Information and Communication Technology on households (ICT)

TYPE

Methodology for Community statistics on the information society: Individuals, households and the information society - under the scope of the draft Commission Regulation implementing Regulation 808/2004 as agreed by the Statistical Programme Committee.

Sampling method

The sample of households for the ICT survey of the year 2005 is consisted of the sample used in the Greek Survey of Income and Living Conditions (EU-SILC) of the year 2004. The EU- SILC is an annual rotating household survey covering the target population of the ICT survey.

Additionally, the EU-SILC is a multistage stratified sampling survey with primary sampling unit the area (one or more unified blocks) and final unit the household. The sample design of EU-SILC of the year 2004 was based on data coming from the population census of the year 2001.

The reasons for using the sample of the EU-SILC to collect the ICT information was the operational convenience, the low cost and the facility in creating a representative sample of the requisite size.

Survey technique

The data collection will be based on telephone interviews.

From data collected through the EU-SILC of the 2004 survey, the composition of the households is already known among all the households' members one, in the age sixteen to seventy four years, will be randomly selected, with the use of random numbers.

The interviewer will telephone the household and will ask to speak to the pre-selected for the survey person, in order to get the information concerning the whole household and him/herself (rest of the questionnaire). If that is feasible, then the interviewer will fill in the whole questionnaire collecting information by the member of the sample. In case, the member of the sample is temporarily away from home (working, in a visit, etc.) the interviewer makes another appointment with the pre-selected member in order to complete the questionnaire.

If the member is moved away from the household permanently, then another household member, fulfilling the age prerequisite of 16-74 years old, will replace him. The same held if the pre-selected member will join the army or is gone away for studies, etc.

Target population

In accordance with the general outline of EUROSTAT's plan for the Household Survey on ICT usage, the target population is the current

population of persons aged sixteen to seventy four who are resident within private households in Greece. The target population consisted from 8.160.045 individuals living in 3.726.806 households. The target population was estimated from the data coming from the Labour Force Survey of the year 2004.

Sampling unit

The sampling unit is the private household, and one person randomly selected among the household members of age sixteen to seventy four years.

Sampling frame

The sampling frame containing the primary units (cluster of households in one or more unified blocks) for the ICT survey is the same as for the EU-SILC of the year 2004, which is an area frame constructed using the necessary information from the recent Greek General Population Census 2001, and provides complete coverage of the target population of this survey.

In each selected primary unit, the sampling frame containing the households is updated on month before the data collection.

Sampling design

The multistage stratified area sampling is applied for the survey. The primary units are the areas (one or more unified blocks) and secondary sampling units selected in each sampling area are the households containing members belonging to the target population. The final unit is one person randomly selected among the household members of age sixteen to seventy four years.

Stratification

The sampling design involves two levels of area stratification:

(i) The first level is geographical stratification based on the partition of the total country area into thirteen standard regions corresponding to the European NUTS II level. The two major city agglomerations of Greater Athens and Greater Thessalonica constitute separate major geographical strata.

(ii) The second level of stratification involves grouping municipalities and communes within each NUTS II region by degree of urbanization, i.e., according to their population size, into four categories. These categories are defined by the population size intervals 0-999, 1000-4999, 5000-29999, 30000 and over. The number of final strata in the thirteen regions, i.e., non-empty strata formed by crossing region and degree of urbanization, was 50. The two major city agglomerations were further partitioned into 31 and 9 substrata (administrative subsections), respectively, on the basis of the city blocks of the

municipalities that constitute them. Thus, the total number of strata for this survey was 90.

Clustering

The sampling design involves clustering of households in the area units that comprise the area frame of each final stratum. By area unit we mean a part of inhabited area ending at artificial or natural boundaries well defined and identifiable on the ground by using a map of the locality. Such a unit could be one or more neighbouring blocks, or part of a rural locality with such boundaries. To reduce field costs and the time length of the fieldwork, the size of the area clusters is limited to an average of approximately fifty households in the thirteen administrative regions, and approximately seventy households in the two major city agglomerations.

Sample size

The sample will be 5.500 households belonging to 1.056 primary sampling units

Stages of probability sampling

The sample of private households and associated eligible residents will be drawn in three stages. In the first stage, the primary sampling units (PSUs), were systematically selected from the frame of clusters in each final stratum (=Geography X Urbanization) with probability proportional to the number of private households in the clusters (according to data coming from the last population census of the year 2001). A total number of 1056 PSUs were selected in the EU-SILC of the year 2004, with an average of 60 households per PSU according to Census 2001 counts. In the second stage, a systematic random sample of households was drawn with a pre-fixed sampling rate from the current population of households (based on a list prepared in the field) of each selected PSU. In the third stage of sampling, one of the residents aged sixteen to seventy four of each selected household will be selected at random. Analytically, the sampling selection process by stage (1st, 2nd and 3rd stage) is as follows:

1st stage: The primary unit of order i in stratum h has probability of being drawn proportional to the target population size as follows:

$$P_{hi} = \frac{N_{hi}}{N_h}$$

N_{hi}: The updated (from EU-SILC survey) target population of households in the hi primary unit

N_h: The updated target population of households in the h stratum.

2nd stage: Out of N_{hi} households, a sample of n_{hi} households was

selected with equal probabilities. Each of n_{hi} households has the same chance to be selected, equal to: $\frac{n_{hi}}{N_{hi}}$. As the estimator of the stratum total Y_h (for any characteristic) should be self-weighting, the n_{hi} was defined, as follows: $n_{hi} = \frac{n_h}{a_h}$, where $n_h = \sum_i n_{hi}$ and a_h is the number of primary units in the h stratum.

3rd stage: In this stage from each household one individual (member of household belonging to the target population) is selected with equal probabilities.
Let be p_{hij} the selection probability of the hij individual, which belongs to the hi household. As one individual was selected with equal probabilities out of m_{hi} members belonging to target population, the p_{hij} is defined as:

$$p_{hij} = \frac{1}{m_{hi}}.$$

Estimation of the survey characteristics
Weightings - Design factor

Let h be one of the final strata of households (Final stratum = Geography x Urbanization), then this will take the following values: $h = 1,2,...,H$ (where $H = 90$). In each of the final strata (let h), if statistical information was selected from a sample of n'_h households, the extrapolation factor of the household of order i was defined as:

$$w_{hi} = \frac{N_h}{n_h} \cdot \frac{1}{r_h} t_{hi}$$ where:

N_h: The target population size (number of households) in the h stratum

n_h : The initial sample size in the h stratum

$\frac{N_h}{n_h}$: The inverse of the initial probability of selection of the sampling households in the the h stratum, as the the estimator of the stratum total

Y_h (for any characteristic) is self-weighting,

$r_h = \frac{n'_h}{n_h}$ is the response rate in the h stratum

t_{hi}: Factor, which adjusts the sample weights of households so that the estimated totals conform to the population totals on a cell-by-cell basis (Population Weighting Adjustment). The auxiliary variable used at household level is the household size (1,2,3,4 or 5+ members) for the definition of cells or classes.

In each of the final strata of households (let h), if statistical information was selected from a sample of m_h individuals, the extrapolation factor of the individual of order j belonging to the hi household is defined as follows:

$w_{hij} = w_{hi} \cdot \frac{1}{p_{hij}} \cdot g_{hij}$ where:

w_{hi}: The extrapolation factor of the hi household in which the hij individual belongs

g_{hij} Factor, which adjusts the sample weights of individuals, so that the estimated distribution conform to the population distribution across a set of classes. The classes are 24, which are defined by crossing sex by age groups (2 sex categories × 12 age groups). The age groups are defined by the year intervals: 16-19, 20-24, 25-29,30-34, 35-39,40-44, 45-49, 50-54, 55-59,60-64, 65-69 and 70-74. The population distribution of individuals by sex and age groups was estimated from data coming from the Labour Force Survey with the same reference period.