Sampling

Sampling Procedure

Introduction

The sampling frame for the EICV1 was based on the data and cartographic materials from the 1991 Rwanda Census of Population and Housing, while the EICV2 was based on the 2002 Rwanda Census frame. There were significant changes in the areas considered urban between the two censuses, but these geographic changes are taken into account in the comparative analysis between the EICV1 and EICV2 data. The sample design for EICV1 is described in the report on EnquÃªte IntÃ©grale sur les Conditions de Vie des MÃ©nages (Avec Volet Budget - Consommation) - Plan de Sondageâ€ (Scott, July 1997). A detailed description of the EICV2 sample design is found in the report on Recommendations on Sample Design and Estimation Methodology for the Rwanda EnquÃªte Integrale sur les Conditions de Vie des MÃ©nages 2005. (Megill, June 2004).

A stratified two-stage sample design was used for both the EICV1 and EICV2. The primary sampling units (PSUs) were the enumeration areas or zones de dÃ©nombrement (ZDs) defined for the census. The sample of ZDs in each stratum was selected with probability proportional to size, where the measure of size was based on the number of households from the census frame. A new listing of households was conducted in each ZD, and a sample of households was selected at the second sampling stage. The units of analysis are the households and the individual members of the household.

One of the objectives of EICV1 and EICV2 was to provide reliable estimates of household consumption and other characteristics at the level of the 12 old provinces, as well as at the national level, City of Kigali, other urban and rural. Later the country was divided into five new provinces; given the larger size of the new provinces, the corresponding estimates will have better precision than those at the old provincial level.

Stratification

The stratification of the sampling frame for both EICV1 and EICV2 was designed to improve the efficiency of the sample design and ensure a sufficient sample size for the major geographic domains of analysis. The sampling frame for these surveys was stratified by the 12 old provinces, as well as by urban and rural areas. At the national level three residential strata were defined: (1) City of Kigali, (2) other urban, and (3) rural. In the case of EICV1, the ZDs in the urban and rural strata for each province were ordered geographically to provide a corresponding implicit stratification.

In the case of the City of Kigali, there is a higher variability in socioeconomic characteristics compared to the other domains. Therefore a socioeconomic stratification was defined for the ZDs in the EICV2 sampling frame for the City of Kigali, using an indicator of bien-Ãªtre (well-being) based on housing characteristics in the 2002 Rwanda Census

data. The ZDs were coded by four socioeconomic quartiles, and this was used as a sorting variable to provide a corresponding implicit stratification. A new stratification code for "semi-rural" was introduced into the sampling frame for EICV2 to identify urban ZDs with at least 70 percent of households with agricultural operations (based on the 2002 Rwanda Census data). This "semi-rural" code was used as one of the sorting criteria for the sampling frame of the City of Kigali and the other urban stratum in each province. Within each stratum, the ZDs in the sampling frame were further sorted geographically to provide an additional level of implicit stratification.

Given that the rural economy is primarily agricultural, the socioeconomic characteristics of the rural households are generally correlated with the crop and livestock activities found in the different bio-climatic zones. Therefore the EICV2 sampling frame for rural strata was sorted by the ten bio-climatic zones as well as geographic codes to provide an effective implicit stratification.

Sample size and allocation

The sample size for EICV1 and EICV2 was determined by the precision required for the survey estimates for each domain, as well as by the resource and operational constraints. The total sample size for EICV1 was 570 ZDs and 6,450 households. For EICV2 this sample size was increased to 620 ZDs and 6,900 households, in order to provide a larger sample for the urban strata. One reason for increasing the urban sample for EICV2 was because of the expansion of urban areas following the 2002 Rwanda Census. The effective sample size for EICV1 was actually 6,420 households, since 30 non-interviews were not replaced for this survey.

Given that one of the objectives of these surveys was to produce reliable estimates for each of the 12 old provinces, a total of 40 sample rural ZDs was allocated to each province. A larger sample was allocated to the City of Kigali because of the larger variability of socioeconomic characteristics; 80 sample ZDs were selected in this domain for EICV1 and 100 ZDs for EICV2. In the case of the other urban strata, a sample of 50 ZDs for EICV1 and 80 ZDs for EICV2 were allocated to the 11 other provinces proportionately to their urban population. Table A.1 presents the distribution of the sample for EICV1, and Table A.2 shows the corresponding distribution for EICV2.

For EICV1 the number of households selected per sample ZD was 9 for the City of Kigali and the other urban stratum, and 12 for the rural stratum. This was an effective sampling strategy given that the urban strata generally have more variability between ZDs and homogeneity of households within ZDs. This approach also provided a reasonable workload for the enumerators in the urban and rural ZDs based on the data collection procedures each cycle. Therefore this same sampling strategy was used for EICV2.

Sample Selection Procedures

For both EICV1 and EICV2 the ZDs within each stratum were selected systematically with probability proportional to size, where the measure of size was based on the number of households in the ZD from the

corresponding census frame (1991 for EICV1 and 2002 for EICV2).
Following a new listing of households in the sample ZDs, at the second
stage 9 sample households were selected systematically in each sample
urban ZD and 12 sample households were selected in each rural ZD. This
sampling strategy provided an approximately self-weighting sample (that
is, the sampling weights were similar) within each stratum. A sample
of possible replacement households was also selected systematically
within each sample ZD. Whenever an original sample household could not
be interviewed for any reason, it was substituted by one of the random
replacement households.

Deviation from Sample Design

As indicated,any household that was not interviewed as per the original
listing and selection was replaced with a reserve household. Each EA
had 4 households on reserve. A total of 522 households were replaced
over the course of the survey. In addition, several EAs were swapped
from their scheduled cyclic visit due to seasonal accessability
problems.

```
Cluster From                    To
337.........Cycle 6..........Cycle 2
364.........Cycle 8..........Cycle 2
131.........Cycle 6..........Cycle 3
269.........Cycle 5..........Cycle 3
270.........Cycle 5.........Cycle 3
264.........Cycle 3.........Cycle 5
271.........Cycle 3.........Cycle 5
132.........Cycle 3.........Cycle 6
327.........Cycle 2..........Cycle 6
362.........Cycle 2..........Cycle 8
```

Response Rates

Out of the 6900 household sample, 92.4% responded. All 7.6% of
households that were not interviewed (for the reasons provided below)
were replaced.

Reason for non-
response...HHolds......
....precent

Yes...
........................6378.............92.4%
No, address not found or uninhabited
dwelling..................254...............3.7%
No, change of
residence..139..
..............2.0%
No, sick or
deceased..
..29................0.4%

```
No,
refused...........................................................
.................27.................0.4%
No,
other.............................................................
.................73.................1.1%

Total.............................................................
.................6900...........100.0%
```

The City of Kigali and the old province of Butare had the highes
refusal rates with about 14% of the original selected households being
replaced. The primary reason given for replacement was the inability
to positively identify the dwelling (or the selected dwelling was found
unihabited).

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.