



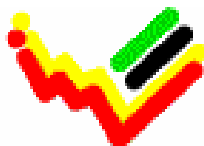
Report from a short-term mission

In support of the 2002-3 IAF survey

17 February to 31 March, 2003

**TA for the 'Bridging Support Program to Strengthen
the Institutional Capacity of the National Statistics, Mozambique**

James Otto



Instituto Nacional de Estatística

This report contains
restricted information
and is for official use only.

Ref. no. 104.Mozambique.1-5

May 18, 2003

Report on a mission to Mozambique: 17 February to 31 March, 2003

In support of the 2002-3 IAF survey

Prepared by James Otto

May 18, 2003

Table of contents

1. Executive Summary	Page 2
2. Introduction	Page 2
3. Activities during the mission	Page 3
4. Findings and recommendations	Page 5
Annex 1: IAF2002 Data and Database summary	Page 6
Annex B: IAF2002 Imputation Summary	Page 21

1. Executive Summary

The system to process the data from the 2002-3 IAF survey was extended to include the integrated validation of all questionnaires, imputation of values for observations with incomplete data, aggregation of data to household and higher levels, preparation of all standard household tables, preparation of provisional income and expenditure tables and the calculation of sampling errors for all principal estimates. The system was tested using data from the first six months of this survey. Results were prepared for the first six months of the survey.

2. Introduction

The 2002-3 integrated household survey (IAF) is a 12 month survey from July, 2002 to June, 2003 using a sample covering 8400 households. The survey is designed to produce estimates of household expenditure, household income and social characteristics of the household at the national, provincial and rural/urban level.

The survey uses four questionnaires

- (1) The general household characteristics questionnaire (Carateristicas Gerais do Agregado Familiar - QCG), based on the questionnaire used for the 2000-1 QUIBB survey records:

Household information

- Housing characteristics and amenities
- Poverty predictors
- Victimization

Household member information

- Demographic characteristics
- Education
- Health
- Employment
- Victimization

- (2) The daily household expenses questionnaire (Despesas Diarias - QDD) records daily observations of

- Household purchases
- Household consumption of own production
- Gifts in kind received by the household

- (3) The monthly household expenses and income questionnaire records (Despesas e Receitas - QDR) records:

- Inventory of durable goods owned by the household
- Number and value of durable goods purchased in the last 12 months
- Education expenditure in the last 12 months
- Household purchases in the last 30 days
- Income and revenue in the last 30 days by household member
- Transfers received and paid in the last 30 days by household member

- (4) In rural areas the community questionnaire (Questionario Comunitario - QC) records:

- Market prices for selected products
- General characteristics of the community

Development of the data processing system for the survey began in April 2002. The system used for the 2000-1 QUIBB survey was adapted to process the general household questionnaire (QCG). This system uses

TELEform to capture data from questionnaires using optical scanning technology. The captured data are transferred to a Microsoft Access database and all subsequent processing is done with Access. The data entry system for the expenditure and income questionnaires was designed and implemented in July and August 2002. The data entry system was developed using Microsoft Access to allow easy integration with the data from the household questionnaire. The data entry system uses very limited data validation procedures: it uses double entry of the questionnaire data to insure the accuracy and completeness of the entered data. Development of complete integrated validation, imputation, aggregation and tabulation modules was deferred until a sufficient amount of data was available to test these procedures. By February, 2003 data from the first 13 periods of the fieldwork, covering the six months from July to December, 2002, had been entered and were available for use to develop the rest of the data processing system for the survey. The team of consultants David Megill, survey sampling specialist, Erwin Triebkorn, survey statistician and James Otto, survey data processing specialist worked with INE staff starting in mid-February, 2003 to define, implement and test all the methods and procedures needed to completely process the data from the first six months of the survey. The main objective of this work was to insure that results for the complete survey would be available by the end of September, 2003.

3. Activities during the mission

a. Terms of reference

Design, develop, test, implement and document the following components of the IAF2002 survey data processing system:

- (1) Integrated validation of household and expenditure/income questionnaires.
- (2) Validation of daily expenses data, including purchase price by product.
- (3) Validation of monthly expenses data, including value of purchase by product.
- (4) Imputation of missing values for observations with missing data.
- (5) Aggregation of expenditure and income data to the household and higher levels.
- (6) Tabulation of data from all questionnaires.

b. Outline of work accomplished

The data from the household questionnaire were validated and errors were corrected immediately after scanning, so the household data was ready for final validation and analysis. The data summary and tabulation programs used for the 2000-1 QUIBB survey were modified for changes made to the questionnaire and were tested using the available data. The final validation revealed a few errors that needed to be corrected. Once these were corrected, unweighted summaries of all questions in the questionnaire were prepared.

David Megill, the survey sampling specialist, produced the household weights needed to produce weighted data summaries and tables, after several iterations of provisional estimates and adjustments. The details of the adjustments are described in his report of February, 2003. The weights were transferred to the IAF database and used to prepare weighted data summaries and tables. Tables were prepared for the entire household questionnaire except the victimization module which had been added to the questionnaire for this survey. David calculated sampling errors for several important estimates using the U.S. Census Bureau standard program, CenVar. These were compared to the sampling errors produced using the CWIQ system and were found to be the same.

We decided to base the principal validation of the purchases, own consumption and gifts received in kind in the daily questionnaire (QDD) on the unit price of the observation. This is calculated when both the quantity and value are recorded. The calculated unit price is compared to the unit price range defined for the product, period of the survey, place (rural or urban) and province. The unit price minimum, maximum and mean were defined using data from the observed daily purchases and the community questionnaire. The mean was included for use in the valuation of own consumption and gifts in kind observations that had quantity, but no value. The method used to define the range is described in detail in Annex A – IAF2002 data and database summary (Page 8).

When the unit price range information was available for a period, the observations in the daily questionnaires for the period were validated. A list of all observations with out of range unit prices was given to the survey specialists and data editors for comparison to the data recorded in the questionnaires. Corrections were written on the validation error list and then corrected in the database. All 13 periods were validated and corrected.

Similar methods were used to define the validation rules for the observations recorded in the monthly expenditure and income questionnaire (QDR). For some products only the value is recorded in the questionnaire so ranges of unit price and total value were defined. These are classified by product and area (rural or urban) for the first six months of the survey from the observed unit prices/value of the monthly purchases, annual purchases of durable goods and annual education expenses.

The unit price/value of observations in the monthly questionnaire were validated in a second validation program. This program included validation of all codes recorded in both daily and monthly questionnaires and consistency checks between all the questionnaires including the household questionnaire.

The second validation was run after the first validation had been completed. This was also done by period and used the same methods to record and enter corrections. The policy for correction of consistency errors was to verify that the information in the database agreed with the data in the questionnaire and to accept the inconsistency as representing reality. Examples are education expenses with no household members in school and no education expenses when there are members in school.

While the validation and correction were proceeding we started to define the procedures for aggregating the income and expenditure data to the household level across questionnaires. We decided to aggregate the income and expenditure data by source (questionnaire and table) to product level for the household. This is necessary because data for some products is recorded in more than one place. The aggregated data are combined to define total daily expenditure for the product by the household. At the same time the individual product totals are aggregated to the 15 income and expenditure categories defined by the first two digits of the product code. The aggregation procedure involves many separate, but fairly straightforward steps. These are described in detail in Appendix A – IAF2002 data and database summary (Pages14).

After the aggregation procedure had been defined we turned to the problem of observations with missing data. There were a certain number of own consumption and gifts received in kind with missing value and/or quantity. For observations with quantity and no value we need to impute a price in order to estimate the value of the observation. We decided to use the same price information that was used for validation of unit prices. The objective was to find a price for the product that was as close as possible in time and space to the observation with missing information. We established a search hierarchy that starts as close as possible to the observation and moves upward to higher (farther) levels to find a price for the product. The source of the price used for imputation is recorded with the observation. The method is described in detail in Appendix A – IAF2002 data and database summary (Page 13)

Further imputation was required for observations with neither quantity nor value. These involved only a few products - firewood, papaya and meals prepared for the household. For these products we used the median value of all observations in the survey that recorded value only (no quantity).

The imputation logic and a summary of the number of cases where imputation was required is in Appendix B – IAF2002 Imputation summary.

The tabulation plan for the income and expenditure data was based on the tables published for the 1996-7 IAF survey. These tables show (1) average daily household expenditure, (2) percentage distribution of average daily expenditure, (3) average household per capita expenditure, and (4) average household expenditure per adult equivalent, all at the two digit product level and aggregated to national, urban/rural, and provincial levels. Additional classifications are expenditure quintile and characteristics of the head of household. Estimates of sampling errors for these estimates were also prepared.

4. Findings and recommendations

a. Next steps.

The basic system for processing the data from the 2002-3 IAF survey has been designed, implemented, tested and used to produce results from the first six months of the survey. The basic procedures are well defined but there are some procedures that need to be evaluated and perhaps modified before producing results for the entire survey. Some obvious examples are rent imputation and the treatment of durable goods. The poverty monitoring unit in Ministry of Finance and Planning together with IFPRI is investigating these and other issues.

The entire system will be reviewed with a data processing team from INE at the advanced CWIQ workshop at EASTC in Dar-es-Salaam in May, 2003. Programs to prepare tabulations from the victimization module of the household questionnaire will also be prepared during this workshop.

The final review of the system will be in September, 2003 after all data have been entered and validated. The final tabulation plan for the income and expenditure data should be defined at this time.

Annex 1: IAF2002 Data and Database summary

Questionnaires

Características gerais do agregado familiar (QUIBB)

Household information

- Housing characteristics and amenities
- Poverty predictors
- Victimization

Household member information

- General characteristics
- Education
- Health
- Employment
- Victimization

Questionario des Despesas Diarias (QDD)

Tabelas

- DD – daily purchases
- AC – daily consumption of own production
- RE – daily gifts in kind

Questionario des Despesas e Receitas

Tabelas

- BD – Inventory of durable goods and purchases in last 12 months
- DE – Education expenditure in last 12 months
- DM – Monthly purchases
- RM – Monthly income and revenue
- TM – Monthly transfers received and paid

Questionario comunitario

Sections

- Market prices

Databases

QUIBB2 – QUIBB (household questionnaire)

IAF2002 – IAF (income and expenditure data)

IAF2002p – Aggregated data, programs

QCP – Prices from Community questionnaire

<u>Table</u>	<u>Contents</u>	<u>Database</u>
Cluster master	Master list of enumeration areas	QUIBB2
HHdata	Household level data from QUIBB	QUIBB2
INdata	Individual data from QUIBB	QUIBB2
HhIE	Household data from QDD/QDR	IAF2002
DD	QDD – DD data	IAF2002
AC	QDD – AC data	IAF2002
RE	QDD – RE data	IAF2002
BD	QDR – BD data	IAF2002
DE	QDR – DE data	IAF2002
DM	QDR – DM data	IAF2002
RM	QDR – RM data	IAF2002
TM	QDR – TM data	IAF2002
AggrProdHH	IE data aggregated for the household and product	IAF2002p
HHaggr	IE data daily total aggregated to two digit product level (15 groups: 12 expenditure, income, transfers received, transfers paid)	IAF2002p
AggrProdHH	IE data aggregated by product (6 digits)	IAF2002p
Productos	Master product list (based on COICOP) Includes master unit for product.	IAF2002
LugarCodigos	Codes for place where purchase occurred	IAF2002
UnidCodigos	Codes for local units	IAF2002

PrecosPUP	Unit prices by province, urban/rural, survey period and product derived from DD observations and the community questionnaire	IAF2002p
PrecosPU	Unit prices by province, urban/rural, and product derived from DD observations and the community questionnaire	IAF2002p
PrecosU	Unit prices by urban/rural, and product derived from derived from DD observations and the community questionnaire	IAF2002p
DMrange	Unit price range by product, urban/rural derived from BD and DM observations	IAF2002p
DMVrange	Value ranges by product, urban/rural derived from DE and DM observations with only value	IAF2002p

Main data processing functions

Build unit price range for products by province, urban/rural and period of the survey using community questionnaire prices and prices from daily purchases.

Validate prices recorded in the daily expenses questionnaire (DD, AC, RE)

Validate consistency between household and income/expenditure questionnaires.

Correct validation errors.

Impute values for missing data:

AC, RE observations with quantity but no price or value

AC, RE observations with neither quantity nor price (principally lenha)

Households without actual or imputed rent.

Durable goods with quantity but no value.

Summarize the results of imputation.

Aggregate income and expenditure data for the household.

Variable construction.

Prepare data summaries, tabulations and sampling errors.

Basic program specifications

mBuildPrecos

Define unit price range for products recorded in the daily expenses questionnaire (QDD)

Input

CMper table: list of AE's for selected period

DD table:

Query – PrecosPUPdd1: calculate unit price where quantity/value defined

Query – PrecosPUPdd2: summarize unit price by Prov-UrbRur-Period

QCPrecos table in QCP.mdb:

Query – PrecosPUPqc: summarize QC prices by Prov-UrbRur-Period

PRlimites table: broad unit price limits by product

Output

PrecosPUP table: unit price information by product summarized by Prov-UrbRur-Period

Processing

A table entry is created for each product with unit price information recorded in the DD (daily purchases) or QCprecos (community questionnaire prices) tables.

Unit price range is defined for each product:

minimum is the smaller of

QC price minimum

QC average price divided by 2

DD average price divided by 2

maximum is the larger of

QC price maximum

QC average price multiplied by 2

DD average price multiplied by 2

Broad price limits defined by product are added to each entry.

mQDDvalidate

Validate unit prices recorded in the daily expenses questionnaire (QDD) tables (DD, AC, RD) against the ranges defined in the PrecosPUP table.

Input

ClusterMaster table – master list of enumeration areas

HhIE table – household income and expenditure data

DD table – daily purchases

AC table – daily consumption of own produce

RE table – daily gifts/receipts in kind

PrecosPUP table – unit price ranges classified by product with Prov-Urb/Rur-Period

Output

List table – validation error listing to be printed with report QuestListR.

Processing

When quantity and value are defined for an observation, the unit price is calculated and checked against the unit price range defined for the product classified by province, urban/rural and survey period. Out of range values are printed for possible correction after consulting the questionnaire.

mQIEvalidate

Validate data recorded in the QDD tables (DD, AC, RD) and the QDR tables (BD, DE, DM, RM, TM); check consistency between QDD/QDR tables and the household questionnaire.

Input

ClusterMaster table – master list of enumeration areas

HhData table – general household data

InData table – individual household member data

HhIE table – household income and expenditure data

DD table – daily purchases

AC table – daily consumption of own produce

RE table – daily gifts/receipts in kind

DM table – monthly purchases

BD table – durable goods

DE table – annual expenditure on education

RM table – monthly receipts

TM table – monthly transfers received and paid

DMrange table – Unit price range from DM and BD observations

DMVrange table – Value range from DE and DM observations (without quantity)

Produtos – master product list

CodigoUnid – master list of local units

CodigoLugar – master list of place of purchase

Output

List table – validation error listing to be printed with report QuestListR.

Processing

DD observations:

Product code must be defined in the master product table (Produtos).
Value must be > 0.

For observations from pre-printed lines in the questionnaire
quantity must be > 0
local unit code must be defined in the local unit table (CodigoUnid)
place of purchase code must be defined in the place table (CodigoLugar)

For value only lines in the questionnaire
local unit code must be blank
place of purchase code must be blank

AC/RE observations:

Product code must be defined in the master product table (Produtos).
Local unit code must be defined in the local unit table (CodigoUnid).

DM observations:

Product code must be defined in the master product table (Produtos).
Value must be > 0.

When quantity and value are defined, the unit price is calculated and checked against the range defined for the product classified by urban/rural in the DMrange table. If only value is defined it is checked against the range defined for the product classified by urban/rural in the DMvrange tables.

For certain products, the place code must be defined in the place table (CodigoLugar).

BD observations:

Product code must be defined in the master product table (Produtos).

When quantity and value are defined for purchases in the last 12 months, the unit price is calculated and checked against the range defined for the product classified by urban/rural in the DMrange table.

DE observations:

Product code must be defined in the master product table (Produtos).

The value is checked against the range defined for the product classified by urban/rural in the DMvrange tables.

RM observations:

VL33 (net revenue) must = VL31 (gross revenue) – VL32 (costs)

Consistency checks:

Rent paid (DM 041000) and imputed rent (DM 042000) can both be present.

If rent paid is present then F1 (housing tenure) must be 2

If F1 (housing tenure) = 2 then rent paid must be present

If the household pays for piped water (DM 044100) then F9 (source of water) must be 1

If the household pays for electricity (DM 045100)

F11 (fuel used for cooking) must be 5

F12 (fuel used for lighting) must be 3

RM/TM observations:

The member number must exist in the household.

If the member has any income recorded in RM1-RM4 then the E7 for the member must be 1, 2 or 4.

If there are any education expenses in the last 12 months (DE observations) there should be household members in school (C5 = 1).

If there are any household members in school (C5 = 1) there should be education expenses in the last 12 months (DE observations).

All households should have some daily purchases (DD) or own consumption (AC) observations for the seven day period.

mQIEprint

Print data from all the questionnaires for a household

Input

ClusterMaster table – master list of enumeration areas

HhData table – general household data

InData table – individual household member data

HhIE table – household income and expenditure data

DD table – daily purchases

AC table – daily consumption of own produce

RE table – daily gifts/receipts in kind

DM table – monthly purchases

BD table – durable goods

DE table – annual expenditure on education

RM table – monthly receipts

TM table – monthly transfers received and paid

Output

List table – validation error listing

mImputeAC

Impute values for AC, RE observations with missing data; identify products within households with both daily and monthly purchases.

Input

HhIE table – household income and expenditure data

AC table – daily consumption of own produce

RE table – daily gifts/receipts in kind

PrecosPUP table – unit price ranges classified by product and Prov-Urb/Rur-Period

PrecosPU table – unit price ranges classified by product and Prov-Urb/Rur

PrecosU table – unit price ranges classified by product and Urb/Rur

DD table – daily purchases

DM table – monthly purchases

Output

DDMind table – products with both daily and monthly purchases with the household

Processing

Imputation

To be imputed:

Value for AC, RE observations with quantity and no value

Value for AC, RE lenha with no quantity and no value

Rent for households with no reported rent

Imputation methods:

For AC, RE observations with quantity and no value, use

- (1) QC average price for province, urban/rural, period, product
- (2) DD average price for province, urban/rural, period, product
- (3) QC average price for province, urban/rural, product
- (4) DD average price for province, urban/rural, product
- (5) QC average price for urban/rural, product
- (6) DD average price for urban/rural, product
- (7) QC average price for urban/rural, product (opposite of urban/rural)
- (8) DD average price for urban/rural, product (opposite of urban/rural)

Adjustment for DD/DM product duplication

Identify DD observations for products that also appear in the DM which were recorded before the DM questionnaire was completed. Exclude these observations from the household product aggregation.

mAggrProdHh

Aggregate income and expenditure data by source (table) to the product level for the household; calculate total daily expenditure and income for the household by product; aggregate total daily expenditure to the two digit classification (15 categories) by household.

Input

HhIE table – household income and expenditure data

DD table – Query: AggrProdDD

AC table – Query: AggrProdAC

RE table – Query: AggrProdRE

DM table – Query: AggrProdDM

BD table – Query: AggrProdBD

DE table – Query: AggrProdDE

RM table – monthly receipts

TM table – monthly transfers received and paid

Input/Output

RMt – Monthly receipts transposed by six digit code

TRt – Monthly transfers received transposed by six digit code

TPt – Monthly transfers paid transposed by six digit code

Six digit codes for receipts and transferred have been defined and added to the Produtos table. The list of codes is in Appendix A.

Output

AggrProdHh – household income and expenditure aggregated by product

HHaggr – household income and expenditure aggregated to two digits

Processing

1. Generate product level classifications of monthly receipts (RM) and transfers (TM) for the household.
2. Impute values for products with indefinite observations:

AC: Lenha (045401) value = 3000

AC: Papaia (011624) value = 2000

RE: Meals prepared for household (111300) value = 3000

Set the ImputeInd to 9 for these observations

3. Aggregate all observations to the product level for the household across tables

DD: aggregate quantity, value, number of days

AC: aggregate quantity, value

RE: aggregate quantity, value

DM: aggregate quantity, value

BD: aggregate quantity, value

DE: aggregate value

RT: aggregate number of persons with data, value

4. Generate entries in AggrProdHh for

Value of own production consumed as income

Imputed rent value for households with no rent observations

5. Calculate total daily value of consumption and expenditure by product, table and household

DD: $\text{total value} / 7$, where the total number of days is less than or equal to 7
 $\text{total value} / \text{total number of days}$, where the total number of days is > 7

AC: $\text{total value} / 7$

RE: $\text{total value} / 7$

DM: $\text{total value} / 30.4$

BD: $\text{total value} / 365$

DE: $\text{total value} / 365$

RT: $\text{total value} / 30.4$

6. Calculate product daily total for the household:

(1) Products with DD and DM observations

Total number of days = $30 + 3$, if DD number of days is less than or equal to 3

Total number of days = $30 + \text{DD number of day}$, if DD number of days > 3

Daily total value is $(\text{DD value} + \text{DM value}) / \text{total number of days}$

(2) All other cases

Daily total value is the sum of the individual daily values: DD, AC, RE, DM, BD, DE, RT

7. Duplicate receipts in kind to be included in expenditure as well as income

Alimentacao	(51.1.0.3.1)	as	(21.1.9.4.2)
Alojamento	(51.1.0.3.2)	as	(21.4.2.0.2)
Transporte	(51.1.0.3.3)	as	(27.3.6.0.0)
Outros	(51.1.0.3.4)	as	(32.7.0.0.0)

7. Duplicate transfers received in kind to be included in income

Received from organizations	(62.1.0.7.2)	as	(51.9.9.9.2)
Received from family inside Mozambique	(62.1.0.8.2)	as	(51.9.9.9.3)
Received from family outside Mozambique	(62.1.0.9.2)	as	(51.9.9.9.4)

8. Aggregate household expenditure and income aggregated to 2 digit COICOP ++ classification.

mConstVar

QUIBB2 variable construction program

Define household level aggregated variables:

Total household consumption (expenditure)

Per capita household expenditure

Expenditure quintile (weighted by population)

Number of adult equivalents in household

Age in years	Male adult equivalent	Female adult equivalent
< 1	0.27	0.27
1 – 3	0.45	0.45
4 – 6	0.61	0.61
7 – 9	0.73	0.73
10 – 12	0.86	0.78
13 – 15	0.96	0.83
16 – 19	1.02	0.77
>= 20	1.00	0.73

Source: FAO Adult equivalent scales.

Tabulation

Basic summary tables showing average household expenditure at the two digit level by main household characteristics: Urban/Rural, Zone, Province, Expenditure quintile, characteristics of the head of household.

Basic summary tables showing percentage of total household expenditure at the two digit level by main household characteristics: Urban/Rural, Zone, Province, Expenditure quintile, characteristics of the head of household.

Basic summary tables showing average household expenditure per capita at the two digit level by main household characteristics: Urban/Rural, Zone, Province, Expenditure quintile, characteristics of the head of household.

Basic summary tables showing average household income per adult equivalent at the two digit level by main household characteristics: Urban/Rural, Zone, Province, Expenditure quintile, characteristics of the head of household.

Appendix A: Product codes for monthly revenue and transfers

Codigo	Designação	UnidadeUP	CodigoUP	Divisão	Grupo	Classe	SubClasse	Categoria
51.1.0.0.0	Receitas			51	511	5110	51100	511000
51.1.0.1.1	Trabalho principal			51	511	5110	51101	511011
51.1.0.2.1	Trabalho secundario			51	511	5110	51102	511021
51.1.0.3.0	Receitas en especie do trabalho			51	511	5110	51103	511030
51.1.0.3.1	Receitas en especie do trabalho - alimentacao			51	511	5110	51103	511031
51.1.0.3.2	Receitas en especie do trabalho - alojamento			51	511	5110	51103	511032
51.1.0.3.3	Receitas en especie do trabalho - transporte			51	511	5110	51103	511033
51.1.0.3.4	Receitas en especie do trabalho - outros			51	511	5110	51103	511034
51.1.0.4.1	Gratificacoes ou remuneracoes extraordinarias			51	511	5110	51104	511041
51.2.0.0.0	Valor das vendas de produtos produzido pelo agregado familiar			51	512	5120	51200	512000
51.2.0.1.1	Valor das vendas de produtos agricolas			51	512	5120	51201	512011
51.2.0.2.1	Valor das vendas de produtos pecuarios			51	512	5120	51202	512021
51.2.0.3.1	Valor das vendas de peixe, camarao e outro pescado			51	512	5120	51203	512031
51.2.0.4.1	Valor das vendas de alimentos, bebidas processadas			51	512	5120	51204	512041
51.2.0.5.1	Valor das vendas de vestuario			51	512	5120	51205	512051
51.2.0.6.1	Valor das vendas de lenha e carvao			51	512	5120	51206	512061
51.2.0.7.1	Valor das vendas de produtos de artesanato			51	512	5120	51207	512071
51.2.0.8.1	Valor das vendas de material de construcao			51	512	5120	51208	512081
51.2.0.9.1	Valor das vendas de mel e produtos de caca			51	512	5120	51209	512091
51.2.0.9.9	Valor das vendas de outros produtos			51	512	5120	51209	512099
51.3.0.0.0	Valor obteve de negocios			51	513	5130	51300	513000
51.3.0.1.1	Valor obteve da actividade			51	513	5130	51301	513011
51.3.0.2.1	Valor gastou para realizar da actividade			51	513	5130	51302	513021
51.3.0.3.1	Valor liquido			51	513	5130	51303	513031
51.4.0.0.0	Valor que obteve no arrendamento			51	514	5140	51400	514000
51.4.0.1.1	Arrendamento da casa			51	514	5140	51401	514011
51.4.0.2.1	Arrendamento de terras agricolas			51	514	5140	51402	514021
51.4.0.3.1	Alugurer de carro			51	514	5140	51403	514031
51.5.0.0.0	Receitas extraordinarios			51	515	5150	51500	515000
51.5.0.1.1	Jogos de sorte (lotaria, totabola, rifa, etc)			51	515	5150	51501	515011
51.5.0.2.1	Heranca			51	515	5150	51502	515021
51.5.0.3.1	Outras receitas ocasionais			51	515	5150	51503	515031

Codigo	Designação	UnidadeUP	CodigoUP	Divisão	Grupo	Classe	SubClasse	Categoria
61.1.0.0.0	Transferencias pagas			61	611	6110	61100	611000
61.1.0.1.1	Pagou pela pensao de divorcio			61	611	6110	61101	611011
61.1.0.2.1	Pagou pela pensao de alimentacao			61	611	6110	61102	611021
61.1.0.3.1	Pagou de juros pelo emprestimo			61	611	6110	61103	611031
61.1.0.4.1	Transferiu para o estrangeiro			61	611	6110	61104	611041
61.1.0.5.1	Pagou para clubes, partidos, associacoes			61	611	6110	61105	611051
61.1.0.6.1	Contribui para instituicoes sem fins lucrativos e religiosas			61	611	6110	61106	611061
61.1.0.7.1	Pagou de xitique			61	611	6110	61107	611071
61.1.0.8.1	Outra transferencia paga			61	611	6110	61108	611081
62.1.0.0.0	Transferencias recibidas			62	621	6210	62100	621000
62.1.0.1.1	Recebeu pela pensao de reforma			62	621	6210	62101	621011
62.1.0.2.1	Recebeu pela pensao de divorcio			62	621	6210	62102	621021
62.1.0.3.1	Recebeu pela pensao de sangue (viuvez)			62	621	6210	62103	621031
62.1.0.4.1	Recebeu pela pensao de alimentacao			62	621	6210	62104	621041
62.1.0.5.1	Recebeu em juros do banco ou dos devedores			62	621	6210	62105	621051
62.1.0.6.1	Recebeu pelos seguros			62	621	6210	62106	621061
62.1.0.7.1	Recebeu de instituicoes sem fins lucrativos e religiosas			62	621	6210	62107	621071
62.1.0.7.2	Recebeu em especie de instituicoes sem fins lucrativos e religiosas			62	621	6210	62107	621072
62.1.0.8.1	Recebeu de familiares que vivem fora do agregado			62	621	6210	62108	621081
62.1.0.8.2	Recebeu em especie de familiares que vivem fora do agregado			62	621	6210	62108	621082
62.1.0.9.1	Recebeu de familiares que trabalham no estrangeiro			62	621	6210	62109	621091
62.1.0.9.2	Recebeu em especie de familiares que trabalham no estrangeiro			62	621	6210	62109	621092
62.1.0.9.5	Recebeu de xitique			62	621	6210	62109	621095
62.1.0.9.9	Outra transferencias recibidas			62	621	6210	62109	621099

Annex B: IAF2002 Imputation Summary

The method used for imputation of incomplete data for IAF2002 is explained in detail in the program specifications for two VBA modules: mImputeAC and mAggrProdHh. These appear in the document IAF2002 database summary on pages 9-12.

The Excel workbook ImpAggregationSummary.xls contains summaries of the various imputation steps.

This spreadsheet is on the IAF2002 CD in \IAF2002\I Semestre\Tables\IE.

Worksheet **ACQnoV** summarizes the number of autoconsumo observations with quantity but no value by product. The value for these observations is imputed using a unit price for the product as close as possible in time and space to observation with missing data. There are a total of 1,337 observations with quantity and no value out of 57,373 observations.

The search for the unit price is done in the following order (where QC is the Questionario Communitario and DD are the daily purchases):

- (1) QC average price for province, urban/rural, period, product
- (2) DD average price for province, urban/rural, period, product
- (3) QC average price for province, urban/rural, product
- (4) DD average price for province, urban/rural, product
- (5) QC average price for urban/rural, product
- (6) DD average price for urban/rural, product
- (7) QC average price for urban/rural, product (opposite of urban/rural)
- (8) DD average price for product

The source used for the price is recorded in the AC table variable ImputeInd.

Worksheet **ACimpSumm** summarizes the number of observations with imputed values by product and source of unit price. Sources of unit prices are numbered 1 to 8 to correspond to the sources outlined above. A “?” indicates that no price was found and no value imputed. 1,261 observations were successfully imputed out of the 1,337 candidates.

Worksheet **ACnoQV** summarizes autoconsumo observations with no quantity and no value. These observations record the use of a product by the household with no definite quantity or value. The most common products are lenha (firewood) for which there are 828 observations and papaya with 34 observations. Values for these two products are imputed using the median value of observations with value but no quantity. A summary of these observations appears in worksheet **ACVnoOrange**. For lenha (firewood) 3,000 meticais is the median and for papaya it is 2,000 meticais. These medians were calculated using 9,639 observations for firewood and 36 observations for papaya. All

other products with no quantity and no value were left unchanged. The variable **ImputeInd** was set to 9 for the imputed observations.

The same procedures were used for *Receitas em Espécie (RE)* observations. The worksheets **REQnoV**, **REimpSumm**, **REnoQV** and **REVnoQrange** record the results of the imputations done for the RE data. There are 139 observations with quantity and no value out of a total of 5636 RE observations. Values were imputed for 133 of these observations.

The only product with no quantity and no value for which a value was imputed is product 111300 “Refeições preparadas pelos agregados” (meals prepared for the household). There were 176 RE observations imputed with a value of 5,000 meticaís based on 323 observations with value but no quantity.

Worksheet **DDMexcluded** summarizes the number of daily purchase transactions that were excluded from the aggregation procedure because of double counting between the daily and monthly purchase questionnaires. These involve products that are recorded on the daily purchases questionnaire, which occurred on days before the monthly questionnaire was completed. These involve products such as soap, matches and oil for lamps that are typically purchased in small quantities. A total of 428 observations worth 2,107,700 meticaís were excluded.

Worksheet **RentRange** summarizes the distribution of the value of rent observations recorded in the monthly expenses questionnaire. Actual rent is recorded for 181 households and imputed (estimated) rent is recorded for 3,752 households. The remaining households have no rent information. The median value of observed rents (actual and estimated) for the province and urban/rural is used to impute rent for households with no data. The worksheet **ImputedRent** summarizes the number and value of imputed rent observations by province and urban/rural.