80	790	32	369	48	421
Maputo Província	80	886	60	711	20	175
Maputo Cidade	100	1,194	100	1,194	-	0
Mozambique	1,060	10,766	436	5,184	624	5,582

The sampling consultant also checked the geographic codes in the sample control file and found a few coding problems in the IOF data. These cases were corrected with the assistance of Carlo Azzarri and Ramiro Mousinho. For some households the number of persons in the sample control file (from the first page of the questionnaire) was different from the actual count of persons in the questionnaire roster. Therefore Carlo Azzarri corrected the sample control file with the actual number of persons in the roster. This information was needed for tabulating the weighted total population used for the calculation of the weight adjustment factors, as explained later in this report.

In the case of the non-interview households that were replaced, the reason for the non-interview of the original household was recorded in a separate replacement control sheet. The summary information on the replacements by segment was used to determine the percent of sample households replaced by province as well as the distribution of the non-interviews by reason. These results are summarized in Table 2. It can be seen that the percentages for the non-interviews by reason do not add up to the total percentage of replacements; this is because of some missing information on the reason for non-interviews. Table 2 indicates that the overall replacement rate for sample households was 9.9%, but the rate varied considerably by province, from 5.7% for Nampula to 18.4% for Manica. According to this table a large proportion of the non-interviews were because there was no respondent at home (absence), and there were few cases of refusals. However, in some provinces the reason for the non-interviews may not have been well controlled, so the non-interview rate of 18.4% for the absent category in Manica probably includes some refusals and other types of non-interview.

Table III.2 Sample Household Replacement Rates by Province, and Distribution of Non-Interviews by Reason

Province	No. Sample Households	Replacements		% Non-Interview Households		
		No. Hhs.	%	Absent	Refusal	Other
Niassa	816	77	9.4%	6.7%	0.5%	1.3%
Cabo Delgado	780	141	18.1%	18.1%	0.0%	0.0%
Nampula	1,584	90	5.7%	3.7%	0.3%	0.6%
Zambézia	1,524	138	9.1%	7.5%	0.4%	1.2%
Tete	768	52	6.8%	6.0%	0.1%	0.7%
Manica	804	148	18.4%	18.4%	0.0%	0.0%
Sofala	852	59	6.9%	6.5%	0.0%	0.1%
Inhambane	804	78	9.7%	8.2%	0.4%	1.0%
Gaza	816	63	7.7%	7.7%	0.0%	0.0%
Maputo Província	900	128	14.2%	6.8%	0.0%	0.0%
Maputo Cidade	1,200	105	8.8%	7.7%	0.2%	0.9%
Mozambique	10,848	1,079	9.9%	8.3%	0.2%	0.6%

As explained in the sampling documentation, the basic weights were calculated as the inverse of the overall probabilities of selection, taking into account the probabilities at each sampling stage. Based on the sample design, the basic weights for the IOF data can be expressed as follows:

$$W_{hi} = \frac{M_h \times M'_{hi}}{n_h \times M_{hi} \times m_{hi}},$$

where:

W_{hi} = basic weight for the sample households in the i-th sample EA in stratum h

M_h = total number of households in the frame for stratum h (based on the preliminary 2007 Mozambique Census frame)

M'_{hi} = total number of households listed in the i-th sample EA in stratum h

n_h = number of sample EAs selected in stratum h

M_{hi} = total number of households in the frame for the i-th sample EA in stratum h (based on the preliminary 2007 Mozambique Census frame)

m_{hi} = number of sample households selected in the i-th sample EA in stratum h (12 households in urban EAs and 9 households in rural EAs)

Although most of the non-interview sample households were replaced, there were some EAs where the number of completed interviews was less than the number of households selected. Therefore it is still necessary to have a weight adjustment factor to take into account any non-interviews or incomplete questionnaires that were not replaced. Since the weights are calculated at the level of the sample EA, the weight adjustment factor for non-interviews is also calculated at this level. The adjusted weight (W'_{hi}) for the sample households in the i-th sample EA in stratum h can be expressed as follows:

$$W'_{hi} = W_{hi} \times \frac{m_{hi}}{m'_{hi}},$$

where:

m'_{hi} = total number of sample households with completed interviews in the i-th sample EA in stratum h, including replacement households

The weighting spreadsheet included all the information from the sampling frame for each EA in the IOF sample, as well as the listing summary data and the number of completed interviews. The spreadsheet includes formulas for calculating the basic sampling weight and adjusting the weight for any non-interviews based on the final distribution of the sample households with completed interviews.

The preliminary distribution of the weighted total number of households and population by province, urban and rural stratum was tabulated from the IOF 2008/09 data using the basic weights adjusted for non-interviews. The weighted total number of households was 4,278,529 (1,271,592 urban and 3,006,937 rural), and the weighted population was 20,013,406

(6,218,233 urban and 13,795,173 rural). This weighted total population is just slightly less than the corresponding population count from the 2007 Census (20,226,298), which is a good validation of the overall weights. However, when the distribution of the weighted total population by stratum from the IOF data was compared to the corresponding census population distribution, it was found that there were significant differences for some provinces, by urban and rural strata. This was probably mostly due to the differential quality of the listing in different areas. For some sample EAs there were considerable differences between the number of households listed and the corresponding number of households in the frame. Therefore it was decided to adjust the weights by province, urban and rural stratum, using population projections based on the 2007 Mozambique Census.

The demographers at INE used the exponential population growth rate for each stratum (province, urban and rural stratum) between the 1997 and 2007 censuses to estimate the projected population of each stratum for 31 March 2009, which is the approximate mid-point of the IOF data collection period. Table 3 presents the total population by province, urban and rural stratum from the 2007 Mozambique Census data, the estimated annual growth rate, and the population projections for 31 March 2009.

The weight adjustment factor based on the projected total population by stratum can be expressed as follows:

$$A_h = \frac{\hat{P}_{08h}}{\sum_{i \in h} \sum_{j} W'_{hi} \times p_{hij}},$$

where:

A_h = adjustment factor for the weights of the IOF sample households in stratum (province, urban/rural) h

$\hat{P}_{08h}$ = projected population for 31 March 2009 for stratum h (based on applying the growth rate to the 2007 Census population)

p_{hij} = number of persons in the j-th sample household of the i-th sample EA in stratum h

The denominator of the adjustment factor A_h is the estimated total population in stratum h from the IOF data using the preliminary weights.

Table III.3.Total Population by Province, Urban and Rural Stratum, from 2007 Mozambique Census, and Corresponding Population Projections for 31 March 2009

	Urban			Rural			Total
Province	Census Population (Aug. 2007)	Annual Pop. Growth Rate	Projected Population 31/3/09	Census Population (Aug. 2007)	Annual Pop. Growth Rate	Projected Population 31/3/09	Projected Population 31/3/09
Niassa	268,171	0.043	288,125	901,666	0.044	969,790	1,257,915
Cabo Delgado	334,476	0.043	359,488	1,271,173	0.017	1,308,004	1,667,492
Nampula	1,139,754	0.041	1,220,586	2,845,531	0.015	2,919,329	4,139,915
Zambézia	548,466	0.053	599,381	3,299,810	0.023	3,425,836	4,025,217
Tete	261,398	0.044	281,292	1,522,569	0.044	1,639,548	1,920,840

Manica	356,987	0.026	372,986	1,055,042	0.041	1,129,708	1,502,694
Sofala	628,469	0.017	646,155	1,014,167	0.029	1,064,710	1,710,865
Inhambane	263,307	0.018	271,242	989,172	0.009	1,004,349	1,275,591
Gaza	303,890	0.015	311,446	922,382	0.014	944,485	1,255,931
Maputo Província	829,679	0.049	900,957	375,874	0.022	390,190	1,291,147
Maputo Cidade	1,094,315	0.012	1,116,409	-	-	0	1,116,409
Mozambique	6,028,912		6,368,068	14,197,386		14,795,949	21,164,018

Table 4 shows the preliminary weighted estimates of total population by stratum from the IOF data, and the corresponding weight adjustment factors. These weighting procedures will result in a final distribution of the IOF weighted total population by stratum that is consistent with the corresponding population projections. Given that the editing of the IOF data has not yet been finalized, the number of completed interviews by EA may change slightly, so the calculation of the weights may have to be done again after the data are considered final.

Table III.4. Estimates of Total Population from the IOF Data Based on Basic Design Weights, by Province, Urban and Rural Stratum, and Corresponding Weight Adjustment Factors

	Urban			Rural		
Province	IOF Weighted Population, Preliminary Weights	Projected Population 31/3/09	Weight Adjustment Factor	IOF Weighted Population, Preliminary Weights	Projected Population 31/3/09	Weight Adjustment Factor
Niassa	248,226	288,125	1.1607	787,240	969,790	1.2319
Cabo Delgado	393,632	359,488	0.9133	1,443,195	1,308,004	0.9063
Nampula	986,197	1,220,586	1.2377	2,299,167	2,919,329	1.2697
Zambézia	652,511	599,381	0.9186	3,502,868	3,425,836	0.9780
Tete	190,648	281,292	1.4755	1,768,778	1,639,548	0.9269
Manica	187,372	372,986	1.9906	750,990	1,129,708	1.5043
Sofala	884,404	646,155	0.7306	923,624	1,064,710	1.1528
Inhambane	391,553	271,242	0.6927	968,068	1,004,349	1.0375
Gaza	390,218	311,446	0.7981	897,449	944,485	1.0524
Maputo Província	815,886	900,957	1.1043	453,794	390,190	0.8598
Maputo Cidade	1,077,586	1,116,409	1.0360	-	-	-
Mozambique	6,218,233	6,368,068		13,795,173	14,795,949	

III.3 FINDINGS AND RECOMMENDATIONS

The mid-mission report by Triebkorn includes a description of the issues related to IOF data edits for the IOF data from the full 12 months. The experience in processing the IOF data for the first six months has helped to improve the final data capture procedures, and the editing of the data is now more systematic.

For future household surveys at INE it is recommended to bring the listing sheets to the central office and store them in an organized manner like the questionnaires following the data collection, so that the listing summary data can be entered and verified. The lack of the listing sheets delayed the availability of the listing summary data for the calculation of the IOF weights. There is a need for improved verification of the geographic codes. There were

lessons learned from the IOF data processing experience also, especially in managing the data entry operations and the systematic editing of the data.

The weights for the IOF data have been validated. However, it was apparent from the weighted population by province, urban and rural stratum, that there may be a problem with the coverage of the listing in some areas. It is possible that it was difficult to identify the exact boundaries for some EAs when the census maps were not clear. Hopefully the census cartography can be improved for the set of sample primary sampling units (PSUs) in the new master sample once it is developed.

It is important to calculate the sampling errors and confidence intervals for the estimates of the most important indicators from the IOF data. The sampling consultant may assist with this work once the IOF data are considered final for tabulating the survey results, and the income and expenditure aggregates have been generated.

APPENDIX III.1 Persons Contacted

Instituto Nacional de Estatística

Dr. João Loureiro, INE President
Manuel Gaspar, INE Vice-President
Arão Balate, Director, *Direcção de Censos e Inquéritos*
Tomás Bernardo, Director-Adjunto
Cristóvão Muahio, Chief, *Departamento de Metodologia e Amostragem*
Eugênio Matavel, Programmer, INE
Carlos Creva, Sampling Statistician, INE
Nordino Titus, INE Data Processing Staff

Scanstat
Irwin Triebkorn, Household Surveys Consultant
Lars Carlsson, Resident Advisor

Michigan State University
Ellen Payongayong, Resident Advisor, Ministry of Agriculture

World Bank
Carlo Azzarri, Economist

STATISTICS DENMARK

Statistisk sentralbyrå
Statistics Norway

Statistiska centralbyrån
Statistics Sweden

Mid-Term Report

for a short-term mission of the specialist in sampling for household surveys

From 15 June to 16 August 2009

David J. Megill

Instituto Nacional de Estatística

Ref: Contract DARH/2008 /004

David J. Megill
davidmegill@yahoo.com

1504 Kenwood Ave.
Alexandria, VA 22302

1-703-824-0292

Table of contents

IV.1 INTRODUCTION AND TERMS OF REFERENCE

The *Instituto Nacional de Estatística* (INE) is conducting the *Inquérito sobre Orçamento Familiar* (IOF) 2008/09, or Household Budget Survey, in a nationally-representative sample of 10,848 households in 1,060 sample census enumeration areas (EAs) over the 12-month period from September 2008 to August 2009. A new listing of households was conducted in each of the sample EAs. In each sample urban EA 12 households are selected to be interviewed, and 9 households are selected in the rural EAs. The sample EAs for the IOF 2008/09 are divided into four nationally-representative subsamples, and one subsample was assigned to each quarter. Therefore the first two quarters (six months) of the IOF data are representative at the national level. INE is processing and analyzing the first six months of data for the IOF 2008/09 to obtain preliminary results. This is a similar approach to that used for the *Inquérito aos Agregados Familiares* (IAF) 2002/03, which was the previous Household Budget Survey. This will also provide an opportunity to finalize all of the processing, tabulation and analytical procedures in preparation for producing the final results from the 12 months of IOF data.

The following are the terms of reference for the sampling consultant:

g. Support INE in calculating the weights for all variables of the survey;
h. Support INE in expanding the results to different levels in all aspects;
i. Support INE in the documentation of the procedures described previously;
j. Support INE with the consultant's experience in this type of surveys and to perform other related activities

Although the original duration of the mission was scheduled for 35 days, it was decided that it would be more effective to have a mission of 5 weeks, followed by remote part-time consulting over a longer period of time for the remaining days of the contract. The consultant's mission in Mozambique began on 15 June and ended on 16 July, and the part-time remote assistance may continue as needed for another 6 to 8 weeks. Given the delay in the availability of the IOF data, this schedule will make it possible for the sampling consultant to remotely assist with the analysis of the data for the first six months after the final editing has been completed and the income and expenditure aggregates have been generated.

The sampling consultant had previously worked with Carlos Creva, INE Statistician, in developing the sample design for IOF 2008/09. The sampling methodology is documented in the report "Sampling Recommendations for the 2008/09 Mozambique *Inquérito Sobre Orçamento Familiar*" (Megill, March 2008). This report also describes the procedures for calculating the weights.

The mid-term mission report by Erwin Triebkorn, the Scanstat specialist in household surveys, has a detailed description of the status of the processing and editing of the IOF 2008/09 data for the first six months. Therefore reference can be made to that report on these issues.

The sampling consultant worked closely with Carlos Creva, INE Statistician, in developing the weighting procedures for the IOF 2009. He also collaborated with other INE staff, his Scanstat consultant colleagues (Irwin Triebkorn, Astrid Mathiessen and Ellen Cathrine Kiøsterud), and Ellen Payongayong, Resident Advisor with Michigan State University, during this mission. He would like to thank Dr. João Loureiro, INE President, Manuel Gaspar, INE Vice-President, Arão Balate, Director, *Direcção de Censos e Inquéritos*, and Cristóvão Muahio, Chief, *Departamento de Metodologia e Amostragem* (DMA), as well as his Scanstat colleagues, for this collaboration.

IV.2 ACTIVITIES DURING THE MISSION

The sample design and estimation procedures for the IOF 2008/09 are described in the report on "Sampling Recommendations for the 2008/09 Mozambique *Inquérito sobre o Orçamento Familiar*" (Megill, March 2008). The first activity during this mission was to obtain all the information needed for calculating the weights for the first six months of IOF 2008/09 data. Creva had compiled all the information on the geographic codes and the number of households for each IOF sample EA from the sampling frame into a spreadsheet that was used for calculating the weights. The sampling consultant used this spreadsheet to extract the information for the IOF subsample of EAs assigned to the first two quarters. Summary data on the number of households listed in each sample EA were then added to this spreadsheet. The original distribution of the selected EAs and households for the first six months of IOF are presented in Table 1. This represents exactly one half of the overall IOF sample size.

Table IV.1.Distribution of Sample EAs and Households Selected for the First Six Months of IOF 2008/09 by Province, Urban and Rural Stratum

	Total		Urban		Rural	
Province	No. Sample EAs	No. Sample Households	No. Sample EAs	No. Sample Households	No. Sample EAs	No. Sample Households
Niassa	40	408	16	192	24	216
Cabo Delgado	40	390	10	120	30	270
Nampula	80	792	24	288	56	504
Zambézia	80	762	14	168	66	594
Tete	40	384	8	96	32	288
Manica	40	402	14	168	26	234
Sofala	40	426	22	264	18	162
Inhambane	40	402	14	168	26	234
Gaza	40	408	16	192	24	216
Maputo Província	40	450	30	360	10	90
Maputo Cidade	50	600	50	600	-	-
Mozambique	530	5,424	218	2,616	312	2,808

Once the first round of the IOF data entry and verification had been completed for the first six months of data, Nordino Titus, INE data processing staff, produced an SPSS control file which included information on each sample household in the IOF data file with the geographic identification number, interview result, number of persons in the household and other variables from the first page of the IOF questionnaire. The

sampling consultant used this file to tabulate the distribution of the sample households by EA and interview status; this information was needed to adjust the weights for non-interviews. When the sample EAs in this table were matched to the sample EAs in the weighting file, there was initially a problem with the geographic identification codes of some of the sample EAs. There were several cases of incorrect EA codes due to data entry problems, and a few sample EAs had been exchanged with EAs assigned to other quarters. Most of these EAs were reassigned between the first and second quarters, but one EA had been inadvertently exchanged with another one in the fourth quarter. Once the final composition of the sample EAs for the first six months was established, it was also found that some EAs had one more or one less sample household than the number selected. These cases were mostly due to data entry errors that resulted in households being assigned to the wrong EA. Therefore it was necessary to print out a list of all the household identification numbers appearing in the IOF data file by EA, which was then checked with a complete inventory of all the sample household identification numbers found in the box of questionnaires for each EA. In this way Ramiro Mousinho was able to identify the households with incorrect codes and correct the corresponding records in the data file. One sample EA was found to have the wrong province number in the data file, and this was also corrected.

During the IOF data collection, any sample households that could not be interviewed were supposed to be replaced by households from a preselected list of reserves. This procedure appears to have been implemented well. However, in many cases the field staff did not fill out the first page of the IOF questionnaire indicating the reason for a non-interview. Instead, this information was compiled separately in a field control form. This information will be summarized later and presented in a table for the report. The overall substitution rate was about 9.9 percent. There were IOF questionnaires identified as incomplete, but when these questionnaires were checked it was found that some of them were actually complete interviews, so the interview status was corrected for these cases.

Given problems found with the IOF data entry, INE decided to have the IOF data for the first six months keyed a third time. This situation is explained in more detail in Triebkorn's mid-term report. After the new IOF data entry was completed for the sample EAs enumerated in the first six months, Ellen Payongayong assisted with updating the SPSS control file with information on the identification codes and interview status for each household in the IOF sample for the first six months, and the process of summarizing the information by EA and matching to the EAs in the weighting file was repeated.

As explained in the sampling documentation, the basic weights were calculated as the inverse of the overall probabilities of selection, taking into account the probabilities at each sampling stage. Based on the sample design, the basic weights for the IOF data from the first six months can be expressed as follows:

$$W_{hi} = \frac{M_h \times M'_{hi}}{n'_h \times M_{hi} \times m_{hi}},$$

where:

W_{hi} = basic weight for the sample households in the i-th sample EA in stratum h, for the first six months of IOF data

M_h = total number of households in the frame for stratum h (based on the preliminary 2007 Mozambique Census frame)

M'_{hi} = total number of households listed in the i-th sample EA in stratum h

n'_h = number of sample EAs selected in stratum h for the first six months of IOF data collection, shown in Table 1; this is half of n_h, the number of sample EAs in the full sample for IOF

M_{hi} = total number of households in the frame for the i-th sample EA in stratum h (based on the preliminary 2007 Mozambique Census frame)

m_{hi} = number of sample households selected in the i-th sample EA in stratum h (12 households in urban EAs and 9 households in rural EAs)

Although most of the non-interview sample households were replaced, there were some EAs where the number of completed interviews was less than the number of households selected. Therefore it is still necessary to have a weight adjustment factor to take into account any non-interviews or incomplete questionnaires that were not replaced. Since the weights are calculated at the level of the sample EA, the weight adjustment factor for non-interviews is also calculated at this level. The adjusted weight (W'_{hi}) for the sample households in the i-th sample EA in stratum h can be expressed as follows:

$$W'_{hi} = W_{hi} \times \frac{m_{hi}}{m'_{hi}},$$

where:

m'_{hi} = total number of sample households with completed interviews in the i-th sample EA in stratum h, including replacement households

The weighting spreadsheet included all the information from the sampling frame for each EA in the IOF sample for the first six months, as well as the listing summary data and the number of completed interviews. The spreadsheet includes formulas for

calculating the basic sampling weight and adjusting it for any non-interviews based on the final distribution of the sample households with completed interviews.

The preliminary distribution of the weighted total number of households and population by province, urban and rural stratum was tabulated from the IOF 2008/09 data for the first six months using the basic weights adjusted for non-interviews. The weighted total number of households was 4,136,662 (1,296,578 urban and 2,840,084 rural), and the weighted population was 19,677,149 (6,509,014 urban and 13,168,135 rural). This weighted total population is just slightly less than the corresponding preliminary population count from the 2007 Census (19,915,382), which is a good validation of the overall weights. However, when the distribution of the weighted total population by stratum from the IOF data was compared to the corresponding census population distribution, it was found that there were significant differences for some provinces, by urban and rural strata. This was probably mostly due to the differential quality of the listing in different areas. For some EAs there were considerable differences between the number of households listed and the corresponding number of households in the frame. Therefore it was decided to adjust the weights by province, urban and rural stratum, using population projections based on the 2007 Mozambique Census.

The demographers at INE used the exponential population growth rate for each stratum (province, urban and rural stratum) between the 1997 and 2007 censuses to estimate the projected population of each stratum for December 31, 2008, which is approximately the mid-point of the IOF 2008/09 data collection for the first six months. Table 2 presents the total population by province, urban and rural stratum from the 2007 Mozambique Census data, the estimated annual growth rate, and the population projections for 31 December 2008.

The weight adjustment factor based on the projected total population by stratum can be expressed as follows:

$$A_h = \frac{\hat{P}_{08h}}{\sum_{i \in h} \sum_{j} W'_{hi} \times p_{hij}} ,$$

where:

A_h = adjustment factor for the weights of the IOF sample households in stratum (province, urban/rural) h for the first six months of IOF data

$\hat{P}_{08h}$ = projected population for 31 December 2008 for stratum h (based on applying the growth rate to the 2007 Census population)

p_{hij} = number of persons in the j-th sample household of the i-th sample EA in stratum h

The denominator of the adjustment factor A_h is the estimated total population in stratum h from the IOF data for the first six months using the preliminary weights.

Table 3 shows the preliminary weighted estimates of total population by stratum from the IOF data for the first six months, and the corresponding weight adjustment factors.

Table IV.2.Total Population by Province, Urban and Rural Stratum, from 2007 Mozambique Census, and Corresponding Population Projections for 31 December 2008

	Urban			Rural			Total
Province	Census Population (Aug. 1997)	Pop. Growth Rate	Projected Population 31/12/08	Census Population (Aug. 1997)	Pop. Growth Rate	Projected Population 31/12/08	Projected Population 31/12/08
Niassa	268,171	0.043	285,090	901,666	0.044	958,905	1,243,995
Cabo Delgado	334,476	0.043	355,683	1,271,173	0.017	1,302,504	1,658,187
Nampula	1,122,027	0.041	1,189,516	2,602,977	0.015	2,660,416	3,849,932
Zambézia	664,606	0.053	716,852	3,133,035	0.023	3,234,755	3,951,607
Tete	261,398	0.044	278,265	1,522,569	0.044	1,621,740	1,900,005
Manica	356,987	0.026	370,581	1,055,042	0.041	1,118,369	1,488,950
Sofala	628,469	0.017	643,515	1,014,167	0.029	1,057,098	1,700,613
Inhambane	263,307	0.018	270,057	989,172	0.009	1,002,096	1,272,152
Gaza	303,890	0.015	310,320	922,382	0.014	941,191	1,251,511
Maputo Província	829,679	0.049	890,068	375,874	0.022	388,044	1,278,112
Maputo Cidade	1,094,315	0.012	1,113,122	-	-	-	1,113,122
Mozambique	6,127,325		6,423,070	13,788,057		14,285,117	20,708,187

Table IV.3.Estimates of Total Population from First Six Months of IOF Data Based on Preliminary Weights, by Province, Urban and Rural Stratum, and Corresponding Weight Adjustment Factors

	Urban			Rural		
Province	IOF Weighted Population, Preliminary Weights	Projected Population 31/12/08	Weight Adjustment Factor	IOF Weighted Population, Preliminary Weights	Projected Population 31/12/08	Weight Adjustment Factor
Niassa	281,252	285,090	1.0136	784,722	958,905	1.2220
Cabo Delgado	450,103	355,683	0.7902	1,620,780	1,302,504	0.8036
Nampula	1,097,608	1,189,516	1.0837	2,145,247	2,660,416	1.2401
Zambézia	563,221	716,852	1.2728	2,973,665	3,234,755	1.0878
Tete	212,720	278,265	1.3081	1,703,741	1,621,740	0.9519
Manica	188,129	370,581	1.9698	668,368	1,118,369	1.6733
Sofala	802,307	643,515	0.8021	1,185,158	1,057,098	0.8919
Inhambane	400,699	270,057	0.6740	811,348	1,002,096	1.2351
Gaza	358,682	310,320	0.8652	743,875	941,191	1.2653
Maputo Província	955,078	890,068	0.9319	531,231	388,044	0.7305
Maputo Cidade	1,199,215	1,113,122	0.9282	0	-	-
Mozambique	6,509,014	6,423,070		13,168,135	14,285,117	

These weighting procedures will result in a final distribution of the IOF weighted total population by stratum that is consistent with the corresponding population projections. Given that the editing of the IOF data for the first six months has not yet been finalized, the number of completed interviews by EA may change slightly, so the calculation of the weights will have to be done again after the data are considered final for the preliminary analysis.

It is important to calculate the sampling errors and confidence intervals for the estimates of the most important indicators from the IOF data for the first six months.

The sampling consultant will assist with this work once the IOF data for this subsample are considered final for tabulating the preliminary survey results, and the income and expenditure aggregates have been generated.

IV.3 RECOMMENDATIONS

The mid-mission report by Triebkorn includes a detailed description of the issues related to IOF data entry for the sample EAs from the first six months and related delays in the data processing. One purpose of the processing and preliminary analysis of the IOF data for the first six months is to resolve such issues related to the data entry and processing, as well as the production of tables and analysis, so that the final IOF results based on the full data set for 12 months can be produced in a timely manner once the IOF data collection is completed. As part of this process it is also useful to produce a report on preliminary results from the IOF data for the first six months, especially given that a nationally-representative subsample of EAs was used for this period. This is similar to the approach used for the IAF 2002/03. The experience from processing the IOF data for the first six months presents a valuable opportunity to review the lessons learned and to make recommendations for the processing of the IOF data for the remaining sample EAs.

In reference to the IOF data processing, it was found that insufficient experienced staff were assigned at the beginning to develop the processing system such as the data entry and editing programs. This was mostly due to the concurrent urgency of the data processing for the 2007 Mozambique Census, which required considerable attention from the most experienced INE data processing staff. Although part-time assistance from a data processing consultant was provided later, much time had been lost because of the need to enter the IOF data more than twice. In the future, it is important to ensure sufficient data processing support from the planning stage. A data processing consultant should be added to the IOF team if necessary.

The data processing issues also affected the work related to the calculation of the preliminary weights. Even though the data entry included 100 percent independent verification, several problems were found with the coding of the EAs that had to be resolved in order to match the geographic codes in the data files to the corresponding codes in the sampling frame. This is critical to ensure that appropriate weights are assigned to the sample household records in the data files. This could have been avoided with effective quality and operational control for verifying the geographic identification codes, as well as batch control during data entry.

The preliminary weights for the IOF data from the first six months have been validated. However, it was apparent from the weighted population by province, urban and rural stratum, that there may be a problem with the coverage of the listing in some areas. It is possible that it was difficult to identify the exact boundaries for some EAs when the census maps were not clear. Hopefully the census cartography can be improved for the set of primary sampling units (PSUs) in the new master sample once it is developed.

The lessons learned from the processing and analysis of the first six months of IOF data can be used for improving the operational and quality control for the IOF data from the second six months. The data processing system is being improved with the assistance of a consultant in order to increase the effectiveness of the data entry and validation systems.

APPENDIX IV.1 Persons Contacted

Instituto Nacional de Estatística

Dr. João Loureiro, INE President
Manuel Gaspar, INE Vice-President
Arão Balate, Director, *Direcção de Censos e Inquéritos*
Tomás Bernardo, Director-Adjunto
Cristóvão Muahio, Chief, *Departamento de Metodologia e Amostragem*
Carlos Creva, Sampling Statistician, INE
Nordino Titus, INE Data Processing Staff

Scanstat
Irwin Triebkorn, Household Surveys Consultant
Astrid Mathiessen, Economist, Statistics Norway
Ellen Cathrine Kiøsterud, Economist, Statistics Norway
Lars Carlsson, Resident Advisor

Michigan State University
Ellen Payongayong, Resident Advisor, Ministry of Agriculture

World Bank
Dr. Antonio Nucifora, Economist, World Bank/Maputo

STATISTICS DENMARK

Statistisk sentralbyrå
Statistics Norway

Statistiska centralbyrån
Statistics Sweden

Mid-Term Report

from a short-term mission on IOF2008/09

From 15 June to July 31 2009

Astrid Mathiassen
Ellen Cathrine Kiøsterud

Instituto Nacional de Estatística

Ref: Contract DARH/2008 /004

Astrid Mathiassen
Phd, Researcher
Divsion for Development Cooperation, Statistics Norway,
e-mail: mss@ssb.no

Ellen Cathrine Kiøsterud
Adviser
Divsion for Development Cooperation, Statistics Norway,
e-mail: eck@ssb.no

Table of contents

V.1 EXECUTIVE SUMMARY

According to the terms of reference the overall objective of the mission was to support INE in making the preliminary HBS report using data gathered during the first six months of data collection in the field. Further the poverty specialist consultants should:

- Support INE in the desegregation of the results in various quintiles and levels of poverty
- Support INE in using the information to assess the levels of poverty
- Support the INE in the documentation of the procedures described previously
- Support INE with the consultants experience in this type of surveys and to perform other related activities

At the arrival the team met with INE President Dr. Loureiro, who told us to contribute towards building capacity in producing the poverty line within INE.

During meetings with INE-staff it became clear that there was not much room to start this work during the mission, but that INE aim to take a longer time perspective on building capacity on poverty line. One workshop, however, was undertaken to give a brief introduction to the poverty line and the approach taken by MPD. In the workshop we discussed the building blocks and the construction of the consumption aggregates. Two slightly different approaches are previously taken for computing the consumption aggregate by INE and MPD. It was decided that one for the first six month round should follow the same definitions as in the last survey. For the 12 months survey INE will, however, do some smaller change in the consumption aggregate. Data programs and explanation on how compute the aggregates have been made by the consultants.

Due to delays in the data processing, data analysis was undertaken only during the two last weeks of the mission by the consultants in INE. At that time the staff were very busy in the data processing and validation, thus, training and capacity building in this field was limited.

We recommend that in the time to come INE should be trained to analyse the data in terms of poverty and changes in welfare and that INE take a long term perspective on building capacity to produce the poverty line for the next IAF-round. Also, it is important that INE-staff get a sound understanding of the building blocks in the welfare aggregate and poverty analyses, to better be able to judge decisions on possible future changes in the questionnaires.

V.2 INTRODUCTION

The overall objective of the mission is to support INE making the preliminary HBS report using data gathered during the first six months of data collection in the field. In particular, the specialist in poverty analyses should contribute by using income and expenditure information from the households;

- Support INE in the desegregation of the results in various quintiles and levels of poverty;
- Support INE in using the information to assess the levels of poverty;
- Support the INE in the documentation of the procedures described previously;
- Support INE with the consultants experience in this type of surveys and to perform other related activities

The role of the poverty specialist was decided to be filled by two persons from Statistics Norway: Astrid Mathiassen and Ellen Cathrine Kiøsterud. Astrid Mathiassen, the principal poverty analysis specialist worked three weeks in INE, divided between the first week and two weeks towards the end of the mission. Ellen Cathrine Kiøsterud worked six weeks in INE.

In the previous IAF's INE has calculated the nominal consumption aggregate, while MPD has calculated the poverty line and the consumption in real terms. However, in the initial meeting with President Loureiro, he expressed an interest in building capacity in INE for producing the poverty line, so that INE eventually take the leading role in this work. Thus, the consultants spent some time elaborating on how INE is to build capacity in producing the poverty line.

As a proposal on how to build capacity during the mission, we submitted an attachment to Terms of Reference to specify our assignment. This resulted in a meeting on 15 July 2009 with Fatima Zacarias, Arão Balate, Lars Carlsson and ourselves with the following guidelines:

- The aim that INE being able to estimate the poverty line must be a long term project. There is an interest to build capacity in poverty analysis, but INE do not have the time/human resources at this point in time. When the production of data is finished there might be time to start such work. This means February 2010 onwards.

- During this mission, INE will not be aiming to understand the details in the method used by MPD in constructing the poverty line and doing poverty analysis, but there is an interest in understanding the framework. Thus it was decided that the poverty specialists should organize a small workshop for the INE staff involved in the validation process and the table report.

Due to delays in the data processing the preliminary IOF-data was not available before towards the end of the mission, and the final 6-months data were not ready when the consultant left so that some of the final analyses will be undertaken remotely.

V.3 ACTIVITIES DURING THE MISSION

The first six months of data were supposed to be ready processed when the Scanstat team arrived. During the first week, though, it became clear that the data would have to be re-entered due to issues which are covered in Mr Triebkorn's reports. The poverty analysis team could do significant parts of the work without final data, so initially this did not create too many problems for us. However, data was finished re-entered only about two weeks before the mission was over. From then on the cleaning process started. Thus, we did not get cleaned data before we left, which means we could not produce the finale 6-month aggregates, quintiles and

tables within the time frame set for the mission. Neither could we assist with analysing the results.

The other main problem caused by data not being ready was that the people who could have been involved in poverty analysis were busy preparing and validating the data all the way to the end of the mission. Significant parts of the work can take place before the final 12-month data arrives. Unfortunately INE did not have people available to participate in this process.

Data constructs

The poverty consultants made the data programs/syntax files for calculating the consumption aggregates and quintiles, which will have to be re-run on the final six months dataset. The syntax files are attached in the appendix. There is also a document describing the aggregates in more details.

The team reviewed the syntax files obtained from MPD showing how poverty line and household consumption aggregates were computed previously and compared it to the aggregate calculated by INE. In the previous IAF's INE has calculated the nominal consumption aggregate, while MPD has calculated the poverty line and thus the consumption in real terms. INEs focus has been to calculate consumption expenditure aggregates for macro level analysis, while MPD focus on poverty analyses, and therefore concerned about ranking the households by their monetary welfare level. Thus, there has been some difference in how the household consumption aggregate has been calculated by the two institutions. While INE has included large and rare expenditure to durables (for example car, bicycle) MPD calculate a "user-value" of such goods. There is, however, some inconsistency in INE's approach in the 2002/03 survey. INE used only the weekly consumption of the purchased food in the aggregate, although total purchase of food during that week is available (the households were asked for how many days the food they purchased was going to last). If the focus of interest is to get the "right" consumption aggregate at macro level, then INE should rather use actual purchase of food in the consumption aggregate.

It was decided at the workshop that INE will create the aggregate in the same way as in the previous survey for the six months report to ensure comparability, but that they will correct the latter for the 12 months report and as well as compute the welfare aggregate to be used for poverty analyses when the full 12 months data is ready. Fatima Zacarias and Arão Balate were present at the workshop and part of making the decisions.

Workshop

There were approximately 15 participants in the workshop, the Scanstat team and INE-staff. MPD was invited to give a brief presentation of the approach taken to calculate the poverty line in Mozambique. Unfortunately they did not attain.

The target group was INE staff working with the IOF-survey, and the workshop had two main goals:

Primarily to increase the understanding of the requirement/the building blocks for constructing household consumption expenditure aggregates, and how to use the information in the IOF-survey for this purpose. In particular, to get a better understanding of the two approaches used by INE and MPD.

Secondly to start the learning process towards understanding the whole process of computing the poverty line in Mozambique.

The PowerPoint presentation can be found in the appendix of the document.

Poverty Analysis

We had informal talks with other stakeholders in the poverty analyses; Antonio Nucifora in the World Bank and Arndt Channing and Azhar Hussain, consultants for MPD. They say that in principle it is a good idea to build capacity on poverty analyses in INE. The MPD consultants are also positive to bringing INE staff into the MPD poverty team for this purpose, however, they emphasize that it requires commitment and much training.

Questionnaire issues

A careful review of differences between the IOF 2008/09 questionnaires with the IAF 2002/03 questionnaire was done to discuss effect this would have on comparison between the surveys. The IOF-questionnaire was also analyzed in terms of quality and the findings have been documented in an attachment to this report. A major part of those findings were revealed through Triebkorn's work and recorded in cooperation with the team.

The final table report

Due to the delay of data, some issues that we had ambitions to cover in the six months report were postponed to the final report. As the six months data were not ready when we left, the responsibility for chapters in the six-months were divided between INE staff and Scanstat consultants. The tables will largely cover the same issues as in the last survey.

We had, however, some initial discussions with people from the analysis team on the final table report, where we commented on the present reports and gave some suggestions for further work.

For the previous household surveys INE has produced two reports. A pure table report quickly after the 12months dataset is ready and the IOF table report with text later on. For IAF 2002/03 a CD-rom was produced with documentation and contents from the CD-rom was made available on the Internet in 2009.

The suggested pure table report that we were shown makes sense because it provide information to people with special interest which cannot all be printed in the table report as it would make it way too long and complex. The Internet solution is a good start and could be a good place for such tables.

The Scanstat team thinks that INE produce a good table report. One way to improve on the report would be to do some more research of who the users of the table report are. What information do they ask for? What are they interested in? How can INE help them? INE could also read other countries' reports and use the best ideas.

It is important to use simple tables and figures in the text report to get messages through to media and other non-experts who reproduce results. Give good explanations for non-experts in the reports. In particular attach explanation for the method used, like the construction of the consumption/expenditure aggregate to the report. Also give the number of respondents, N, with the tables.

There was also a request that from INE that the Scanstat team to present other methodology/angles than the poverty line, that can be used to illustrate poverty and may be

included in the 12 months report. For this purpose The Scantstat team suggests to include some more tables on the consumption pattern for each quintiles, in particular the food consumption, and for the regions. Further, it would be interesting to look at seasonal poverty, whether there are large differences between the season and in particular for the lowest quintiles. We also are writing together a chapter focusing on changes in living conditions between the surveys. This could be done by simple illustrative tables, for example changes in the share that had good housing conditions (define good housing conditions) and changes in consumption pattern for deciles.

Recommendations

We think it is timely and important that INE takes over the production of the poverty line. This will strengthen the authority of INE as an institution, and build relevant capacity in INE. We also believe that it would give more credibility to the poverty line if it is produced by the statistical office and not by its main user (MPD), although the approach taken would be more or less the same.
As poverty line construction and poverty analyses require sound understanding of the field and good programming capabilities we recommend that INE take a long term perspective on building this competence. INE must decide on who is responsible for poverty analysis within the organisation, and assign suitable and committed people for this task.

We recommend that INE identify two or more staff for the purpose of poverty line/poverty analyses. This should be staff with an interest as well as programming capacity for doing this type of work. We suggest that INE design a long term plan and follow up for capacity building for these staff. Preferably one should aim to building capacity directly related to real assignments. These staff could also attain poverty line - and programming courses at other institutions.

In particular we suggest that INE consider a follow up course on the poverty prediction models as was undertaken by Statistic Norway consultants together with INE staff in 2005. That time we used the IAF2002/03 model to predict poverty for 2004/05 using IFTRAB survey data. The follow up would be to test the modelling approach using the new survey.

An advantage of building capacity in poverty analyses is that the staffs involved in the design of the questionnaire get a fuller understanding of the actual use of the various components in the questionnaire for poverty analyses, and thus are more competent to take decisions regarding suggestions for changes in the questionnaire. Thus the problems encountered in this year's questionnaire may be avoided in the future.

Further we suggest that INE produce the real term consumption/expenditure aggregates. This involves using the capacity already in INE for producing IOF-deflators.

Recommendations for follow-up mission

The consultants suggest taking part in two follow-up missions in the Household Budget Survey 2008/09 work.

First, we suggest a short mission overlapping with the end of Mr. Triebkorns mission in October/November 2009. The second mission could be around February 2010 as suggested by Vice-president Gaspar. The latter could be organized as workshops concerning poverty analysis in connection to the production of the table report.

The aim with the first mission would be to actually compute the consumption aggregates and quintiles together with INE staffs. Practically that means to sit together and write the programmes that generate the aggregate and the quintiles, and INE staff documenting the process as they do it. We have already made the programmes; however, the best way to learn for INE staff is by doing it themselves.

A part of the final work on the consumption aggregates is to decide on the practical solutions to the various aspects of the consumption aggregates (as opposed to theoretical elements as we discussed in the workshop), for example whether one use the impute rent, how to handle mission observations etc.. For this part it is crucial that it happens while Triebkorn is still around. Another reason for not simply reproducing this aggregate in February is that is much more inspiring to be a part of the team when the work is actually done. The other part of the capacity building would be the actual programming. We also think it would be a good start to work together on the construction of the consumption aggregate before we eventually come back in February to actually use this construct.

Attachments available on request

Description of construction of the consumption and expenditure aggregates

Data programme to create the Consumption and Expenditure aggregate

APPENDIX V.1 Persons Contacted

INE: President Dr. João Dias Loureiro, Vice-President Manuel da Costa Gaspar, Director Fátima Zacarias and Director Arão Balate

Other INE staff members: Cassiano Chipembe, Maria Alfeu, Xadreque Maunze, Olimpio Zavale, Gilberto Nhapure, Ramiro Mousino, Ernesto Samo and others.

MPD: Nelson Maximiano and Maymona,
Azhar Hussain and Arndt Channing, consultants for MPD

World Bank: Antonio Nucifora, Senior Economist

Norwegian Embassy: Tove Bruvik Westberg Ambassador and Thor Oftedal First Secretary

Scanstat: Erwin Triebkorn, David Megill and Lars Carlsson

APPENDIX V.2 Constructing consumption aggregate and poverty line A brief introduction - Workshop at INE, Maputo 20 July 2009 Astrid Mathiassen

Constructing consumption aggregate and poverty line

A brief introduction

Workshop at INE

Maputo 20 July 2009

Astrid Mathiassen

Introduction

- "Poverty can be said to exist in a given society when one or more persons do not attain a level of material well-being deemed to constitute a reasonable minimum by the standards of that society"
  - Ravallion (1992)

- the poverty line is the cut off point that defines "that reasonable minimum"

Introduction

- Choose a welfare indicator for monetary well-being.
  - Income or expenditure?
  - Perspective:
    - ranking/comparing households/individuals
    - average consumption per capita for populations

- Define the poverty line: The cut-off point between poor and non-poor.

Outline

- How to compute/construct consumption aggregates, with empirical reference to IAF/IOF
  - Difference in approach previously taken by INE and MPD

- How to compute a poverty line, with empirical reference to Mozambique

Consumption versus income

Income is considered less suitable than consumption

- **Income often comes in lump sums. Consumption is smoother. Particularly in agricultural economies like Mozambique.**
- **Practically, it can be difficult to gather income data (for example when self-employed)**
- **Consumption can be viewed as realized welfare, while income is a measure of potential welfare.**
- **Households are often more willing to truthfully report their consumption than their income.**
- **Consumption data is more expensive to collect, but concept is clearer.**

Household consumption expenditure aggregate

- **Monetary measure of welfare is the sum of the values of all goods and services used by a household.**
- **As each questionnaire is unique, there will be considerable variation among countries and surveys with respect to how this is done.**

Consumption-expenditures normally included

Purchases, consumption of own production and gifts

- **Food**
- **Non-food consumable goods**
- **Durable goods**
- **Housing and shelter**
- **Other housing expenses**
- **Transport**
- **Education**
- **Health**

Consumption-expenditures normally excluded because it is difficult to measure

- **Value of goods and services received for free.**
  - **Public goods; Leisure; cooking etc.**
  - **Theoretical and practical difficulties. Major problems: use of - and shadow price.**

- **Lumpy and relatively infrequent expenditures such as marriages an funerals. (Rare, but high share of expenditure, problems with ranking).**

Consumption-expenditures that should not be included

- **Expenditure that have no impact on current welfare level or represent investment in household production: i.e. taxes and monetary transfers and repayments of loans.**

Food

- **Should express total value of food consumed in the reference period: food bought, consumption from own production and gifts/remittances/in-kind.**
  - **Recall period**
  - **Several designs in the questionnaires for capturing food consumption information**

- **IAF 2002/03:**
  - **Collects purchases the last 7 days and ask how many days the product was for. Excludes value of consumption over 7 days in the aggregate.**
  - **Value of consumption of own production last 7 days.**
  - **Value of daily gifts/payment in kind last 7 days.**
  - **INE and MPD used same methodology.**

Durable goods

- **Last for several years, it is the *use* of durables that is important for welfare.**
- **Compute "user-value". Ideally one needs information on lifespan on assets, age, value at purchase time and current value. (all this information is not available in IOF)**
- **IAF 2002/03:**
  - **INE: added expenditure to durables in the last 12 months. Reasonable when computing the average consumption in a population, but for ranking the households better to use MPD approach**
  - **MPD: compute user value assuming the average age on the durables if they were not new.**

Housing and shelter (durables)

- **For those who rent house, use the rental cost**
- **For owner occupied housing:**
  - **Use estimate on rent given by households themselves.**
  - **Impute rent based on housing characteristics**

- Relation between reported rent and housing characteristics
- Estimate property value and than user value

- IAF 2002/03:
  - INE: for those with missing, National Accounts estimated values for type of dwelling
  - MPD: estimate hedonic regression models. Enough observation for this approach?

Issues in the questionnaries IAF/IOF

- Consumption from stock (purchases made before the reference period)
  - Asks for some products, but lack information on reference period so the answers cannot be used in the aggregate.
  - Can be solved on macro level by using the full amount of purchased food for all households, but problem for poverty analysis.
- Consumption of food received as gifts
  - Do not have number of days for consumption.
- More products added to monthly expenditure questionnaire
  - Effect of this?

Poverty line

- Poverty line = food component + non-food component
- Food component/Food poverty line
  - Anchored in caloric requirement
  - What is the cost of the average caloric requirement in the population?
    - An adult man require more calories than a child
    - Requirement differs according to activity level
  - Define a food basket, consisting of the food the poor normally eat. Food poverty line is the cost of that basket adding up to the caloric requirement. Using the prices the poor normally pay.

Poverty line, continued

- Non-food component
  - Basic non-food consumption needs
  - In IAF non-food share is added based on the actual share to non-food for those household with total consumption “around/near” the food-poverty line.

Poverty line

- When a new expenditure survey is available
  - Update poverty lines by adjusting for price changes?
  - Make new poverty line (new basket)?
- One poverty line for whole country or separate for urban/rural/regions (issue: differences in relative prices)

The most important choices of MPD

- 'Flexible' food basket
  - With flexible food basket, poverty went down from 69% to 54% from 1996/97 to 2002/03
  - With fixed food basket poverty went down from 69% to 64% from 1996/97 to 2002/03
- Seven 'poverty lines regions'
  - (do not correspond geographically to the provinces)

Recommendations on the computation of the household aggregate

- Different needs – different aggregates:
  - Average consumption or ranking of the households?

Core readings

- Specific recommendation for consumption aggregate.
  - See Deaton and Gosh (1998): "Consumption" in Designing Household Survey Questionnaires for Developing Countries
  - Deaton and Zaidi (2002) http://siteresources.worldbank.org/INTPRS1/Resources/Thematic-Workshops/415743-1089658785131/Training_2001-27-02_Deaton_ConsumpAgg_doc.pdf
- Poverty line
  - Ravallion (1994) "Poverty Comparison"
  - Ravallion (1998) "Poverty lines in theory and practice"
- The poverty line Mozambique
  - "Poverty and Well being in Mozambique" http://www.mpd.gov.mz/gest/documents/3E_Poverty.pdf

STATISTICS DENMARK

Statistisk sentralbyrå
Statistics Norway

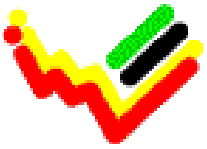
Statistiska centralbyrån
Statistics Sweden

Report

for a short-term mission as specialist in household surveys, especially in the area of household budget and expenditure

***From 1 June to 8 August* 2009**

Erwin Triebkorn



Instituto Nacional de Estatística

Ref: Contract DARH/2008 /004

Erwin Triebkorn

Erwin.t@gmx.de

Karl Liebknecht Str.14
D-15370 Petershagen

49-33439-65637

Table of contents

VI.1 EXECUTIVE SUMMARY

The National Institute of Statistics (INE) started a new Household Budget Survey (IOF2008/9) in September 2008. The data collection in the field will last until August 2009. According to the sampling techniques applied to IOF 2008/9 a sample size of 10848 households has been selected in 1060 enumeration areas to be covered by the survey. Taking into account the urgent need for these data INE decided not to wait until the final data will be available but to process the data for the first 6 months of the survey, based on a nationally-representative subsample, and produce a preliminary report on the main results. The report is for internal use and will not be published. The whole procedure of data validation and aggregation will be tested on the basis of the data for the 6 months and the experience can be used to produce the final report.

A combined advisory mission was requested for assisting INE with these activities. This mission consisted of

1. Erwin Triebkorn, specialist in household surveys,
2. David Megill, specialist in sampling for household surveys,
3. Astrid Mathiassen, specialist in poverty analysis,
4. Ellen Cathrine Kiosterud, specialist in poverty analysis.

The specialist on household surveys started his activities on June 3. At this time the data entry for the 6 months had entered its final stage. However, the validation had shown that the data quality was insufficient. One reason was the weakness of staff working on data processing. In part this problem was solved by the assistance of Ellen M. Payongayong from Michigan State University, who is on a mission with the Ministry of Agriculture in Mozambique. She started around mid-June on a part-time basis to assist INE in the data processing. Given the poor quality of the IOF data files she suggested re-entering all the data for the 6 months. The data entry started again on July 4 using a new application of CSPro worked out by Ellen Payongayong. The new data for the first 6 months including all provinces were available on July 28 only.

A working group was created to work on validation. The household budget survey specialist (HBS specialist) prepared the instructions and procedures for the validation and gave continuous assistance to the group.

The main objective of the mission was to support INE making the preliminary report using the data gathered during the first six months of data collection. Because of delay in data processing it was not possible to finish the report. In close cooperation with Director Fatima Zacarias, a plane was made how to elaborate the report. In a meeting with all the staff engaged in this task the plane was discussed and the responsibilities determined. The HBS specialist should work on the report remotely.

The major achievement of the mission is a better organisation of the work. Now the staff knows better what to do and is more motivated. Provided that Ellen M. Payongayong will continue giving support to data processing the expected results will be achieved.

VI.2 INTRODUCTION

The National Institute of Statistics (INE) started a new Household Budget Survey (IOF2008/9) in September 2008. The data collection in the field is to be finished in August 2009. The interview period for each household lasted for one week. Each household was visited at least 3 times in order to fill out the questionnaires.

A total of 5 questionnaires have been used for data collection.

Questionário do Agregado Familiar which contains the following modules:
- General demographic information such as age, sex etc.
- Migration,
- Education,
- Health,
- Employment,
- Housing,
- Agriculture,
- Calamities,
- Poverty predictors,
- Monetary section,
- Weight and height of children

Questionário das Despesas Diárias with the following modules:
- Daily purchases,
- Daily consumption of own production,
- Daily gifts in kind,

Questionário das Despesas Anuais, Mensais e Receitas:
- Inventory of durable goods and purchases in last 12 months,
- Monthly purchases,
- Monthly transfers received and paid,
- Monthly income,
- Transfers from children and husband or wife of head of household
- Informal sector,

Questionário Comunitário including also market prices.

Questionário sobre Turismo.

Compared to the last Household Budget Survey (IAF 2002/3) the extent and range of information collected has been increased substantially.

A total of 10848 households have been selected in 1060 enumeration areas (AE) what will ensure representative results at the national and provincial level. In urban areas in each AE 12 households and in rural areas 9 households have been selected for the interview. The sample has been divided into four subsamples, in a way to ensure representative data on national level for every 3 months of the survey. Taking into account the urgent need for these data INE decided not to wait until the final data will be available but to process the data for the first 6

months of the survey, and produce a preliminary report on the main results. This procedure provides also an opportunity to finalize the processing, tabulation and analytical procedures in preparation for producing the final results from the 12 months of IOF data.

For the first 6 months of the survey 5424 households had been originally selected, out of which 4895 had been interviewed. That gives a response rate of 90, 4 %. A total of 520 non-interviewed households have been substituted. The main reason for substitution was the absence of the selected households.

A combined advisory mission was asked to support INE. The mission consisted of

1. Erwin Triebkorn, HBS specialist,
2. David Megill, specialist in sampling for household surveys,
3. Astrid Mathiassen, specialist in poverty analysis,
4. Ellen Cathrine Kiøsterud, specialist in poverty analysis.

The main objective of the mission was to support INE achieving the following results:

1 Completion of tabulation plan;

2 Completion of the consistency check program;

3 A clean database containing data from the first six months;

4 Tables generated on the basis of the first 6 months;

5 A draft preliminary report.

VI.3 ACTIVITIES DURING THE MISSION

At the beginning of the mission the data entry for the 6 months was in its final stage. A staff of 55 people worked in two shifts on data entry. Following a double-entry of the survey data, the independent data files from the two entries were compared and the inconsistencies after consulting the questionnaires were corrected. One could expect that this procedure should ensure a reasonable quality of data. However, the validation of the data files showed that the quality was very low. The main reason was the lack of experience of staff managing the data processing operations. Changes in the record structure during the process of data entry caused confusion and resulted in many errors. Ellen M. Payongayong from Michigan State University, who is at present on a mission in the Ministry of Agriculture in Mozambique, was asked to give technical support for processing the IOF 2008/9 data. After analysing the situation she suggested re-entering the data for the first 6 months. Finally, on June 22 the data entry was stopped. Ellen M. Payongayong made a new application for the data entry program. The data entry staff was reduced by almost half. The best people were chosen after a test. On July 4 the data entry started again. Actually, the new data for the first 6 months including all provinces were available on July 28.

Because of the limited capacity in programming the data entry program contains only a few checks of data consistency. A team of 9 INE staff members was created to work on the validation of various parts of the questionnaires. This team also includes specialists from other parts of INE, such as the Department of National Accounts and the Price Department. The HBS specialist defined the methods and procedures needed. The following documents containing the rules for validation and consistency checks have been prepared:

1. Rules for consistency checks in the demographic part of the Household Questionnaire
2. Rules for validating and processing expenditures
3. Rules for validating and processing income

A Portuguese version of the mentioned documents is attached as annexes 2, 3 and 4 to the report.

The validation of data on daily expenditures, consumption of own produced products and durable goods was done using the average prices from the observations recorded in the daily expenditure questionnaire

The monthly and annual expenditure observations have been checked in a similar way. Observations with out-of-range values are printed and compared to the data recorded in the questionnaire and corrected where necessary. This was combined with the consistency checks between the general household questionnaire and the income and expenditure questionnaires. This included consistency checks between rent expenditure and housing tenure; expenditure on water and source of water; salary income and employment status; education expenditure and children in school; expenditure on health and visits to the doctors. Corrections have been made when the data printed in the error lists did not correspond to the data recorded in the questionnaires.

The main reason of errors was the weak monitoring of the data entry process. To ensure better quality a strict control of the quantity and quality of data entered by each staff member has been introduced. In regular meetings with all data entry personnel the results are analyzed and the staff with bad quality of work is being dismissed.

By the end of the mission the first round of the validation of the first 6 months of the survey for all provinces had been finished, but the corrections in the data base had been made only for 3 provinces.

The delay in data processing did not allow finishing the validation in time. The HBS specialist and the data processing specialist agreed on the following activities which remained to be done:

Finalize the corrections of the first round,
Establish the price limits,
Produce error lists according to the validation rules,
Impute values for observations with incomplete data,
Aggregate the data to household level,
Produce tables according to the table program.

Because of the mentioned problems with data processing the sampling specialist was not able to produce in time the weights and the sampling errors. This will be done remotely when the data will be available (see also the mid-term report of David Megill).

A draft of the table program was prepared by Maria Alfeu on the basis of the tables from the IAF 2002/3. The HBS specialist studied the program and gave advice how to perfect it. Actually, it's a good basis for internal work but the tables are too broad to be published. The poverty specialists worked with Maria Alfeu on the table program. They will continue giving their support.

The main objective of the mission was to support INE making the preliminary report using the data gathered during the first six months of data collection. Because of delay in data processing it was not possible to finish the report. In close cooperation with the Director of the Directorate of Demographic, Vital and Social Statistics, Fátima Zacarias, a plane was made how to elaborate the report. In a meeting with all the staff engaged in this task the plane was discussed and the responsibilities determined. It has been agreed that the members of the mission should work on the report remotely.

VI.4 RECOMMENDATIONS

The experience proved that INE has developed its capacity to collect survey data efficiently. However, its ability to elaborate the questionnaires and Data Processing management should be improved.

The design of the tabulation plan and the data processing system should be an integral part in preparing the survey. In that way the data of the pilot survey could be used to check the data processing programs.

The programming and data processing management has shown that it needs to be considerably strengthened.

Although the mission could not realize all expectations it did contribute effectively to create the conditions that the IOF 2008/9 will succeed. The mission should continue supporting INE in producing the final results of the survey.

APPENDIX VI.1. Persons met

INE

Mr. Loreiro, João, Dias, President
Mr. Gaspar, Manuel da Costa, Vice-President
Mr. Balate, Arão, Director of surveys and censuses
Ms. Zacarias, Fátima, Directora of demographic and vital statistics
Mr. Muahio, Cristovão, Head of department
Mr. Guiliche, Firmino, Head of department
Mr. Megill, David, Consultant in sampling
Ms. Mathiassen, Astrid, Consultant in poverty analysis
Ms. Kiosterud, Ellen, Consultant in poverty analysis
Ms. Payongayong, Ellen, Consultant in data processing

Mr. Nucifora, Antonio, World Bank
Mr. Channing, Arndt, Ministry of Planning and Development

APPENDIX VI.2 Rules for checking consistency in the demographic part of the Household Questionnaire

ID07 O código da Área de Enumeração (AE) deve responder a lista das AE seleccionadas,
ID08 O código do Agregado Familiar (AF) deve responder a lista dos AF seleccionados
Os números de membros do AF devem ser consecutivos,
O chefe do AF deve ter o numero 1,
Não pode haver mais de 1 chefe do AF,
AF03 (relação de parentesco com o chefe) = 1 então AF05 (idade)> =12
AF03 = 03 (filho), então AF05 (idade) <= idade do chefe – 12,
AF03 = 04 (pais), então AF05 (idade)> = idade do chefe + 12.
A idade da mãe (AF18B, AF05) deve ser pelo menos 12 anos mais da dos filhos,
A idade do pai (AF19B, AF05) deve ser pelo menos 12 anos mais da dos filhos,
AF03 =2 (cônjuge), então AF11 = 2ou3 ou 4,
O chefe e cônjuge devem ter sexo diferente,
O chefe e o cônjuge devem ter AF11 (estado civil) igual,
AF11 (estado civil) <> 1 (solteiro), então AF05 (idade)> = 12,
AF05 (idade) <12, então AF11 fica branco,
AF05 (idade)> =18, então AF18 e AF19 ficam em branco,
AF05 (idade) <18, então AF18 e AF19 não podem ficar em branco,
Se no AF viver cônjuge, então para o chefe do AF o estado civil (AF11) =2,3ou 4,
Se no AF viver mais de um cônjuge, então o chefe e os cônjuges vivem em regime de poligamia (AF11=4),
Se AF05 (idade) <5, então AF20 (sabe ler e escrever) ate AF31 (pagou transporte) fica em branco),
Se AF05 (idade)> =5, então AF20 e AF21 (frequentou a escola) devem ter resposta,
Se AF21 (frequentou a escola) =2 (não), então AF21A ate AF31D ficam em branco,
Se AF21 (frequentou a escola) = 1 (sim) e AF05> =5 e <=24, então AF22 ate AF23B devem ter respostas,
Se AF24 (estuda actualmente) = 2 (não) então AF25A ate AF28 ficam em branco,
Se AF24 (estuda actualmente) = 1 (sim) então AF25A ate AF28 devem ter resposta,
Se AF30 (frequentou a escola) = 2 (não), então AF31A ate AF31D ficam em branco,
Se AF30 (frequentou a escola) = 1 (sim) então AF31A ate AF 31D devem ter resposta,
Se AF23 (nível mais alto frequentado) =01 ou 02 (primário), então AF05 (idade) mínimo 5 anos,
Se AF23 (nível mais alto frequentado) = 3 ou 4 (secundário), então AF05 (idade) mínimo 11
AF25 (nível que frequenta actualmente) deve ser > = AF23 (nível completado),
SE AF24 (actualmente estuda) = 1 (sim), então AF29 fica em branco,
Se AF24 (actualmente estuda) = 2 (não), então AF (porque não estuda) deve ter resposta,
AF33 (esteve doente) não pode ficar em branco.
Se AF33 = 2 (não), então AF34 ate AF38 ficam em branco.
Se AF33 = 1 (sim) então AF34 ate AF35 não podem ficar em branco.
Se AF35 (consultou?) = 2 (não), então AF36 (tipo de agente) e AF37 (problemas na consulta) ficam em branco
Se AF35 = 1 (sim), então AF38 fica em branco.
Se AF35 = 2 (não), então AF38 deve ter resposta.
Se AF39A (quanto pagou pela consulta no sector publico?)> 0, então AF39 (quantas vezes consultou?) deve ter resposta,

SE AF39C (sector privado de saúde)> 0, então AF39B deve ter resposta
Se AF39E (sector tradicional)> 0, então AF39D deve ter resposta,
Se AF05 (idade) <7, então AF40 (trabalhou nos últimos 7 dias) ate AF68 ficam em branco,
Se AF05 (idade)> = 7, então AF40 (trabalhou nos últimos 7 dias) deve ter resposta,
Se AF40 (trabalhou nos últimos 7 dias) = 1 (sim), então AF41 ate AF45A ficam em branco, mas AF46 (ocupação principal), AF47 (para quem trabalhou), AF48 (trabalhador permanente?) e AF 49 (actividade económica) devem ter resposta,
Se AF40 (trabalhou nos últimos 7 dias) = 2 (não), então AF41 (não trabalhou mas tem algum emprego) deve ter resposta,
Se AF42 (estava disponível) = 1 (sim), então AF43 (porque não estava disponível)) fica em branco,
Se AF42 (estava disponível) = 2 (não), então AF43 porque não) deve ter resposta,
Se AF43 (porque não estava disponível?) = 01 (estava a estudar), então AF24 (estuda actualmente) = 1,
Se AF43 = 07 (muito velho), então AF05 (idade)> 65,
O código da ocupação principal (AF46) deve constar na Classificação das Ocupações,
O código da actividade económica (AF49) deve constar na Classificação das Actividades Económicas,
Se AF53 (ultimo salário)> 0, então AF50 (trabalhador assalariado) = 1,
O código da outra ocupação (AF58) deve constar na Classificação das Ocupações,
O código da outra actividade económica (AF 61) deve constar na Classificação das Actividades Económicas,
Se AF56 (mais de uma actividade) = 2, então AF57 ate AF66 ficam em branco,
Se AF56 = 1, então AF57 ate AF61 devem ter resposta,
Se AF62 = 2, então AF 63 ate AF66 ficam em branco,
Se AF62 =1, então AF63 ate AF66 devem ter resposta,

APPENDIX VI.3 Rules for validating and processing expenditures

Validação do questionário de DESPESAS DIARIAS

O código do produto deve responder ao respectivo código no COICOP,
Os códigos da unidade local, da unidade padrão e do lugar de compra devem constar nas listas dos respectivos códigos.
O preço deve responder ao valor dividido pela quantidade em unidades padrão.
Produzir limites de preços para cada produto no questionário de DESPESAS DIARIAS.
Para os produtos que tem preços no questionário comunitário seguir as regras seguintes:
O preço mais baixo de cada produto serve como limite mínimo,
O preço mais alto de cada produto serve como limite máximo,
Para os produtos que não tem preço no questionário comunitário calcular preços com base nos dados sobre quantidade e valor no questionário das despesas diárias, conforme as regras seguintes:
Preço mínimo = preço médio dividido por 2
Preço máximo = preço médio multiplicado por 2
(Nota: Para alguns produtos podem-se ampliar os limites de preços)
Comparar o preço de cada produto com os limites de precos e listar as divergências.
Sugere-se de verificar, se devia-se aplicar a mediana para os preços porque ela está menos afectada pelos valores extremos.
Verificar as divergências com base no questionário e corrigir os ficheiros.
Quando a coluna 4 (quantidade)>0, mas a coluna 5 (valor) esta em branco aplica-se o preço medio por provincia, Urbano/Rural e periodo para o calculo do valor.

Verificaçao dos produtos comprados

Se VP3 = 1 (sim), então o produto deve constar nas despesas diarias pelo menos num dos 7 dias.
Se VP3 = 2, então VP4, VP5, VP6 e VP7 não podem ficar brancos.

A validação das posições do AUTOCONSUMO e das RECEITAS EM ESPECIE PELO TRABALHO segue as mesmas regras como as definidas para as despesas diárias.

Validação do questionário de BENS DURAVEIS, DESPESAS ANUAIS e MENSAIS

Os códigos de bens duráveis devem constar na COICOP.
Se a coluna 5 >0 (o que significa que o AF comprou este bem duravel) a coluna 6 não pode estar em branco.
Calcular o preço médio que o AF pagou, dividindo o valor pago (coluna 6) por numero de bens comprados (coluna 5),
Calcular o preço medio com base nos dados sobre quantidade e valor.
Comparar o preço de cada produto com a lista dos limites de preços para os bens duraveis e listar as divergências.
Verificar as divergências e corrigir os respectivos ficheiros.

Em casos quando a coluna 5 (quantidade)>0, mas a coluna 6 (valor) esta em branco aplica-se o preço medio para o cálculo do valor.

A validação das despesas mensais segue em principio as mesmas regras como a de bens duraveis.Os códigos dos bens e serviços devem constar no COICOP.
Os códigos das unidades de medida devem responder a lista destes códigos.
Se a coluna 4 não está em branco (o que significa que o AF comprou este bem), a coluna 5 deve >0
Calcular o preço médio para cada produto ou cerviço [valor(col. 5) /quantidade(col. 4)].
Calcular os preços com base dos dados sobre quantidade e valor no questionário das despesas diárias,conforme as regras seguintes:
Preço mínimo = preço médio dividido por 2
Preço máximo = preço médio multiplicado por2
Comparar os preços calculados com a lista dos limites de preços.
Verificar as divergencias no questionario e fazer as correcções .
Em casos quando a coluna 4 (quantidade)>0, mas a coluna 5 (valor) está em branco aplica-se o preço medio para o calculo do valor.

Validação das despesas pela habitação
Se o código 04.10.00 (rendas efectivas pela habitação> 0 então a pergunta 71 deve ser 2 (alugada),
Se o código 04.20.00 (rendas imputadas pela habitação)> 0 então a pergunta 71 deve ser 1 (própria),
Se o código 04.41.01 (consumo de agua canalizada)> 0 então a pergunta 78 (principal fonte de agua) deve ser 1 (agua canalizada dentro da casa) ou 2 (agua canalizada fora de casa).
Se o código 04.51.01 (consumo de electricidade)> 0 então a pergunta 84 (principal fonte de energia para cozinhar) e (ou) a pergunta 85 (principal fonte de energia para iluminação) devem ser 01 (electricidade),

Validação das despesas pela saúde
Se AF33 (doente durante as duas últimas semanas) = 2 (não), então AF34 ate 39E ficam brancos.
Se AF33 = 1 (estava doente), então ver se consultou o agente de saúde AF35.
Se AF35 = 1 (significa que a pessoa fez a consulta), então ver AF36 (tipo de agente: posto de saúde, clínica privada....)
Se AF36 = 01 (posto de saúde do estado), 02 (centro de saúde) ou 03 (hospital), então ver AF39 (quantas vezes consultou) e AF39A (quanto pagou pela consulta),
Se AF36 = 04 (clínica privada), 05 (farmácia) e 06 (medico privado), então ver AF39B (quantas vezes) e AF39C (quanto pagou).
Se AF36 = 07 (curandeiro), então ver AF39D (quantas vezes) e AF39E (quanto pago),

Validação das despesas pela educação
Se AF24 (estuda actualmente?) = 2, então Af25A (nível de educação) ate AF28 (problemas na escola) ficam brancos.
Se AF24 = 1 (a pessoa estuda), então verificar AF25A (nível), AF25B (classe), AF27 (a quem pertence a escola),
Se as perguntas AF31A (matricula e propinas), AF31B (livros escolares), AF31C (uniforme escolar) e AF31D (transporte escolar) não estiveram em branco, então pelo menos um membro do AF deve ter frequentado a escola.
Se AF30 (nos últimos 12 meses frequentou a escola) =1, então verificar FA31A (propinas), FA31B (livros escolares), FA31C (uniforme escolar) e AF31D (transporte)

Os códigos AF31a_1 (propinas), AF31C_3 (uniforme) e AF31D_4 (transporte) devem constar na divisão 10 no COICOP.

APPENDIX VI.4 Rules for validating and processing income

Validação de dados no questionário do Agregado Familiar (QAF)

Salário na actividade principal

Se P 50 (trabalhador assalariado) = 1, então P53 (Salário no último mes) não pode ficar em branco,
Se P53> 0, então P50 deve ser 1
Se P53> 0, então P 54 (salário diário, semanal ou mensal) não pode ficar em branco,
Para transferir dados da P53 (salário recebido no ultimo mes) em salário mensal segue as regras seguintes:

Se P54 = 1 (diário), então o valor da P53 deve-se multiplicar com o número de dias trabalhadas durante de mês.

Se P54 = 2 (semanal), então o valor da P53 deve-se multiplicar com o número de semanas trabalhadas durante de mes,

Se P54 = 3 (mensal), não precisa alterar o valor contido na P53,

Nota: No questionário falta a informação sobre o número de dias e semanas trabalhadas durante de mês. Assumindo, que nestes casos as receitas respondem as despesas e sabendo o salário recebido por dia ou por semana, pode-se estimar o numero de dias e semanas trabalhadas.

Salário na actividade secundaria

Se P 62 (trabalhador assalariado) = 1, então P65 (Salário no ultimo mes) não pode ficar em branco,
Se P65> 0, então P62 deve ser 1
Se P65> 0, então P 65A (salário diário, semanal ou mensal) não pode ficar em branco,
Para transferir dados da P65 (salário recebido no ultimo mes) em salário mensal segue as regras seguintes:
Se P65A = 1 (diário), então o valor da P65 deve-se multiplicar com o número de dias trabalhadas durante de mês.
Se P65A = 2 (semanal), então o valor da P65 deve-se multiplicar com o número de semanas trabalhadas durante de mes,
Se P65 = 3 (mensal), não precisa alterar o valor contido na P65,

Nota: No questionário falta a informação sobre o número de dias e semanas trabalhadas durante de mês. Assumindo, que nestes casos as receitas respondem as despesas e sabendo o salário recebido por dia ou por semana, pode-se estimar o numero de dias e semanas trabalhadas.

Receitas do trabalho assalariado além do salário

Se P55 (valor recebido)> 0, então P50 não pode ficar em branco,
Se P66 (valor recebido)> 0, então P62 não pode ficar em branco,

Rendimento da machamba própria

Os códigos das culturas praticadas devem responder a CNBS (Classificação Nacional de Bens e Serviços de Moçambique).
A unidade deve responder a lista das unidades de medida.
P103 (estado) não pode ser> 9
Se P100 (vendeu uma cultura?) = 1 (sim), então P101 (quantidade), P102 (unidade) e P103 (estado) não podem ficar em branco,
Se P101> 0, então uma das perguntas P104 (valor da venda) ou P105 (preço por unidade de venda) deve ter resposta,
Listar Área, AF, P100, P101, P102, P103, P104 e P105 por produto,
Se P104 estiver em branco, o valor de cada produto vendido calcula-se na base de P105, multiplicando a quantidade (P101) por preço de venda.
Se P110 (cria ou criou animais) = 1 (Sim), então P111 (quantos tem), P112 (quantos foram vendidos) e P 113 (valor de venda) não podem ficar em branco.
Listar Área, AF, P109, P110, P111, P 112 e P 113,
Se P117 (vendeu produtos) = 1, então P118 (Valor de venda) não pode ficar em branco.
Se P118> 0, então P117 deve ser 1,
Listar AE, AF, P116, P117, P118,

Validação de insumos e de outros custos
Se P108> 0, então P107 deve ser 1,
Se P114 = sim, então P115> 0

Calcular o rendimento da produção agrícola
Listar AE, AF, P104, P113, P118, P108, P115,

A. Somar o valor da venda P104 (inclusive os valores calculados com base de P105) + P113 + P118,

B. Somar os custos P108 + P115,

Calcular as receitas por AF: A (valor de venda) – B (custo de produção),

Validação do rendimento de pequenos negócios
Se P1 = 2, então todo o resto do ficheiro fica em branco.
Se P1 = 1, então P1A ate P11 devem ter resposta,
O código da P1A deve constar na CAE,
Se P8 não estiver em branco, então P9 deve ter resposta,
Se P9> 0, então P8 não pode ficar em branco,
Calcular o lucro que vem dos pequenos negócios:
P7 (rendimento) – [P9 (salário pago) + P10 (despesa em material) + P11 (outras despesas)]
Listar AE, AF, P7, P9, P10, P11 para os AF que tem lucro negativo.

Validação de Receitas do ultimo mes
O Número de Ordem da Pessoa nesta parte de Questionário das Despesas Anuais, Mensais e Receitas deve responder ao Numero de Ordem no Questionário do Agregado Familiar,
Se RUM 1.2 (Receita de arrendamento de terras agrícolas)> 0, então AF91 = 1 (Renda de machamba a outros).

Se RUM 1.3 (Receita de arrendamento de carro)> 0, então BD3 (Quantos possui) pelo menos um dos códigos 07.11.10 (carro comprado novo) ou 07.11.20 (Carro comprado a segunda mão) deve ter resposta.

April 14, 2009

TERMS OF REFERENCE

for three short-term missions on a

The Household Budget Survey 2008/9

25 May – 24 July, 2009

within the

AGREEMENT ON CONSULTING IN INSTITUTIONAL CAPACITY BUILDING, ECONOMIC STATISTICS AND RELATED AREAS
between INE and Scanstat.

Consultants: 1._____, 2. _____, 3. _____

Counterparts: The consultants will further work in close coordination with the staff of the Family Budget Survey (IOF 2008/09) at INE, headed by INE Vice President Manuel Gaspar, Director Arão Balate and Director Fatima Zacarias.

Background
In the Strategic Plan 2008 - 2012 one of the priority activities is the execution of the Household Budget Survey 2008/09, which has as some of its main objectives to gather data making it possible to evaluate the Action Plan of Reduction of Absolute Poverty (PARPA II) and monitor the fulfillment of the Millennium Goals.

The last HBS survey was conducted in 2002/03 and when analyzed a reduction of poverty by 14% compared with the HBS 1997/98 was found.

Main reasons for the Missions
The National Institute of Statistics has a lot of experience in conducting household surveys, but there are still weaknesses in database cleaning and in analysis of the information, taking into account the multi sectoral data that this survey include. For this reason a combined advisory mission to support INE is necessary. The following expertise is needed:

6. **A specialist in household surveys**, especially in the area of household budget and expenditure;

7. **A specialist in sampling for household surveys**, with experience in household budget surveys;

8. **A specialist in poverty analysis**, using income and expenditure information from the households.

Beneficiaries of the mission
The mission will benefit the National Statistical System in general and in particular the National Statistical Institute which will get further technically trained by this mission.

Objectives
The overall objective of the mission is to support INE making the preliminary HBS report using data gathered during the first six months of data collection in the field. The three experts are supposed to contribute in the following way:

1. The specialist in household surveys, especially in the area of household budget and expenditure;
    h. Support INE in the conclusion of the consistency check program;
    i. Support the INE in cleaning the database;
    j. Support INE in the conclusion of the tabulation plan;
    k. Support INE in preparing the draft preliminary report;
    l. Support the INE in the documentation of all procedures described above.
    m. Support the INE with the consultants experience in this type of surveys and to perform other related activities;
    n. Duration of the mission; 9 weeks (May 25 to July 24, 2009).

2. The specialist in sampling for household surveys with experience in household budget surveys;
    a. Support INE in calculating the weights for all variables of the survey
    b. Support INE in expanding the results to different levels in all aspects;
    c. Support INE in the documentation of the procedures described previously;
    d. Support INE with the consultants experience in this type of surveys and to perform other related activities;
    e. Duration of the mission; 7 weeks (June 08 to July 24, 2009)

3. The specialist in poverty analysis, using income and expenditure information from the households;
    a. Support INE in the desegregation of the results in various quintiles and levels of poverty;
    b. Support INE in using the information to assess the levels of poverty;
    c. Support the INE in the documentation of the procedures described previously;
    d. Support INE with the consultants experience in this type of surveys and to perform other related activities;
    e. Duration of the mission; 7 weeks (June 08 to July 24, 2009)

Expected results
1. Completion of tabulation plan;
2. Completion of the consistency check program;
3. A clean database containing the first six months;
4. Tables generated on the basis of the first 6 months;
5. A draft preliminary report.

Activities
The detailed agenda of the missions will be specified in the first day together with the counterparts.

Tasks to be done by INE to facilitate the mission

1. Elaborate ToR for the mission
2. Prepare and supply the consultants with necessary documents and information, such as mission reports, strategies, plans etc.
3. Supply good working conditions for the consultants.

Source of Funding
Project: MPD – 2008 – 0006 – Inquérito Sobre Orçamento Familiar – IOF

PAAO09 – 1.4.6 Inquérito ao Orçamento Familiar - IOF 2008/2009

Timing of the mission
See above

Place
The premises of the National Institute of Statistics in Maputo

Language
Portuguese

Report
The consultants will prepare a draft report in Portuguese with findings and recommendations which must be agreed with the counterparts of the INE. The final draft must be submitted for final comments by INE, within a week after the completion of its mission. The structure of the report will be agreed with the counterparts and most of the content will be the recommendations.

Statistics Denmark as Scanstat Lead will publish the final version on www.dst.dk/mozambique within 3+ weeks of the end of the mission.

These Terms of Reference were prepared by Arão Balate INE/DCI

Day / / .....

Approved by Luis Mungamba, Contract Manager for the INE – Scanstat Contract

Day / / .....