



HELLENIC REPUBLIC



HELLENIC STATISTICAL AUTHORITY

# Community Survey on ICT usage in households and by individuals 2010 Quality Report

## Please read this first !!!

General guidelines on using this template

- In the title line of this page, please delete the non-applicable term (Metadata / Quality Report).
- Fill in the required information in the space (box) foreseen next to or below the item heading, if a box is irrelevant for your national survey, indicate 'non-applicable' to avoid we have to come back to you on this item. An increase of the box' size after inserting several lines or paragraphs is no problem. However, when reporting several pages for one item, we kindly ask you to give a short summary and refer to the full text in an annex.
- Keep the numbering of the chapters and items. Additional comments can be given at the end of the report.
- This template is designed to serve both the requirements for the Metadata as well as the Quality report. **Chapters 1 to 6 shall be completed for the metadata report (deadline for submission: before 31/05/2010), chapters 7 to 10 can be postponed until the Quality report (deadline for submission: not later than 05/11/2010).** However, where provisional information for the Quality Report topics is already available, we invite you to provide us with this data in the Metadata report (and update it in the Quality Report).
- Please submit the national questionnaire used (in national language and if available in English) annexed to the metadata report.
- Please replace in the header field the code 'EU' with your country code.
- **All information provided in this report on coverage of questions/items, net sample sizes, number of respondents, proportions, etc. should be in line with the transmitted data file(s).**

We kindly thank you for respecting these guidelines.

## 1. Cover information

1.1	Country	GREECE
1.2	<b>Organisation responsible for the survey</b> <i>Please also indicate the organisation <u>running</u> the survey if different from the organisation responsible (e.g. because of sub-contracting).</i>	HELLENIC STATISTICAL AUTHORITY

<b>1.3</b>	<b>Contact person(s)</b> (name, unit, e-mail, phone, fax)	<p>1. <i>CHALKIADAKI MARIA</i> UNIT OF HOUSEHOLD SURVEYS TEL. 0030 -213-135 2896 FAX. 0030 -213-135 2906 E-MAIL: MCHALK@STATISTICS.GR</p> <p>2. <i>ZOULIATIS IOANNIS</i> UNIT OF HOUSEHOLD SURVEYS TEL. 0030 -213-135 2941 FAX. 0030 -213-135 2906 E-MAIL: ZOULIATI@STATISTICS.GR</p>
<b>1.4</b>	<b>Name of the collection</b> The name of the survey in its original language(s) and in English (e.g. name used in the statistical office's English website).	<p>SURVEY ON THE USE OF INFORMATION AND COMMUNICATION TECHNOLOGIES FROM THE HOUSEHOLDS</p> <p>ΕΡΕΥΝΑ ΧΡΗΣΗΣ ΤΕΧΝΟΛΟΓΙΩΝ ΠΛΗΡΟΦΟΡΗΣΗΣ &amp; ΕΠΙΚΟΙΝΩΝΙΑΣ ΑΠΟ ΤΑ ΝΟΙΚΟΚΥΡΙΑ</p>
<b>1.5</b>	<b>Last update of this report</b>	27-10-2010

## 2. General methodological information

<b>2.1</b>	<p data-bbox="228 313 1428 353"><b>Reference period(s)</b></p> <p data-bbox="228 353 1428 414">The main reference period for the ICT variables as well as the background variables, e.g. <i>first quarter of the year</i> or <i>last three months before the interview</i> (with an indication of the respective months), or a specific date.</p> <ul data-bbox="284 436 1428 638" style="list-style-type: none"> <li>• 31st of March 2010 for educational level completed</li> <li>• Day of the survey conduction for activity status, employment situation, A1-A5, C4, E1, E4, E5, E6, E7, E8, F1, G1, G2, G6, G7, G8, G10, G11, G12, G13, G14</li> <li>• Three first months of 2010 for questions B2, C2, C3, C5, C6.</li> <li>• Last 12 months (April 2009-March 2010) for questions D2, D3, D4, D5, E2, E3.</li> </ul>
<b>2.2</b>	<p data-bbox="228 672 1428 712"><b>Survey period</b></p> <p data-bbox="228 712 1428 772">The beginning and end date – if already known – of the data collection period.</p> <p data-bbox="228 772 1428 873">28 April 2010-30 June 2010 (maybe an extension of 15 days will be used)</p>
<b>2.3</b>	<p data-bbox="228 907 1428 947"><b>Survey vehicle</b></p> <p data-bbox="228 947 1428 1030">Stand-alone or embedded in another survey. If embedded, give a short description of the survey the ICT modules are inserted in.</p> <p data-bbox="228 1030 1428 1120">ICT is a stand-alone survey.</p>
<b>2.4</b>	<p data-bbox="228 1153 1428 1193"><b>Survey type</b></p> <p data-bbox="228 1193 1428 1276">Short description of the survey type (face-to-face interview, self-administered mail survey, telephone interview, combination of techniques, other; etc.).</p> <p data-bbox="228 1276 1428 1366">Telephone interview</p>
<b>2.5</b>	<p data-bbox="228 1400 1428 1440"><b>Survey participation</b></p> <p data-bbox="228 1440 1428 1500">Please indicate whether the survey is mandatory or voluntary.</p> <p data-bbox="228 1500 1428 1603">Participation is mandatory according to Greek law.</p>

## 2.6 Main methodological differences compared to previous survey(s)

If any, indicate the changes in methodology that may have an impact on the (comparability over time of the) results delivered to Eurostat, e.g. change in reference period, new reference sampling frame, different scope, different grossing-up method, different treatment of non-response, etc. No need for giving detailed technical analyses, a bullet point overview of the main differences and the expected impact is sufficient.

The sample of households for the ICT survey of the year **2010** is consisted of a sub-sample of the samples used in the National Health Survey of the year 2009 and the EU-SILC of the years 2005-2008.

### 1. Sample design of the National Health Survey

#### 1.1. Target population

The surveyed units are the private households and the individuals. One individual is surveyed from each household of the sample. The target population of individuals is consisted from the individuals with age 15 years or over, at the whole country.

Only the private households participate in the survey. Individuals permanently reside in collective houses (as hospitals, hotels, asylums, houses of old people, orphanages etc) are not covered by the survey. These individuals are - as a rule - members of institutional households. If however, we subtract from this population, the conscripts and the imprisoned, the actual percentage not covered by the survey procedure, accounts for 2% of the total population, and in its major part concerns economically non-active persons.

##### 1.1.1 National population size

- Households: 4.114.150 private households
- Individuals: 9.305.935 individuals aged 15<sup>+</sup>

#### 1.2. Sampling frame

The multi-stage area sampling was adopted for the survey. The primary units are the areas (one or more unified city blocks), the secondary sampling units selected in each sampling area are the households and the ultimate sampling units selected in each sampling household are the individuals belonging to the target population. The sampling frame of the primary units is based on the data from the last population census of the year 2001. The sampling frame containing the secondary units (households) in the selected sampling primary units is a list of households *updated* before the selection of households.

#### 1.3 Sampling method

The multi-stage stratified sampling method was applied with stratification variables the (a) Region (NUTS 2) and (b) the degree of urbanization. In each Region (NUTS 2), the stratification of primary units was conducted by allocating the Municipalities and Communes according to the degree of urbanization (urban, semi-urban, and rural regions). Except for the two former Major City Agglomerations (Athens and Thessaloniki), the created strata according to the degree of urbanization are:

Urban	Agglomerations and Municipalities with 10.000 inhabitants or more
Semi-urban	Municipalities and Communes with 2.000 to 9.999 inhabitants
Rural	Communes up to 1.999 inhabitants

The former Greater Athens Area was divided into 31 strata of about equal size (equal number of households) on the basis of the lists of city blocks of the Municipalities that constitute it and taking into consideration socio-economic criteria. Similarly, the former Greater Thessaloniki Area was divided into 9 equally sized strata. The two former Major City Agglomerations account for 40% of total population and for even larger percentages in certain socio-economic variables.

## 1.4 Selection of the sampling units

**1<sup>st</sup> stage of sampling:** The primary sampling units in each stratum (stratum=Region x degree of urbanization) were selected with probabilities proportional to their sizes (number of households in the last population census of the year 2001)

**2<sup>nd</sup> stage of sampling:** The secondary sampling units (households) in each selected primary unit were drawn with equal probabilities.

**3<sup>rd</sup> stage of sampling:** In each selected household, one individual belonging to the target population was selected from the members of household with equal probability.

## 2. Sample design of the EU-SILC

### 2.1 Type of sample design

The two-stage area sampling is applied for the EU-SILC survey.

### 2.2 Sampling units

The sample of private households was selected in two stages. The primary units are the areas (one or more unified building blocks) and the ultimate sampling units selected in each sampling area are the households.

### 2.3 Stratification and sub-stratification criteria

There are two levels of area stratification in the sampling design. The first level is the geographical stratification based on the partition of the total country area into thirteen Regions corresponding to NUTS 2 level. The former two major city agglomerations of Greater Athens and Greater Thessaloniki constitute separate major geographical strata.

The second level of stratification entails grouping municipalities and communes within each Region (NUTS 2) by the degree of urbanization, i.e., according to their population size. The scaling of urbanization was finally designed in four groups:

- a.  $\geq 30.000$  inhabitants
- b. 5.000 - 29.999 inhabitants
- c. 1.000- 4.999 inhabitants
- d. 0 - 999 inhabitants

The number of final strata in the thirteen (13) Regions is 50. The former Greater Athens Area was divided into 31 strata of about equal size (equal number of households) on the basis of the lists of city blocks of the Municipalities that constitute it and taking into consideration socio-economic criteria. Similarly, the former Greater Thessaloniki Area was divided into 9 equally sized strata. The two Major City Agglomerations account for about 40% of total population and for even larger percentages in certain socio-economic variables. Thus, the total number of strata of the survey was 79.

### 2.4 Sample size and allocation criteria

The initial sample size is 8.000 households (sampling fraction  $\cong 2\%$ ). This fraction was the same in each administrative region. The Regions (NUTS 2) in Greece are 13 in number. However, throughout this survey the 2<sup>nd</sup> Region (Central Macedonia) was considered without the former Greater Thessaloniki and the 9<sup>th</sup> Region (Attiki) without the former Greater Athens area, while either of these two major agglomerations were treated as a geographical region.

## 2.5 Sample selection schemes

### **1<sup>st</sup> stage of sampling**

In this stage, from any ultimate stratum (crossing of Region with the degree of urbanization), say stratum ***h***, ***n<sub>h</sub>*** primary units were drawn. The number ***n<sub>h</sub>*** of draws was approximately proportional to the population size ***X<sub>h</sub>*** of the stratum (number of households according to the last population census of the year 2001).

Each area unit (primary unit) of the stratum had a selection probability proportional to its size. So, if ***X<sub>hi</sub>*** was the number of households according to the 2001 general population census- of the unit in the sample of order ***i***, then the probability of being drawn was:

$$P_{hi} = \frac{X_{hi}}{X_h}$$

The total number of the primary sampling units is 1.056 areas.

As in each year the 25% of the sample households is replaced, the new households belong to different primary sampling units.

### **2<sup>nd</sup> stage of sampling**

In this stage, from each primary sampling unit, the sample of ultimate units (households) was selected. Actually, in the second stage, we drew a sample of dwellings. However, in most cases, there is one to one correspondence between household and dwelling. If the selected dwelling constitutes of one or more households then all of them are interviewed.

### 3. Statistical unit(s), scope and target population

<b>3.1 Statistical unit</b>	Please indicate whether the statistical unit follows recommendations by ticking <i>Yes</i> or <i>No</i> (and specify the deviations, if any):	
	<b>Yes</b>	<b>No</b> (please specify the deviations)
Module A in the Eurostat model questionnaire: households with at least one member aged 16 to 74	<b>X</b>	
Modules B to E in the Eurostat model questionnaire: individuals aged 16 to 74	<b>X</b>	

<b>3.2 Age groups covered</b>	Please indicate the age scope (in the <i>Yes</i> column), or tick <i>No</i> if not applicable	
	<b>Yes</b> (please specify, e.g. 12-15 or 75-...)	<b>No</b>
Individuals younger than 16 ?		<b>X</b>
Individuals aged 16 to 74 ?	(compulsory)	
Individuals older than 74 ?		<b>X</b>

3.3	<b>Territorial coverage</b>
	If applicable, indicate the parts of the country that are not included as well as an estimate of the resulting percentage of undercoverage (non-covered population compared to total country population).
	All private households of the country and the members of them are covered in the survey, independently of their size or any socio-economic characteristics they may have.
	Excluded are collective households such as hotels, hospitals, military camps, nursing homes, etc. As collective households were also considered households with more than 5 lodgers. Households having as members foreigners in diplomatic missions.

<b>Universe</b>	<b>Households</b>	<b>Individuals</b>
<b>3.4 Target population</b> The number of <i>households</i> and <i>individuals</i> in the target population (scope, universe). Please restrict the numbers to the <i>Eurostat scope</i> (if additional age groups are covered in the national survey, these can be reported separately between brackets). If not directly available, please provide an estimate (e.g. based on other social surveys). If not applicable, please indicate why.	<b>3.737.938</b>	<b>8.274.189</b>
<b>3.5 Non-target population</b> The approximate number of <i>households</i> and <i>individuals</i> outside the general scope of the survey (e.g. <u>individuals younger than 16 or older than 74; households with all members over 74 years old</u> ), i.e. the difference between the total population (in terms of households or individuals) in the country and the target population). If not applicable, please indicate why.	<b>376.212</b>	<b>2.635.867</b>

## 4. Questionnaire

<b>4.1</b>	<b>Adoption of <i>MANDATORY</i> questions and items from the Eurostat model questionnaire (2010, v3.2)</b> <p>The questions listed below and its items reflect the required coverage of subjects and characteristics of Annex 2 of the Commission Regulation (EC) No 1023/2009 of 29 October 2009. Please indicate in the table possible comments on the question/item coverage in your national questionnaire, e.g. insertion of additional items, different reporting periods, deviations in the routing of ordering of the questions and/or items (see also §4.4), differences in definitions or classifications, alternative sources used (esp. in the background characteristics).</p> <p>Where applicable, please report on the coverage of the questions for age groups beyond the standard scope, i.e. for respondents younger than 16 or for respondents older than 74. Possible comments on this issue can be added to the general column on deviations.</p>			
	Question	Any deviations from question / items in Model Questionnaire	Covered for other age groups?	
			<16	>74
	<b>Module A : Access to selected ICTs</b>		→ if not applicable, please leave blank	
A1	Does the household via one of its members has access to a computer at home?	X		
A2	Does any member of this household have access to the Internet at home?	X		
A3	On which of these devices is the Internet accessed at home?	X		
A4	What types of Internet connection are used?	X		
A5	What are the reasons for not having access to the Internet at home?	X		
	<b>Module B : Use of computer</b>		→ if not applicable, please leave blank	
B1	When did you most recently use a computer?	X		
B2	How often on average have you used a computer in the last 3 months?	X		
	<b>Module C : Use of the Internet</b>		→ if not applicable, please leave blank	
C1	When did you last use the Internet?	X		
C2	On average how often did you use the Internet in the last 3 months?	X		
C3	Where have you used the Internet in the last 3 months?	X		
C4	Do you use any of the following mobile devices to access the Internet?	X		
C5	For which of the following activities did you use the Internet in the last 3 months for private use?	X		
C6	For which of the following activities relating to interaction with public services or administrations and during which period did you use the Internet for private purpose?	X		
	<b>Module D : E-Commerce and Trust</b>		→ if not applicable, please leave blank	

	Question	Any deviations from question / items in Model Questionnaire	Covered for other age groups?	
			<16	>74
D1	When did you last buy or order goods or services for private use over the Internet (excluding manually typed e-mails)?	X		
D2	What types of goods or services did you buy or order over the Internet for private use in the last 12 months?	X		
D3	Were any of the following products that you bought or ordered over the internet downloaded or accessed from websites rather than delivered by post etc.?	X		
D4	From whom did you buy or order goods or services for private purpose over the Internet in the last 12 months?	X		
D5	Did you place a bet (e.g. sports betting) or play gambling or lotto over the internet in the last 12 months?	X		
	<b>Module E : Internet security</b>		→ if not applicable, please leave blank	
E1	How concerned are you about the following possible problems related to Internet usage for private purposes?	X		
E2	Have you experienced any of the following security related problems through using the Internet for private purposes in the last 12 months?	X		
E3	Have security concerns kept you from doing the following activities via the Internet for your own private purposes in the last 12 months?	X		
E4	Do you use any kind of IT security software or tool (anti-virus, anti-spam, firewall etc.) in order to protect your private computer and data?	X		
E5	Which IT security software or tool do you use?	X		
E6	Do you update one or more of your security products (virus checking, spyware program etc.)?	X		
E7	Why you do not update your security products?	X		
E8	How often do you make safety copies or back up files from your computer on any external storage device or to disk space on Internet servers?	X		
	<b>Module F : E-Skills</b>		→ if not applicable, please leave blank	
F1	Which of the following internet related activities have you already carried out?	X		
	<b>Socio-demographic background variables</b>		→ if not applicable, please leave blank	
G1	Age	X		
G2	Sex	X		
G5	Educational level (according to ISCED)	X		
G6	Employment situation	X		

	Question	Any deviations from question / items in Model Questionnaire	Covered for other age groups?	
			<16	>74
G7	Occupation (according to ISCO)	X		
G8	Region of Residence, NUTS 1	X		
G10	Geographical location (former y/n Objective 1 : from 2007 corresponding to « Convergence region « and all others, i.e. « Regional Competiveness and Employment region «)	X		
G11	Type of locality (degree of urbanisation)	X		
G12	Number of members in the household	X		
G13	of which, number of children under 16	X		

<b>4.2 Adoption of <i>OPTIONAL</i> questions and items from the Eurostat model questionnaire (2010, v3.2)</b> Please indicate in the table below if and which <u>optional</u> variables and questions were included in the national questionnaire. For each question or item, an "x" in the column named Question included means that it was included in the national questionnaire. Where applicable, please report also on the coverage of the questions for age groups beyond the standard scope, i.e. for respondents younger than 16 or for respondents older than 74.					
	Question / Item	Question included ?	Any deviations from question / items in Model Questionnaire	Covered for other age groups?	
				<16	>74
A3c1	Via Internet enabled mobile phone (GPRS, UMTS, etc.)	X			
A3c2	Via handheld computer (palmtop, PDA)	X			
A4c4	Mobile broadband via 3G handset				
A4c5	Mobile broadband via 3G modem				
B3	Where have you used a computer in the last 3 months?				
B4	When did you last take a training course (of at least 3 hours) on any aspect of computer use ?	X			
C3e1	Public library				
C3e2	Post office				
C3e3	Public office, town hall, government agency				
C3e4	Community or voluntary organisation				
C3e5	Internet café	X			
C3e6	Hotspot	X			
C5n	Selling of goods or services	X			
G3	Country of birth	X			
G4	Country of citizenship	X			
G9	Region of Residence, NUTS 2				
G14	Household income	X			

4.3	<b>Additional questions introduced in the national questionnaire, if any</b>
	An additional question on “green” habits has been introduced for testing reasons, as well as a question on tele-working for the respondent.
4.4	<b>Effects of deviations from the routing used in the Eurostat model questionnaire, if any</b>

## 5. Sampling frame

<b>5.1</b>	<b>Name and short description of the sampling frame or register used</b>
	<p>The sample of households for the ICT survey of the year <b>2010</b> was consisted of a sub-sample of the samples used in the National Health Survey of the year 2009 and EU-SILC of the years 2005-2008.</p> <p>The multi-stage stratified area sampling was adopted for the survey. The primary sampling units are the areas (one or more unified city blocks) participating in the National Health Survey of the year 2009 and the EU-SILC of the years 2005-2008.</p> <p>The secondary sampling units are the sampling households of the National Health Survey and the EU-SILC containing individuals aged 15-74 years old and belonging to the selected primary sampling units.</p> <p>The final sampling unit is one person randomly selected among the household members aged sixteen to seventy four years.</p>
<b>5.2</b>	<b>Known shortcomings of the sampling frame, if any</b>
	<p>Shortcomings in terms of timeliness (e.g. time lag between last update of the sampling frame and the moment of the actual sampling), geographical coverage, coverage of different subpopulations, etc.</p> <p>Only the private households participate in the survey. Individuals who permanently reside in collective houses (as hospitals, hotels, asylums, houses of old people, orphanages etc) are not covered by the survey. These individuals are - as a rule - members of institutional households. If however we subtract from this population the conscripts and the imprisoned, the actual percentage not covered by the survey procedure, accounts for 2% of the total population, and in its major part concerns economically non-active persons.</p>

## 6. Sampling design

### 6.1 Sampling method

Please give a description of the sampling method used (e.g. stratified random sample, quota sampling, cluster sampling; one-stage or two-stage sampling; if not directly selected from the register, how are individuals selected within the household; one or all individuals within a household; etc.) and the method used for determining the sample size and sample selection. If stratification was used, please specify which variables were used to stratify, the categories of those variables and the final number of strata.

The sample of households for the ICT survey of the year **2010** was consisted of a sub-sample of the samples used in the National Health Survey of the year 2009 and EU-SILC of the years 2005-2008.

The multi-stage area sampling was adopted for the survey. The primary sampling units are the areas (one or more unified city blocks) participating in both the National Health Survey of the year 2009 and the EU-SILC of the years 2005-2008. The secondary sampling units are the households of the National Health Survey and the EU-SILC containing members belonging to the target population (individuals aged 15-74 years old). The final sampling unit is one person randomly selected among the household members of age sixteen to seventy four years.

The multi-stage stratified sampling method was applied with stratification variables (a) the Region (NUTS 2) and (b) the degree of urbanization.

In each Region (NUTS 2), the stratification of primary units was conducted by allocating the Municipalities and Communes according to the degree of urbanization (urban, semi-urban, and rural regions). Except for the former two Major City Agglomerations (Athens and Thessaloniki), the created strata according to the degree of urbanization are:

Urban	Agglomerations and Municipalities with 10.000 inhabitants or more
Semi-urban	Municipalities and Communes with 2.000 to 9.999 inhabitants
Rural	Communes up to 1.999 inhabitants

The former Greater Athens Area was divided into 31 strata of about equal size (equal number of households) on the basis of the lists of city blocks of the Municipalities that constitute it and taking into consideration socio-economic criteria. Similarly, the former Greater Thessaloniki Area was divided into 9 equally sized strata. The two Major City Agglomerations account for 40% of total population and for even larger percentages in certain socio-economic variables.

The three-stage area sampling method was adopted for the survey with (a) primary units the areas (one or more unified city blocks), (b) secondary sampling units the households and (c) ultimate sampling units the individuals belonging to the target population.

The sampling fraction in each of the 79 strata (stratum= Region x degree of urbanization) is equal to  $f = \frac{1}{\lambda} = \frac{n}{N} \cong 1,6 \text{ ‰}$ , where  $n=6.000$  is the total sample size of households and  $N=3.737.938$  is the total number of households belonging to the target population. The quantity  $N$  is a projection, based on the time trend of the Labour Force Survey data of the period 2001-2009.

The number of the sampling households in each of the 79 strata (let  $h$ ) was defined by applying the proportional allocation as follows:

$$n_h = n \cdot \frac{N_h}{N},$$

where:

$N_h$ : the population size of the stratum  $h$ , being estimated from data of the Labour Force Survey with reference period the 2<sup>nd</sup> quarter of the year 2009.

### Stages of probability sampling

The sample of households for the ICT survey of the year 2010 was selected of the sample used in the National Health Survey and the EU-SILC. Although the sample of households for the ICT survey of the year 2010 was selected of the sample used in the National Health Survey and the EU-SILC, the following measures were taken for improving the selection probabilities, so that the requirements of the ICT survey to be met. The selection probabilities were defined as following:

#### 1<sup>st</sup> stage of sampling

In this stage, for any ultimate stratum (crossing of NUTS 2 Regions with the degree of urbanization), say stratum  $h$ ,  $a_h$  primary units were drawn with probabilities proportional to their sizes. The number  $a_h$  of draws is approximately proportional to the population stratum size  $N_h$ , as defined above.

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

$$P_{hi} = \frac{N_{hi}}{N_h} \quad (1)$$

where:

$N_{hi}$ : the updated (from the National Health Survey and the EU-SILC) target population of households in the  $hi$  primary unit.

#### 2<sup>nd</sup> stage of sampling

In the  $hi$  primary unit, a sample of  $n_{hi}$  out of  $N_{hi}$  households was selected with equal probabilities. Each one of the  $n_{hi}$  households had the same chance to be selected, equal to:

$$\frac{n_{hi}}{N_{hi}} \quad (2).$$

The total number of households to be interviewed of the  $a_h$  sampling primary units is:

$$n_h = \sum_{i=1}^{a_h} n_{hi}$$

Within each primary sampling unit the calculation of the sampling interval  $\delta_{hi} = \frac{N_{hi}}{n_{hi}}$  was carried out, so that the following two desired conditions to be satisfied.

- a. The expectation of the fraction  $\frac{n_h}{N_h}$  should be the predetermined above sampling fraction

$$\frac{1}{\lambda} \cong 1,6\% \text{ in each stratum: } E\left(\frac{n_h}{N_h}\right) = \frac{1}{\lambda} \quad (3)$$

- b. The estimator of the stratum total  $Y_h$  (for any characteristic) should be self-weighting. In other words, the estimation should be derived as product of the sum of the values of the characteristic over the  $n_h$  sample households by the overall raising factor  $\lambda$ , which is equal in each stratum.

The conditions (a) and (b) are satisfied when:

$$\frac{1}{a_h} \cdot \frac{1}{P_{hi}} \cdot \frac{N_{hi}}{n_{hi}} = \lambda \quad (4) \Rightarrow$$

$$\frac{1}{a_h} \cdot \frac{1}{P_{hi}} \cdot \delta_{hi} = \lambda \Rightarrow$$

$$\delta_{hi} = \frac{N_{hi}}{n_{hi}} = \lambda \cdot a_h \cdot P_{hi} \quad (5)$$

From the relations (1) and (5)  $\Rightarrow \frac{N_{hi}}{n_{hi}} = \lambda \cdot a_h \cdot \frac{N_{hi}}{N_h} \Rightarrow n_{hi} = \frac{N_{hi} \cdot N_h}{\lambda \cdot a_h \cdot N_{hi}} \Rightarrow$

$$n_{hi} = \frac{N_h}{\lambda \cdot a_h} \quad (6)$$

From relation (3), it is deduced that:  $\frac{1}{\lambda} = \frac{n_h}{N_h} \Rightarrow \lambda = \frac{N_h}{n_h} \quad (7)$

From relations (6) and (7), we have:  $n_{hi} = \frac{n_h}{a_h} \quad (8)$

### 3<sup>rd</sup> stage of sampling

In this stage from each household one individual (member of household belonging to the target population) was selected with equal probabilities. Let  $p_{hijk}$  be the selection probability of the  $hijk$  individual, which belongs to the  $hij$  household of the PSU of order  $i$  belonging to the stratum  $h$ . As one individual was selected with equal probabilities out of  $m_{hij}$  members

belonging to target population, the  $p_{hijk}$  was defined as:  $p_{hijk} = \frac{1}{m_{hij}}$

## 6.2 Additional measures taken at the time of sampling design to improve representativeness

If any, and if not covered under §6.1. E.g. corrections for sampling frame under-coverage, etc.

The initial selection probabilities of the sampling households of both National Health Survey and EU-SILC were based on the population sizes (from the National General Population Census of the year 2001), which differ considerably from the new population sizes that better suit the demands of the current ICT survey. Additionally, the target populations of National Health Survey, EU-SILC and ICT do not coincide. The measures of National Health Survey and EU-SILC were based on all persons, but the current sample for ICT was restricted to the households with individuals aged 16 to 74 years old. Thus, although the sample of households for the ICT survey of the year 2010 was selected of the sample used in the National Health Survey of the year 2009 and the National Survey of Income and Living Conditions (EU-SILC of the years 2005-2008), the following measures were taken for improving the representativeness:

- The 1st stage selection probabilities of the primary units were modified taking into consideration the updated target population size in each stratum using estimated data from the Labour Force Survey with reference period the 2nd quarter of the year 2009
- The 2nd stage selection probabilities of the households were modified taking into consideration the updated register of the households in the primary sampling units.
- The allocation of sampling households in each separate stratum was carried out proportionally to the target population size, which was estimated from data coming from Labour Force Survey with reference period the 2nd quarter of the year 2009.

After the application of the above measures, the sampling households for the ICT have not the same probability of selections (1st and 2nd) with the sampling households of the National Health Survey and EU-SILC, after the modification of the selection probabilities of the National Health Survey and EU-SILC households, so that the probabilities of the ICT households to correspond to the updated target population.

Sample size		Households	Individuals (aged 16 to 74)	Individuals (younger than 16)	Individuals (older than 74)
6.3	<b>Gross sample size</b> The number of households/individuals initially selected from the sampling frame (if not applicable, please indicate why). Please restrict the numbers in the first two columns to the <i>Eurostat scope</i> (if additional age groups are covered, these can be reported separately in the last two columns).	6.000	6.000	→ if not applicable, please leave blank	
6.4	<b>Net sample size</b> The number of households/individuals that can be used in the final database (if not applicable, please indicate why).	To be filled in under §7.C (final report)			

## 7. Response and non-response

(quality report)

**Note:** This chapter only deals with non-response error. Other non-sampling error such as frame errors, measurement and processing errors or model assumption errors are discussed elsewhere or outside the scope of this methodological report.

### UNIT NON-RESPONSE

Unit non-response occurs when not all elements (households and/or individuals) of the gross sample (i.e. the initial sample drawn from the reference sampling frame) participate in the survey and are thus not included in the net sample.

However, not all types of non-response are taken into account when calculating the response rate (in §7.D) as they can be rather related to the quality of e.g. the sampling frame than to the quality of the survey data.

Note: In this report - for reasons of comparability across countries - all non-contacts are considered to be *non-response of eligible cases* (where in reality some of the non-contacts may concern ineligible cases).

**If additional age groups were covered, please report separately for individuals in the general scope (16-74), and any additional age groups covered (see the last two columns).**

**If no additional age groups were covered (see also §3.2 and §4.1), the last two columns can be left blank.**

		Number of households	Number of individuals (aged 16-74)	Number of individuals (<16)	Number of individuals (>74)
<b>7.A</b>	<b>Gross sample size</b> The number of households/individuals initially selected from the sampling frame (if not applicable, please indicate why).	6000	6000	→ if not applicable, please leave blank	

	Type of unit non-response (ineligible cases)	Number of households	Number of individuals (aged 16-74)	Number of individuals (<16)	Number of individuals (>74)
<b>7.1</b>	<b>Ineligible: out-of-scope</b> E.g. selected household is not in the target population because all members are over 75 years old.	17	17	→ if not applicable, please leave blank	
<b>7.2</b>	<b>Other ineligible</b> E.g. no dwelling exists at the selected address or selected individual has died between the reference data of the sampling frame (cf. §5.2) and the moment of the interview.	-	-		

<b>7.B</b>	<b>Number of eligible elements</b> I.e. the gross sample size corrected for the ineligible cases. ► $[\$7.B] = [\$7.A] - [\$7.1] - [\$7.2]$	5983	5983		
------------	---	------	------	--	--

	Type of unit non-response (eligible cases)	Number of households	Number of individuals (aged 16-74)	Number of individuals (<16)	Number of individuals (>74)
<b>7.3</b>	<b>Non-contact</b> E.g. no one was home or postal survey was never sent back.	925	925	→ if not applicable, please leave blank	

<b>7.4</b>	<b>Refusal</b> E.g. selected household or individual was contacted but refused to take part in the survey.	<b>448</b>	<b>448</b>		
<b>7.5</b>	<b>Inability to respond</b> E.g. selected household or individual was unable to participate due to language barriers or cognitive or physical incapacity to respond.	-	-		
<b>7.6</b>	<b>Rejected interviews</b> E.g. the selected household/individual did take part but the survey form cannot be used (poor quality - e.g. strong inconsistencies; unacceptable item-response - e.g. individual left most of the questions unanswered; survey form got lost and interview cannot be repeated; etc.).	-	-		
<b>7.7</b>	<b>Other non-response</b> Please specify the other types of non-response encountered. <u>Note:</u> please add the other non-response related to ineligibility of the selected elements under §7.2. <ul style="list-style-type: none"><li>• ...</li><li>• ...</li></ul>	-	-		

<b>7.C</b>	<b>Net sample size</b> The number of households/individuals that can be used in the final database (if not applicable, please indicate why). This notion corresponds to the <i>final sample</i> in the Tabulation Scheme. ► $[\$7.C] = [\$7.B] - [\$7.3] - [\$7.4] - [\$7.5] - [\$7.6] - [\$7.7]$	<b>4610</b>	<b>4610</b>		
------------	---	-------------	-------------	--	--

		<b>Households</b>	<b>Individuals (aged 16-74)</b>	<b>Individuals (&lt;16)</b>	<b>Individuals (&gt;74)</b>
<b>7.D</b>	<b>Unit response rate</b> The unit response rate is the <b>ratio</b> of the <i>number of in-scope respondents</i> (= the number of achieved interviews or the net sample size, see §7.C) to the <i>number of eligible elements</i> selected from the sampling frame (see §7.B). The number of eligible elements equals <i>the gross sample size</i> (see §7.A) <i>minus the ineligible cases</i> (see §7.1 and §7.2). ► $[\$7.D] = [\$7.C] / [\$7.B]$	<b>77,06</b>	<b>77,06</b>	<b>→ if not applicable, please leave blank</b>	

<b>7.8</b>	<b>Comments on the unit response rate, if any</b>

	<b>Methods used for minimizing unit non-response</b> Where applicable, give a description of measures taken to reduce the unit non-response: <ul style="list-style-type: none"><li>• advance notification in the form of a letter or phone call;</li><li>• system of reminders, number of visits, number of attempts for phone calls, etc.</li><li>• showing respondents how the data they are providing are being used;</li><li>• etc.</li></ul>
--	---

7.9	<p><b>Methods used for minimizing unit non-response</b></p> <p>Where applicable, give a description of measures taken to reduce the unit non-response:</p> <ul style="list-style-type: none"> <li>• advance notification in the form of a letter or phone call;</li> <li>• system of reminders, number of visits, number of attempts for phone calls, etc.</li> <li>• showing respondents how the data they are providing are being used;</li> <li>• etc.</li> </ul>
7.9 7.10	<p>An advanced notification letter was sent to all households, one month before the survey conduction, among others providing information on how the data collected are being used.</p> <p><b>Methods used for dealing with unit non-response</b></p> <p>Indicate whether imputations are made for unit non-response and give a short description of the methods used (e.g. correction factor in the weighting procedure, imputation based on background characteristics known from the sampling frame, etc.).</p> <p><b>Correction factor in the weighting procedure:</b> In each stratum <math>h</math>, the inverse of the initial inclusion probabilities of the respondents are multiplied by the inverse of the response rate <math>r_h</math> of the stratum <math>h</math>. Consequently, the extrapolation factor takes into account not only the selection probability of the sampling units but also the inverse of response rate in each stratum.</p>
7.11	<p><b>Proxy answers</b></p> <p>Please indicate whether the instructions to interviewers allow for proxy interviews (another person in the household than the one who was randomly selected can answer the questions).</p> <p>If yes, give an estimate of the percentage of proxy interviews (compared to the total number of interviews).</p> <p>Not-applicable</p>

## ITEM NON-RESPONSE

Item non-response occurs when a respondent provides some, but not all, of the requested information, or if the reported information is not useable (note that entirely non-useable questionnaire are already counted in the *unit* non-response, see §7.6).

It may occur for a variety of reasons. Items may be missing because the respondent broke off the interview after partially completing it (but enough data were provided so that the questionnaire is not classified as a unit non-response). Items may be missing because the respondent inadvertently skipped an item, a module or a page (especially in self-administered mail surveys). Or a respondent may simply not have the information on the question (and no don't know option is foreseen) or refuse to give the requested information.

As item non-response usually goes hand-in-hand with systematic bias (e.g. the proportion of *No* answers may be higher among people with item non-response compared to those who did answer on a specific item), it is useful to assess the degree and impact of this type of non-response.

7.12	<p><b>Questions or items with item response rates below 90%</b></p> <p>If any, identify the items with low response rates (the cut-off value to be used is 0.90) and indicate their respective response rates. The item non-response rate should of course be calculated taking into account the routing and filtering in the questionnaire.</p>
	<p><b>Methods used for dealing with item non-response</b></p> <p>Indicate whether imputations are made for item non-response and give a short description of the methods used (e.g. nearest-neighbour imputation, hot deck imputation, mode imputations within classes, etc.).</p> <p>The data entry program didn't allow for missing items.</p>

**7.14 Other comments relating to the item non-response**

If any, please use this box to inform on additional issues on the non-response calculation (e.g. method used in national publications, etc.).

## 8. Grossing-up procedures

### (quality report)

Please give a description of the extrapolation or weighting procedures used to gross up the *households* (§8.1) and the *individuals* (§8.2) in the net sample to the (target) population, discussing the different steps taken or factors applied to the design weighting to take into account the (post)stratification, balancing for unit non-response, etc. In case similar methods are used for grossing-up the net samples of households and individuals, the discussion can be integrated under one heading.

#### 8.1 Grossing-up procedures for households

Let  $h$  be one of the final strata of households (Final stratum = Region x Degree of Urbanization), then this will take the following values:  $h = 1, 2, \dots, H$  (where  $H = 79$ ). 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  $j$  belonging to the PSU of order  $i$  was defined as:

$$w_{hij} = \frac{N_h}{a_h \cdot N_{hi}} \cdot \frac{N_{hi}}{n_{hi}} \cdot \frac{1}{r_h} t_{hij} = \frac{N_h}{a_h \cdot n_{hi}} \cdot \frac{1}{r_h} t_{hij} \quad (9)$$

From relations (8) in paragraph 6.1 and (9), we have:

$$w_{hij} = \frac{N_h}{a_h \cdot \frac{n_h}{a_h}} \cdot \frac{1}{r_h} t_{hij} \Rightarrow w_{hij} = \frac{N_h}{n_h} \cdot \frac{1}{r_h} t_{hij} \quad (10)$$

where:

$N_h$  : the target population size in the  $h$  stratum according to LFS of the 2<sup>nd</sup> quarter 2009

$n_h$  : the initial sample size in the  $h$  stratum

$\frac{N_h}{n_h}$  : the inverse of the initial inclusion selection probability 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

$t_{hij}$  : Factor, which adjusts the sample weights of households so that the sample 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.

The distribution of households by size class is estimated from EU-SILC, with reference period the year 2009.

## 8.2 Grossing-up procedures for individuals

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  $k$  belonging to the  $hij$  household is defined as follows:

$$w_{hijk} = w_{hij} \cdot \frac{1}{p_{hijk}} \cdot g_{hijk} \quad (8.2)$$

where:

$w_{hijk}$ : The extrapolation factor of the  $hijk$  household in which the  $hijk$  individual belongs

$p_{hijk}$ : The selection probability of the  $hijk$  individual, which belongs to the  $hij$  household. As one individual was selected with equal probabilities out of  $m_{hij}$  members belonging to the target population, the

$p_{hijk}$  is defined as: 
$$p_{hijk} = \frac{1}{m_{hij}}$$

$g_{hijk}$ : Factor, which adjusts the sample weights of individuals, so that the sample 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  $\times$  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 is estimated from data coming from the EU-SILC, with reference period the year 2009.

## 9. Sampling error (quality report)

### Standard error (for a selection of indicators)

The sampling error reflects the fact that only a particular sample was surveyed rather than the entire population. It is estimated by the standard error and can be expressed by the square root of the estimate of the sampling variance ( $\hat{\sigma}_{(\hat{\theta})}$ ). The estimation of the sampling variance should ideally take into account the sampling design (e.g. the stratification).

Please comment on the approach for calculating sampling errors in §9.6. In case the standard errors are derived using the variance formula for simple random sampling and incorporating a factor which reflects the multi-stage, clustered nature of the sampling design, please provide more detailed information in §9.6 d and e.

Please indicate below the number of respondents (absolute value for Yes answers), the estimated value of the proportion (in %) as well as the respective *standard error* (in %) for the indicators and subindicators mentioned.

**Please note that the accuracy measure used, i.e. the STANDARD DEVIATION, was also addressed in the 2006, 2007, 2008 and 2009 report templates but differs from the 2004 and 2005 report templates (where the coefficient of variation was used).**

The section 9.7 should be completed with comments on reliability and representativeness of results and completeness of dataset. The two questions should be left blank if not applicable, i.e. if standard errors found were adequate or if subgroups of the population had always a sufficient number of respondents.

<b>Indicator or subindicator</b>  <b>- on households and individuals in the general scope (16-74) and related subgroups -</b>		<i>Number of respondents</i>	<i>Estimated proportion (%)</i>	<i>Standard error (%)</i>	<i>C.V. (%)</i>
<b>9.1</b>	<b>Proportion of households having access to the Internet at home</b> (item 'Yes' in variable A2 of the 2010 model questionnaire)	<b>1.732.577</b>	<b>46,35</b>	<b>0,80</b>	<b>1,73</b>
<b>9.2</b>	<b>Proportion of households using a broadband connection</b> (a 'Yes' on option b or c in variable A4 of the 2010 model questionnaire)	<b>1.539.321</b>	<b>88,85</b>	<b>0,75</b>	<b>0,84</b>
<b>9.3</b>	<b>Proportion of individuals regularly using the Internet: overall</b> (indiv. who ticked option 1 or 2 in variable C2 of the 2010 model questionnaire)	<b>3.381.170</b>	<b>40,86</b>	<b>0,89</b>	<b>2,18</b>
<b>9.3.1</b>	Proportion of ind. regularly using the Internet: <b>males</b> (as % of all men)	<b>1.857.103</b>	<b>45,52</b>	<b>1,28</b>	<b>2,81</b>
<b>9.3.2</b>	Proportion of ind. regularly using the Internet: <b>females</b> (as % of all women)	<b>1.524.068</b>	<b>36,34</b>	<b>1,18</b>	<b>3,25</b>
<b>9.3.3</b>	Proportion of ind. regularly using the Internet: <b>age group 16-24 years</b> (as % of all individuals aged 16-24 years)	<b>882.333</b>	<b>80,50</b>	<b>2,42</b>	<b>3,00</b>
<b>9.3.4</b>	Proportion of ind, regularly using the Internet: <b>age group 25-34 years</b> (as % of all individuals aged 25-34 years)	<b>1.005.377</b>	<b>62,28</b>	<b>2,26</b>	<b>3,64</b>
<b>9.3.5</b>	Proportion of ind, regularly using the Internet: <b>age group 35-44 years</b> (as % of all individuals aged 35-44 years)	<b>807.862</b>	<b>48,76</b>	<b>1,81</b>	<b>3,72</b>
<b>9.3.6</b>	Proportion of ind, regularly using the Internet: <b>age group 45-54 years</b> (as % of all individuals aged 45-54 years)	<b>482.175</b>	<b>32,22</b>	<b>1,87</b>	<b>5,81</b>
<b>9.3.7</b>	Proportion of ind, regularly using the Internet: <b>age group 55-64 years</b> (as % of all individuals aged 55-64 years)	<b>159.660</b>	<b>12,82</b>	<b>1,29</b>	<b>10,08</b>
<b>9.3.8</b>	Proportion of ind, regularly using the Internet: <b>age group 65-74 years</b> (as % of all individuals aged 65-74 years)	<b>43.763</b>	<b>3,76</b>	<b>0,73</b>	<b>19,35</b>
<b>9.3.9</b>	Proportion of ind, regularly using the Internet: <b>low educational level</b> (as % of all individuals with low education)	<b>474.078</b>	<b>14,39</b>	<b>1,06</b>	<b>7,37</b>
<b>9.3.10</b>	Proportion of ind, regularly using the Internet: <b>medium educational level</b> (as % of all individuals with medium education)	<b>1.556.840</b>	<b>49,92</b>	<b>1,48</b>	<b>2,97</b>
<b>9.3.11</b>	Proportion of ind, regularly using the Internet: <b>high educational level</b> (as % of all individuals with high education)	<b>1.350.252</b>	<b>72,55</b>	<b>1,57</b>	<b>2,16</b>

9,3,12	Proportion of ind, regularly using the Internet: <b>students</b> (as % of all students)	573.988	85,45	2,77	3,24
9,3,13	Proportion of ind, regularly using the Internet: <b>employees or self-employed</b> (as % of all employees or self-employed)	2.129.613	53,23	1,29	2,42
9,3,14	Proportion of ind, regularly using the Internet: <b>unemployed</b> (as % of all unemployed)	327.610	54,00	3,50	6,47
9,3,15	Proportion of ind, regularly using the Internet: <b>retired, other inactive</b> (as % of all retired and other inactive)	349.959	11,69	0,87	7,43
9,4	<b>Proportion of individuals having downloaded official forms in the last 3 months</b> (individuals who ticked item b in variable C6 of the 2010 model questionnaire)	448.265	12,19	0,94	7,69
9,5	<b>Proportion of individuals having ordered goods or services for private use over the internet in the last 3 months</b> (individuals who ticked option 1 in variable D1 of the 2010 model questionnaire)	723.309	18,40	1,10	6,00

## 9.6 Calculation of the standard error

There exist different methods by which the standard error of an estimated proportion can be assessed based on the distribution in the sample. Please, describe below the approach which you have followed. This information will help Eurostat to evaluate the comparability of the standard errors supplied in the previous section by the different statistical institutes participating in the survey.

### a) Name and brief description of the applied estimation approach

Let  $w_{hijk}$  ( $>0$ ) stand for the survey weight attached to the sample individual  $k$  ( $k = 1$ , as one individual is surveyed, in each sampling household) belonging to the sampling household of order  $j$  ( $j = 1, \dots, n_{hi}$ ), belonging to the selected cluster of order  $i$ , of the stratum  $h$ .

#### *Estimation of survey characteristics*

Let  $y_{hijk}$  be the value of the characteristic  $y$  of the ultimate unit (individual) of the household of order  $j$ , belonging to the  $hi$  primary sampling unit (cluster). Moreover,  $Y$  stands for the total population, which is derived by adding the characteristic  $y$  of all ultimate units included in all strata  $h$ . The form of the estimator on the basis of the two-stage design is:

$$\hat{Y}_h = \sum_{h=1}^H \sum_{i=1}^{a_h} \sum_{j=1}^{n_{hi}} w_{hijk} \cdot y_{hijk} \quad (9.6.1)$$

#### **Estimation of a ratio**

Let  $x_{hijk}$  be the value of the characteristic  $x$  of the ultimate unit of the household of order  $j$ , belonging to the  $hi$  primary sampling unit (cluster). Moreover,  $X$  stands for the total population, which is derived by adding the characteristic  $x$  of all ultimate units included in all strata  $h$ . The form of the estimator  $\hat{R}$  on the basis of the two-stage design is:

$$\hat{R} = \frac{\hat{Y}}{\hat{X}} = \frac{\sum_{h=1}^H \sum_{i=1}^{a_h} \sum_{j=1}^{n_{hi}} w_{hijk} \cdot y_{hijk}}{\sum_{h=1}^H \sum_{i=1}^{a_h} \sum_{j=1}^{n_{hi}} w_{hijk} \cdot x_{hijk}} \quad (9.6.2)$$

## b) Basic formula

### Variance estimation

In order to calculate the variance of the estimated characteristics, the following steps should be followed:

- a) For every selected primary sampling unit (cluster)  $i$  of the stratum  $h$ , we calculate the quantity  $T_{hi}$  using the following formula:

$$T_{hi} = a_h \cdot \sum_{j=1}^{n_{hi}} w_{hijk} \cdot y_{hikj} \quad (9.6.3)$$

- b) Since  $T_{hi}$  has been calculated for every primary sampling unit (cluster)  $i$  ( $i = 1, \dots, a_h$ ) of the stratum  $h$ , then  $V(\hat{Y})$  is calculated as (Rao, 1988):

$$V(\hat{Y}) = \sum_{h=1}^H \frac{1}{a_h \cdot (a_h - 1)} \cdot \left[ \sum_{i=1}^{a_h} T_{hi}^2 - \frac{\left( \sum_{i=1}^{a_h} T_{hi} \right)^2}{a_h} \right] \quad (9.6.4)$$

For the estimation of the variance and the coefficient of variation of a ratio  $R = \frac{\hat{Y}}{\hat{X}}$  additional steps should be followed, below:

- a) For every selected primary sampling unit (cluster)  $i$  of the stratum  $h$ , we calculate the quantity  $F_{hi}$  using the following formula:

$$F_{hi} = a_h \cdot \sum_{j=1}^{n_{hi}} w_{hijk} \cdot x_{hijk} \quad (9.6.5)$$

- b) Since  $T_{hi}$  and  $F_{hi}$  have been calculated for every primary sampling unit (cluster)  $i$  ( $i = 1, 2, \dots, a_h$ ) of the stratum  $h$ , then  $V(\hat{X})$  is calculated as:

$$V(\hat{X}) = \sum_{h=1}^H \frac{1}{a_h \cdot (a_h - 1)} \cdot \left[ \sum_{i=1}^{a_h} F_{hi}^2 - \frac{\left( \sum_{i=1}^{a_h} F_{hi} \right)^2}{a_h} \right] \quad (9.6.6)$$

The variance of  $\hat{R}$  can be calculated using the following formula:

$$V(\hat{R}) = \frac{V(\hat{Y}) + \hat{R}^2 \cdot V(\hat{X}) - 2 \cdot \hat{R} \cdot Cov(\hat{Y}, \hat{X})}{\hat{X}^2} \quad (9.6.7)$$

where:

$$Cov(\hat{Y}, \hat{X}) = \sum_{h=1}^H \frac{1}{a_h \cdot (a_h - 1)} \cdot \left[ \sum_{i=1}^{a_h} T_{hi} \cdot F_{hi} - \frac{\left( \sum_{i=1}^{a_h} T_{hi} \right) \cdot \left( \sum_{i=1}^{a_h} F_{hi} \right)}{a_h} \right] \quad (9.6.8)$$

**c) Main reference in the literature**

**Rao, J.N.K (1988).** Variance Estimation in Sample Surveys. In Handbook of Statistics, Vol. 6, (Eds. P.R. Krishnaiah and C.R. Rao). Amsterdam: Elsevier Science, 427-447

**d) How has the stratification been taken into account?**

As a multistage stratified sampling design is applied, the variance estimation procedure pools stratum variance estimates to compute the overall variance estimate. Additionally, for a multistage sample design, in which the first-stage sampling fraction is small and the first-stage sample is drawn with replacement, the variance estimation method depends only on the first stage of the sample design. Thus, the required input for variance estimates includes only first-stage cluster (PSU) and first-stage stratum identification. You do not need to input design information about any additional stages of sampling.

**e) Which strata have been considered?**

The overall variance estimate comes from the sum of stratum variance estimates. These strata were created using as stratification variables (a) the Region (NUTS 2) and (b) the degree of urbanization (see paragraph 6.1)

**9.7**

**Comments on reliability and representativeness of results and completeness of dataset**

These comments should reflect on the standard errors reported for the indicators and subgroups in 9.1 to 9.5 above as well as on the other indicators and breakdowns. The estimated standard error shall not exceed 2% for the overall proportions and shall not exceed 5% for the proportions relating to the different subgroups of the population where these subgroups constitute at least 10% of the total population in the scope of the survey. If problems were found, these could have implications for future surveys (e.g. need to improve sampling design or to increase sample sizes for households or individuals).

**Indicators and breakdowns in sections 9.1 to 9.5 above:**

According to the survey design, the estimated proportions that are greater or equal to 7,5% achieve a sampling error lower or equal to 5%.

**Other indicators and breakdowns:**

**FOR ALL INDICATORS AND BREAKDOWNS:**

**If applicable, please tick:**

**If significant sampling errors were found, were data cells suppressed in the transmitted dataset and flagged as unreliable?**

**Yes**

**No**

✓

**If a low number of respondents were found for one or more groups of population (e.g. <10), were data cells suppressed in the transmitted dataset and flagged as unreliable?**

**Yes**

**No**

✓

## 10. Microdata

10.1	<b>Please indicate below "Yes" or "No" if microdata were/will be transmitted in addition to the aggregated data. If necessary, any comments related to microdata transmission may be added.</b>
	Yes

## 11. Closing remarks

11.1	<b>Problems encountered and lessons to be learnt</b> These comments can relate to methodological issues as well as to the questionnaire itself (item construction, clarity of definitions to interviewers and respondents, routing and filtering, outcome of pre-tests, etc.)
11.2	<b>Other comments, if any</b>

## 12. Annexes

**Note:** Please also provide the annexes in a computer-readable format and in English

12.1	<b>Questionnaire in national language - Yes</b>
12.2	<b>Questionnaire in English (if available) - Yes</b>
12.3	<b>Interviewer instructions (if available) in national language -Yes</b>
12.4	<b>National reports on methodology (if available)</b>
12.5	<b>Analysis of key results, backed up by tables and graphs (if available)</b>
12.6	<b>Other annexes</b> Please give an overview of other annexes (whether or not referred to in the preceding chapters of this report)
	<ul style="list-style-type: none"> <li>• ...</li> <li>• ...</li> </ul>