Rwanda - Integrated Household Living Conditions Survey, Wave 1, 2000-2001.

National Institute of Statistics, Rwanda

Report generated on: March 2, 2021

Overview

Identification

ID NUMBER
RWA_2000-2001_EICV-W1_v01_EN_M_v01_A_OCS

Overview

ABSTRACT
The HLCS, with an expanded budgets and consumption module, was primarily intended to provide policy planners and decision-makers with basic data on household living standards in Rwanda. In addition, the survey was to be used to:

- calculate weights for the Consumer Price Index and estimate final household consumption,

- measure the effect of macro-economic policies and projects on the conditions and living standards of the population,

- produce key indicators of household welfare in order to assist policy-makers and development partners to improve the design of their development strategy,

- identify policy target groups with a view to ensuring that state interventions are better targeted,

- provide information on the socio-economic characteristics of households with a view to setting up a socio-economic data base,

- carry out in-depth studies, for example on poverty, nutrition, housing conditions, etc,

- improve the national capability to conduct statistical surveys, however complex they may be.

KIND OF DATA
Sample survey data [ssd]

UNITS OF ANALYSIS
Households

Scope

NOTES
The information gathered during the survey will be used primarily to provide information on assorted household and personal level characteristics which can be analyzed vis a vis the household's consumption. The primary household and person characteristics that are gathered in this survey in order to provide relevant indicators are:

-School attendance and literacy. This includes information to compute net and gross enrollment rates

-Health and fertility. Some indicators such as maternal mortality are outside the scope of the survey. In this case, a more appropriate survey like the DHS may be recommended

-Migration

-Employment and economic activity.

-Land ownership and other agricultural based indicators.

The survey is also designed to provide important information for the computation of National Accounts and rebasing the Consumer Price Index.

TOPICS

Topic	Vocabulary	URI
ECONOMICS [1]	CESSDA	http://www.nesstar.org/rdf/common
EDUCATION [6]	CESSDA	http://www.nesstar.org/rdf/common
HEALTH [8]	CESSDA	http://www.nesstar.org/rdf/common
DEMOGRAPHY AND POPULATION [14]	CESSDA	http://www.nesstar.org/rdf/common

Coverage

GEOGRAPHIC COVERAGE
National coverage.

GEOGRAPHIC UNIT
cell level

UNIVERSE
Household members (institutional and itinerant populations excluded).

Producers and Sponsors

PRIMARY INVESTIGATOR(S)

Name	Affiliation
National Institute of Statistics, Rwanda	Government of Rwanda

FUNDING

Name	Abbreviation	Role
Department for Intenational Development	DFID	Bilateral funding assistance
World Bank	WB	Financial assistance
United Nations for the Children	UNICEF	Financial assistance
United Nations for Development Program	UNDP	Financial assistance
African Developmennt Bank	ADB	Financial assistance

OTHER ACKNOWLEDGEMENTS

Name	Affiliation	Role
Oxford Policy Management	DFID	International Technical Assistance
MINECOFIN	Government of Rwanda	Primary user of data (EDPRS)

Metadata Production

METADATA PRODUCED BY

Name	Abbreviation	Affiliation	Role
Office of Chief Statisitician	OCS	Food and Agriculture Organization	Metadata adapted for FAM
National Institute of Statistics of Rwanda	NISR	Ministry of Finance	Data and metadata producer and deposit

Department of International Development	DFID	British Government	Provided technical assistance for archiving the data set
Ruben MUHAYITETO		NISR	Revision of DDI

DDI DOCUMENT VERSION
RWA_2000-2001_EICV-W1_v01_EN_M_v01_A_OCS_v01

DDI DOCUMENT ID
DDI_RWA_2000-2001_EICV-W1_v01_EN_M_v01_A_OCS_FAO

Sampling

Sampling Procedure

The sampling plan was drawn up with the technical support of the late Christopher SCOTT, Survey Consultant, during his mission in July 1997.

The two main factors considered in designing the sampling plan were:
- the objectives of the survey,
- the fieldwork methodology given the available logistical resources.

For the survey, one objective was determinant; the Government wanted statistically reliable results at the level of each province, Kigali city and the "other urban sector". Thus, the objective called for 13 domain of analysis. Experience of conducting this type of survey shows that a minimum sample of 500 households per domain of study is required for sound analyses.

Thus, the total sample size was 6,450 households, with 1,170 households for urban areas and 5,280 households for rural areas. A two stage stratified sample was used: sampling at area level and at household level.

At the area level, the chosen sampling base (or at the enumeration district) was the "cellule"in the rural areas and the zone in urban areas, since they are usually fairly homogeneous in size and are well demarcated.

Knowledge of the size of each cellule enabled the use of the classical method of sampling with probability proportional to size at the first stage. A list of all cellules including estimates of the number of households in each was compiled from information provided by the local authorities.

For sampling at the household level, an up-dated list of households was prepared for each of the selected first stage cellule by carrying out a listing in each sampled cellule simultaneously but with a lag in data collection before or while collecting the data. Part of this operation was carried out in collaboration with the National Population Office (ONAPO) and the Food Security Research Project (FSRP) of MINAGRI.

The first stage selection involved drawing cellules using systematic sampling (i.e. fixed interval drawing), with probability proportional to size. In order to ensure a good spatial distribution of the sample, the selected cellules were ordered according to geographical location. The second stage involved selecting households by systematic sampling. A list of households by cellule was drawn up according to proximity by listing agents. This ordering of sampling units before selection constitute an implicit stratification, which makes the sample more representative.

Weighting

In order for the estimates from each survey to be representative at the national level, it is necessary to apply sampling weights to the survey data. The weights for the sample households were calculated as the inverse of the overall probability of selection, taking into account each sampling stage. Given the nature of the sample design and the new listing of households, the weights vary by sample ZD. An Excel spreadsheet with all the sampling frame information for the sample ZDs was used for calculating the weights, which were then attached to the corresponding records in the survey data files.

There are two kinds of weighting: spatial weighting and temporal weighting. Use of these methods enabled annual estimates to be obtained for the whole of the Rwandan population, and are defined below:

1. Spatial weighting enables results relating to the sample to be extrapolated for the whole of the population for the same period. It was calculated using the inverse of the overall probability of selection of a particular household. The details of the theory for calculating the various probabilities are shown in Annex I. Starting from the overall probability formula (Fhi=p1hi x p2hi) where p1hi is the probability proportional to size of drawing cellule i in stratum h and p2hi is the conditional probability of drawing a household knowing that unit i of stratum h has been selected. The numbers 1 and 2 indicate the stage or level of sampling. Spatial weighting is given by the formula (Whi=1/Fhi=Mhi/ahbhi) where Mhi is the total number of households in unit i of stratum h and ah is the number of sample units in stratum h, while bhi is the number of households surveyed in unit i of stratum h .

2. Temporal weighting is intended to produce annual estimates of values relating to the survey period. Thus, the temporal weighting coefficient depends on the length of the collection period. By using CPTmj to designate the coefficient of temporal

weighting of the variable ymj for household m, and Jmj to designate the number of collection days;
Ymj = CPTmj x ymj or CPTmj= 365/Jmj
Ymi being the annual value of the variable ymj for household m.

Questionnaires

Overview

The questionnaires is published in french

Three types of questionnaire were used in the field for data collection:
- the household questionnaire comprising of 12 modules divided in two parts, A and B.
- the community questionnaire for collecting data on economic and social infrastructures in the sample units in rural areas and
- a conversion form for non-standard units used by households.

Household questionnaires

Part A collects data on each member of the household. It covered the following areas:
- demographic and migration characteristics,
- education and health,
- employment and housing.

Part B deals with the economic activity of the household. It comprises of the following five modules:
- agro-pastoral activities and own-produce consumption,
- household expenditure,
- non-agricultural economic activities,
- transfers,
- durable goods, access to credit and savings.

Data Collection

Data Collection Dates

Start	End	Cycle
1999-10-24	2000-12-24	First cycle
2000-07-19	2001-07-10	Second cycle

Time Periods

Start	End	Cycle
1999-10-24	2001-07-19	10

Data Collection Mode

Face-to-face paper [f2f]

Questionnaires

The questionnaires is published in french

Three types of questionnaire were used in the field for data collection:
- the household questionnaire comprising of 12 modules divided in two parts, A and B.
- the community questionnaire for collecting data on economic and social infrastructures in the sample units in rural areas and
- a conversion form for non-standard units used by households.

Household questionnaires

Part A collects data on each member of the household. It covered the following areas:
- demographic and migration characteristics,
- education and health,
- employment and housing.

Part B deals with the economic activity of the household. It comprises of the following five modules:
- agro-pastoral activities and own-produce consumption,
- household expenditure,
- non-agricultural economic activities,
- transfers,
- durable goods, access to credit and savings.

Data Collectors

Name	Abbreviation	Affiliation
National Institute of Statistics, Rwanda	NISR	Government of Rwanda

Data Processing

Data Editing

Questionnaires were reviewd by the controller in the field before they were dispatched for data entry. A control sheet was provided to the contollers to assist in the process of manually editing the questionnaires. Questionnaire structures were verified when the questionnaires were checked in prior to data entry. Three contracted persons reviewed the questionnaire and filled in a form that served as a primary data control sheet. Automated data editing was largely done during the data entry phase (see "Other Data Processing" for details). Some batch edit programs were used to identify inconsistent data.

Data iimputation was largely done during the analysis phase by analysts. However, a "structural" imputation on the microdata was required for the own consumption data. This was done to adjust for erroneous pricing when the unit for measuring own consumption was buckets. For more information, please refer to the SPSS syntax files orthe data processing report.

Coding of products was based on sequential codes for each section.

Data Appraisal

Estimates of Sampling Error

Given that the survey estimates are subject to sampling variability, it is important to calculate the sampling errors for the most important estimates from each survey. The sampling error is measured by the standard error, or square root of the variance of the estimate. The CENVAR software, a component of the Integrated Microcomputer Processing System (IMPS) developed by the U.S. Census Bureau, was used for tabulating the standard errors and other measures of precision, taking into account the stratification and clustering in the sample design. The CENVAR output tables show the value of the estimates, standard errors, coefficients of variation, 95 percent confidence intervals, design effects and number of observations. Given that the confidence intervals provide a user-friendly interpretation of the sampling variability, an annex was produced with tables showing the 95 percent confidence intervals for the most important estimates from the EICV1 and EICV2 data appearing in the preliminary report. These tables provide a quick conservative test to determine whether any difference between the EICV1 and EICV2 estimates is statistically significant.

The INSR was also provided with tables showing the full CENVAR results. The design effect is defined as the variance of an estimate based on the actual sample design divided by the corresponding variance based on a simple random sample of the same size; it is a measure of the relative efficiency of the sample design. In comparing the CENVAR results from EICV1 and EICV2, it was found that the design effects are generally lower for EICV2, indicating that the stratification used for this survey was very effective. Given that the EICV1 was based on an older sampling frame from the 1991 Rwanda Census, this also contributed to the higher design effects for the EICV1 estimates.

Other forms of Data Appraisal

In the process of filling in the questionnaires and data entry, various types of error slipped into the data. Controls were carried out on a number of levels; in the field by the controllers and supervisors, and at the Statistics Department after data entry. More detailed checks and controls were carried out after data entry, since the process can itself introduce errors.

In addition to corrections made at the time of data entry, the data was processed using the following 6 main stages:

- Exhaustivity control: This involved checking the use of identical geographical codes in various data files and verification that questionnaires had not been entred more than once or omitted.
- Consistency between variables: With the aid of absolute frequency tables, verification is made whether eligible respondents for all the questions replied and whether those not eligible did not in effect reply.
- Standardisation: Some quantitative variables were aggregated over the year before validation. Variables arising from local measurements were converted to the conventional measurement system.
- Re-coding: Certain continuous, quantitative variables were divided into classes:
- Creation of derived variables: This involved variables (which are derived from other variables.) not in the questionnaire or the data dictionary
- Imputation of values: During processing, extreme values were encountered for some variables. These were confined to values that deviated more than three standard deviations from the mean. After verification, they were replaced by the mean value of the variable.

A number of programming software and languages were used from capturing the data to preparing tables of results, inter alia IMPS, CS PRO, MS ACCESS, Visual Basic and COBOL, SPSS.