

Community Survey on ICT usage in households and by individuals 2013

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 31/05/2013), chapters 7 to 10 can be postponed until the Quality report (deadline for submission: not later than 05/11/2013).** 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	Organisation responsible for the survey <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
1.3	Contact person(s) (name, unit, e-mail, phone, fax)	1. CHALKIADAKI MARIA UNIT OF HOUSEHOLD SURVEYS TEL. 0030 -213-135 2896 FAX. 0030 -213-135 2906 E-MAIL: MCHALK@STATISTICS.GR 2. IOANNIS NIKOLAIDIS HEAD OF METHODOLOGY, ANALYSIS & SURVEYS UNIT TEL. 0030 -213-135 2195 FAX. 0030 - 210 4853100 E-MAIL: GIANNIKOL@STATISTICS.GR

1.4	Name of the collection The name of the survey in its original language(s) and in English (e.g. name used in the statistical office's English website).	Survey on the use of information and communication technologies from the households. Έρευνα χρήσης τεχνολογιών πληροφόρησης και επικοινωνίας από τα νοικοκυριά.
1.5	Last update of this report Please indicate the date of the last update of this report, for the case we have several versions (i.e. use different dates for the metadata report and the quality report).	4/11/2013

2. General methodological information

2.1	<p>Reference period(s)</p> <p>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 style="list-style-type: none"> • 31st of March 2013 for educational level completed • Day of the survey conduct for socio demographic characteristics (activity status, employment situation, country of citizenship, legal marital status, etc.), A1-A4, F1, F2, F3, G16. • First three months of 2013 for questions B2, C2, C3, C5. • Last 12 months (April 2012-March 2013) for all D module and questions E2, E3.
2.2	<p>Survey period</p> <p>The beginning and end date – if already known – of the data collection period.</p> <p>Survey has already started since April 2013 and is estimated to finish in mid June 2013.</p>
2.3	<p>Survey vehicle</p> <p>Stand-alone or embedded in another survey. If embedded, give a short description of the survey the ICT modules are inserted in.</p> <p>ICT is a stand-alone survey.</p>
2.4	<p>Survey type</p> <p>Short description of the survey type (face-to-face interview, self-administered mail survey, telephone interview, combination of techniques, other; etc.).</p> <p>If a combination of techniques was applied, please indicate the proportion of each technique related to the total number of achieved interviews.</p> <p>Telephone interview</p>
2.5	<p>Survey participation</p> <p>Please indicate whether the survey is mandatory or voluntary.</p> <p>Participation is mandatory according to Greek law.</p>
2.6	<p>Main methodological differences compared to previous survey(s)</p> <p>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. survey type, 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.</p>

The sample of households for the ICT survey of the year **2013** consisted of the samples used in the European Survey on Income and Living Conditions (EU-SILC) of the years 2008 - 2011. The EU-SILC is an annual rotating household survey covering the target population of the ICT survey. The EU-SILC is a multistage stratified sample survey with primary sampling unit the area (one or more unified blocks) and final unit the household. The sample design of the EU-SILC was based on data coming from the General Population Census of the year 2001.

For the ICT, primary units are the areas and secondary units the households with members that belong to the target population. The final unit is one person randomly selected among the household members 16-74 years old.

The sample design involves two levels of area stratification: (i) the first level is geographical stratification based on the partition of the whole Country into thirteen standard Regions corresponding to the European NUTS 2 level. The two former major city agglomerations of Greater Athens and Greater Thessaloniki constitute separate major geographical strata. (ii) The second level of stratification involves grouping municipal and local communes within each NUTS 2 Region by degree of urbanization, i.e., according to their population size, into four categories. These categories are defined by population size intervals 0-999, 1000-4999, 5000-29999, 30000 and over. The number of final strata in the thirteen Regions was 50. The two former major city agglomerations were further partitioned into 31 and 9 substrata 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.

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

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

3.2 Age groups covered	Please indicate the age scope (in the <i>Yes</i> column), or tick <i>No</i> if not applicable		
	Yes		No
Individuals younger than 16 ?	<input type="checkbox"/>	If ticked, please specify (e.g. 12-15): <div></div>	<input checked="" type="checkbox"/>
Individuals aged 16 to 74 ?	X (compulsory)		
Individuals older than 74 ?	<input type="checkbox"/>	If ticked, please specify (e.g. 75-80): <div></div>	<input checked="" type="checkbox"/>

3.3 Territorial coverage	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.	

Universe	Households	Individuals
3.4 Target population 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.	3,772,908	8,379,864

3.5	Non-target population The approximate number of <i>households</i> and <i>individuals</i> outside the general scope of the survey (e.g. <i>individuals younger than 16 or older than 74; households with all members over 74 years old</i>), 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.	381,990	2,719,294
-----	---	---------	-----------

4. Questionnaire

4.1	Adoption of <i>MANDATORY</i> questions and items from the Eurostat model questionnaire 2013			
<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 1083/2012 of 19 November 2012. 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
	Module A : Access to selected ICTs		→ if not applicable, please leave blank	
A1	Do you or anyone in your household have access to a computer at home?	X		
A2	Do you or anyone in your household have access to the Internet at home?	X		
A3	What types of Internet connection are used at home?	X		
A4	What are the reasons for not having access to the Internet at home?	X		
	Module B : Use of computer		→ if not applicable, please leave blank	
B1	When did you last use a computer?	X		
B2	How often on average have you used a computer in the last 3 months?	X		
	Module C : Use of the Internet		→ 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 (using a computer or any other means)?	X		
C4	Do you use any of the following mobile devices to access the Internet away from home or work?	X		
C5	For which of the following activities did you use the Internet in the last 3 months for private use? (18 items)	X		
	Module D : Use of e-Government			
D1	Did you contact or interact with public authorities or public services over the internet for private purposes in the last 12 months for the following activities?	X		

	Question	Any deviations from question / items in Model Questionnaire	Covered for other age groups?	
			<16	>74
D2	Did you use websites of public authorities or public services in the last 12 months for any of the following?	X		
D3	Have you experienced any of the following problems when using websites of public authorities or public services for private purposes in the last 12 months?	X		
D4	Are you satisfied or dissatisfied with the following aspects of using websites of public authorities or public services in the last 12 months?	X		
D5	Did you contact public authorities or public services using methods other than websites for private purposes in the last 12 months?	X		
D6	What were the reasons for not submitting completed forms to public authorities' websites for private purposes in the last 12 months?	X		
	Module E : Use of e-Commerce		→ if not applicable, please leave blank	
E1	When did you last buy or order goods or services for private use over the Internet (excluding manually typed e-mails)?	X		
E2	What types of goods or services did you buy or order over the Internet for private use in the last 12 months?	X		
E3	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		
E4	From whom did you buy or order goods or services for private purpose over the Internet in the last 12 months?	X		
	Module F : E-Skills		→ if not applicable, please leave blank	
F1	Which of the following Internet related activities have you already carried out?	X		
F3	Do you judge your current computer skills to be sufficient if you would need to take up a new job on the labour market or change your job within a year?	X		
	Socio-demographic background variables		→ if not applicable, please leave blank	
G1	Age	X		
G2	Sex	X		
G3	Country of birth	X		
G4	Country of citizenship	X		
G7	Educational level (according to ISCED)	X		

	Question	Any deviations from question / items in Model Questionnaire	Covered for other age groups?	
			<16	>74
G8	Employment situation – mandatory variables -	X		
G9	Occupation (according to ISCO)	X		
G10	Region of Residence, NUTS 1	X		
G12	Geographical location: «Convergence region» and all others, i.e. « Regional Competitiveness and Employment region »	X		
G13	Degree of urbanisation	X		
G14	Number of members in the household – mandatory variable -	X		
G15	<i>of which</i> , number of children under 16 – mandatory variable -	X		

4.2 Adoption of **OPTIONAL** questions and items from the Eurostat model questionnaire 2013

Please indicate in the table below if and which optional 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
A3a	Broadband connection	<input checked="" type="checkbox"/>			
C3e1	Public library	<input type="checkbox"/>			
C3e2	Post office	<input type="checkbox"/>			
C3e3	Public office, town hall, government agency	<input type="checkbox"/>			
C3e4	Community or voluntary organisation	<input type="checkbox"/>			
C3e5	Internet café	<input checked="" type="checkbox"/>			
C3e6	Hotspot	<input checked="" type="checkbox"/>			
C3e7	Other	<input type="checkbox"/>			
D6i	Lack of or problems with electronic signature or electronic ID/certificate as reason for not submitting completed forms	<input checked="" type="checkbox"/>			
F2	Do you judge your current internet skills to be sufficient?	<input checked="" type="checkbox"/>			
G5	Legal marital status	<input checked="" type="checkbox"/>			
G6	De facto marital status	<input type="checkbox"/>			
G8	Full time employment	<input checked="" type="checkbox"/>			
G8	Part time employment	<input checked="" type="checkbox"/>			
G8	Employee	<input checked="" type="checkbox"/>			

	Question / Item	Question included ?	Any deviations from question / items in Model Questionnaire	Covered for other age groups?	
				<16	>74
G8	Employee, permanent	<input checked="" type="checkbox"/>			
G8	Employee, temporary	<input checked="" type="checkbox"/>			
G8	Self-employed (incl. family workers)	<input checked="" type="checkbox"/>			
G8	Economic sector (10 items)	<input type="checkbox"/>			
G8	Status for other, not in the labour force (5 items)	<input checked="" type="checkbox"/>			
G9	Coverage of <i>all</i> ISCO-08 2-digit categories	<input checked="" type="checkbox"/>			
G11	Region of Residence, NUTS 2	<input type="checkbox"/>			
-	Region of Residence, NUTS 3 for production of new regional breakdowns (rural/intermediate/urban)	<input type="checkbox"/>			
G14	Number of persons aged 16-24	<input checked="" type="checkbox"/>			
G14	Number of students aged 16-24	<input checked="" type="checkbox"/>			
G14	Number of persons aged 25-64	<input checked="" type="checkbox"/>			
G14	Number of persons aged more than or equal 65	<input checked="" type="checkbox"/>			
G14	Number of persons aged from 14 to 15	<input checked="" type="checkbox"/>			
G14	Number of persons aged from 5 to 13	<input checked="" type="checkbox"/>			
G14	Number of persons aged less than or equal to 4	<input checked="" type="checkbox"/>			
G16	Household income in quartiles	<input checked="" type="checkbox"/>			
G16	Household income (equivalised) in quintiles	<input type="checkbox"/>			

4.3	Additional questions introduced in the national questionnaire, if any

4.4	Effects of deviations from the routing used in the Eurostat model questionnaire, if any

5. Sampling frame

5.1	Name and short description of the sampling frame or register used		
<p>Please mention the frame population and the units listed therein (e.g. districts, municipalities, addresses, households, persons, telephone numbers, etc.).</p> <p>If the sample is selected from a sample of another survey, from a micro-census or from a master sample (in the case of multi-phase sampling designs), then please mention the frame population used for the other survey/the micro-census/the master sample.</p> <p>If more than one sampling frame are used e.g. one sampling frame for each sampling stage or one sampling frame for each national region, then please mention all of them.</p> <p>Please describe if different frames are used to draw the sample and to gross up.</p> <p>Please mention if RDD (Random Digit Dialling) is used.</p>			
<p>The sample of households for the ICT survey of the year 2013 consisted of a sub-sample of the sample that was used in EU-SILC survey of the years 2008-2011.</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 EU-SILC survey of the years 2008-2011.</p> <p>The secondary sampling units are the households of the sample of the EU-SILC survey with individuals aged 16-74 years old that belong to the selected primary sampling units.</p> <p>The final sampling unit is one person randomly selected among the household members 16-74 years old.</p>			
5.2	Is the sample drawn from another survey sample, master sample or micro-census?	Yes	No
<p>If yes, please, name the survey: EU-SILC 2008-2011</p> <p>If yes, then the sampling stages used to select the other survey sample have to be further included in the description of sampling design.</p> <p>If yes, then we have a case of multi-phase sampling.</p>		<input checked="" type="checkbox"/>	<input type="checkbox"/>
5.3	Known shortcomings of the sampling frame, if any		
<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 1.28% of the total population, and in its major part concerns economically non-active persons.</p>			

6. Sampling design

6.1 Is the sampling design a probability design?

Yes



No



A probability sampling design ensures known probabilities for units selected. In practice, non-response generally makes samples depart from the probability ones. However, the point here is to report on whether or not the gross sample (net sample plus non-respondents) has been selected in a probability way.

The sample of households for the ICT survey of the year **2013** consisted of a sub-sample of the samples used in the EU-SILC of the years 2008 - 2011.

The sample of households for the ICT survey of the year **2013** consisted of a sub-sample of the samples used in the EU-SILC of the years 2008 - 2011.

The sample of households for the ICT survey of the year **2013** is consisted of the samples used in the European Survey on Income and Living Conditions (EU-SILC) of the years 2008 - 2011. The EU-SILC is an annual rotating household survey covering the target population of the ICT survey. The EU-SILC is a multistage stratified sample survey with primary sampling unit the area (one or more unified blocks) and final unit the household. The samples design of the EU-SILC was based on data coming from the General Population Census of the year 2001.

For the ICT primary units are the areas and secondary units the households containing members belonging to the target population. The final unit is one person randomly selected among the household members of 16-74 years old.

The sampling design involves two levels of area stratification: (i) the first level is geographical stratification based on the partition of the total country into thirteen standard Regions corresponding to the European NUTS 2 level. The two major city agglomerations of Greater Athens and Greater Thessaloniki constitute separate major geographical strata. (ii) The second level of stratification involves grouping municipalities and communes within each NUTS 2 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 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.

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

Sample selection schemes

1st stage of sampling

In this stage, from any ultimate stratum (crossing of Region with the degree of urbanization), say stratum ***h***, ***n_h*** primary units were drawn. The number ***n_h*** of draws was approximately proportional to the population size ***N_h*** of the stratum (number of households according to the General Population Census of the year 2001).

Each area unit (primary unit) of the stratum had a selection probability proportional to its size. So, if ***N_{hi}*** is 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 is:

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

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.

2nd 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 consists of one or more households then all of them are interviewed.

6.2 What is the number of sampling stages?

If the survey sample is selected from a sample of another survey, from the micro-census or from the master sample, then please include the number of sampling stages from all sampling phases into the total number of sampling stages.

If there are differences in the same country with regard to the number of sampling stages for different population groups, e.g. one-stage sampling in urban areas and two-stage sampling in rural areas, then report the number of sampling stages for each of the population groups.

Two Stages for households, three stages for individuals.

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 EU-SILC of the years 2008 - 2011, while the secondary sampling units are the households of the EU-SILC containing members belonging to the target population (individuals aged 16-74 years old). The final sampling unit is one person randomly selected among the household members 16-74 years old.

6.3 Is there (explicit) stratification at the first stage?

If there are differences as regards stratification at stage 1 between population groups (e.g. rural/urban, etc.), then please answer separately for each case.

Yes



No



6.4 What are the stratification variables at the first stage?

Examples:

- region/ province/ county/ district/ code of administrative territories;
- size/ population density/ degree of urbanisation;
- type of municipality/ settlement;
- type of residence: urban/ rural;
- age, gender, etc.

The stratification variables of the survey are the:

- a. Region (NUTS 2)
- b. Degree of urbanization.

In each Region (NUTS 2), the stratification of primary units was conducted by allocating the Municipal and Local Communities according to the degree of urbanization. Except for the former two Major City Agglomerations (Athens and Thessaloniki), the created strata according to the degree of urbanization are:

1	Municipal Communities with 30.000 inhabitants or more
2	Municipal Communities with 5.000 to 29999 inhabitants
3	Municipal or Local Communities with 1.000 to 4.999 inhabitants
4	Local Communities up to 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. Thus, the total number of strata of the survey was 90, that is the 'final strata'. The two former Major City Agglomerations account for about 38% of total population and for even larger percentages in certain socio-economic variables.

6.5 What is the sampling method at the first stage?

The sampling method (for the sampling units) refers to the way the sample is selected. For example, the sampling method can be a simple random sampling, whereby all samples are given the same probability of selection. Other possible methods include systematic sampling with equal or unequal probabilities, other proportional-to-size sampling (πps), etc.

Please mention if the systematic sampling has stratification effect (gives rise to implicit stratification). If there are differences as regards the sampling design at stage 1 between population groups (e.g. rural/urban, etc.), then please answer separately for each case.

- ☐ Exhaustive selection
- ☐ Simple random sampling
- ☐ Systematic sampling with equal probabilities
 - ☐ With stratification effect,
please mention the related auxiliary variable:
 - ☐ Without stratification effect
- ☒ Systematic sampling with probabilities proportional-to-size
 - ☒ With stratification effect,
please mention the related auxiliary variable:
 - ☐ Without stratification effect
- ☐ Other proportional-to-size (πps) sampling,
please indicate:
- ☐ Other,
please indicate:

For stage 1 it is important to know if there are self-representing primary sampling units (with probability of selection equal to 1). Please mention if this is the case:

6.6	<p>What is the sampling unit at the first stage (PSU)?</p> <p>Examples: census enumeration areas, sections, municipalities, communes, villages, settlements, households, individuals, etc.</p> <p>If there are differences in the same country with regard to the type of primary sampling units (PSUs), e.g. households as PSUs in urban area and villages as PSUs (and households as SSUs) in rural areas, then report the relevant sampling unit at stage 1 for each of the population groups. Please do this also for the sampling units at further stages at the next questions.</p> <p>The primary sampling units are the areas (one or more unified city blocks) participating in the EU-SILC of the years 2008 - 2011.</p> <p>1st stage of sampling</p> <p>In this stage, for any ultimate stratum ('final strata'), 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.</p> <p>The primary unit of order i in stratum h has probability of being drawn proportional to the population size as follows:</p> $P_{hi} = \frac{N_{hi}}{N_h} \quad (1)$ <p>where:</p> <p>N_{hi}: the updated (from the the EU-SILC) target population of households in the hi primary unit</p>		
6.7	<p>Are the sampling units at the first stage self-representing?</p> <p>In case of self-representing primary sampling units i.e. of PSUs purposefully included in the sample (selected with probability of one), these are treated as primary strata and their secondary sampling units are treated as primary sampling units.</p>	<p>Yes</p> <p><input type="checkbox"/></p>	<p>No</p> <p><input checked="" type="checkbox"/></p>
6.8	<p>If yes to 6.7</p> <p>What is the sampling unit at the second stage (SSU)?</p> <p>Examples: dwellings, households, individuals, etc.</p> <p>In case of self-representing primary sampling units, the secondary sampling units are treated as primary sampling units.</p>		

6.9 What is the sampling unit at the ultimate stage?

Examples: dwellings, households, individuals, etc.

The secondary sampling units are the households of the EU-SILC containing members belonging to the target population (individuals aged 16-74 years old). The third –and final– sampling unit is one person randomly selected among the household members of 16-74 years old.

In more detail:

2nd 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 $\frac{n_h}{N_h}$ was constant in each stratum. That is: $E\left(\frac{n_h}{N_h}\right) = \frac{1}{\lambda} = 1.72\%$ (3) and

b) The estimator of the stratum total Y_h (for any characteristic) will be self-weighting. In other words, the estimate of the survey characteristics is derived as product of the sum of the values of the characteristics over the n_h sample households by the overall raising factor λ , 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 the relation (3), it is deducted that: $\frac{1}{\lambda} = \frac{n_h}{N_h} \Rightarrow \lambda = \frac{N_h}{n_h} \quad (7)$

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

3rd 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_{hij} is 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} was defined as: $p_{hij} = \frac{1}{m_{hi}}$

6.10	How many individuals are interviewed in the household? Interviewed units are units from which data are collected. The interviewed unit can be different from the ultimate sampling unit. For instance, the sampling unit at an ultimate stage can be a household and the interviewed unit can be an individual (all eligible individuals in the household are interviewed — this is a cluster sampling). Furthermore, the sampling unit at the ultimate stage can also be an individual and the interviewed unit can be all eligible individuals in the same household. This is an indirect cluster sampling.	One or some <input checked="" type="checkbox"/>	All <input type="checkbox"/>
6.11	Does the survey have a longitudinal component? The survey collects data from the same sample elements on multiple occasions over time.	Yes <input type="checkbox"/>	No <input checked="" type="checkbox"/>
6.12	If yes to 6.11, please provide additional information! If the survey sample is based on rotation groups, please specify the number of rotation groups. Are the rotation groups of equal sizes? What is the frequency of rotation of groups? How are new rotation groups selected? If the survey sample is drawn from another survey sample/micro-census/master sample, then does the rotation take place at the level of the other survey sample/micro-census/master sample?		
6.13	Please provide any additional information on the sampling design. The sampling fraction in each of the 90 strata (Stratum= Region x Degree of urbanization) is equal to $f = \frac{1}{\lambda} = \frac{n}{N} \cong 1.72\%$, where $n=6,500$ is the total sample size of households and $N=3,773,826$ is the total number of households belonging to the target population, for the 2 nd quarter of the year 2012, according to LFS. The number of the sampling households in each of the 90 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		

<i>Sample size</i>		<i>Households</i>	<i>Individuals</i> (aged 16 to 74)	<i>Individuals</i> (younger than 16)	<i>Individuals</i> (older than 74)
6.14	Gross sample size 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).	6,500	6,500	→ if not applicable, please leave blank	
6.15	Net sample size 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)
7.A	Gross sample size The number of households/individuals initially selected from the sampling frame (if not applicable, please indicate why).	6500	6500	→ if not applicable, please leave as is	
				-1	-1

	Type of unit non-response (ineligible cases)	Number of households	Number of individuals (aged 16-74)	Number of individuals (<16)	Number of individuals (>74)
7.1	Ineligible: out-of-scope E.g. selected household is not in the target population because all members are over 75 years old.	25	25	→ if not applicable, please leave as is	
				-1	-1
7.2	Other ineligible 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.	-1	-1	-1	-1

7.B	Number of eligible elements I.e. the gross sample size corrected for the ineligible cases. ► $[\$7.B] = [\$7.A] - [\$7.1] - [\$7.2]$	6475	6475	-1	-1
------------	---	------	------	----	----

	Type of unit non-response (eligible cases)	Number of households	Number of individuals (aged 16-74)	Number of individuals (<16)	Number of individuals (>74)
7.3	Non-contact E.g. no one was home or postal survey was never sent back.	1693	1693	→ if not applicable, please leave as is	
				-1	-1

7.4	Refusal E.g. selected household or individual was contacted but refused to take part in the survey.	573	573	-1	-1
7.5	Inability to respond E.g. selected household or individual was unable to participate due to language barriers or cognitive or physical incapacity to respond.	-1	-1	-1	-1
7.6	Rejected interviews 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.).	-1	-1	-1	-1
7.7	Other non-response 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. Insert text	-1	-1	-1	-1

7.C	Net sample size 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. ► $[\$7.C] = [\$7.B] - [\$7.3] - [\$7.4] - [\$7.5] - [\$7.6] - [\$7.7]$	4209	4209	-1	-1
------------	---	-------------	-------------	-----------	-----------

		Households	Individuals (aged 16-74)	Individuals (<16)	Individuals (>74)
7.D	Unit response rate The unit response rate is the ratio 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). ► $[\$7.D] = [\$7.C] / [\$7.B]$	65.0	65.0	→ if not applicable, please leave as is	
				-1	-1

7.8	Comments on the unit response rate, if any Insert text
------------	---

7.9 Methods used for minimizing unit non-response

Where applicable, give a description of measures taken to reduce the unit non-response:

- advance notification in the form of a letter or phone call;
- system of reminders, number of visits, number of attempts for phone calls, etc.
- showing respondents how the data they are providing are being used;
- etc.

An advanced notification letter was sent to all households, approximately two months before the survey conduct, among others providing information on how the data collected are being used. In cases where the households couldn't be approached, mainly due to temporary absence, a number of attempts for phone calls (up to three) were used.

7.10 Methods used for dealing with unit non-response

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.).

Correction factor in the weighting procedure: In each stratum h , the inverse of the initial inclusion probabilities of the respondents are multiplied by the inverse of the response rate r_h of the stratum h . 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.

7.11 Has substitution been used?

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).

If yes, please give an estimate of the percentage of proxy interviews (compared to the total number of interviews).

Substitution rate (%)

On which criterion has the selection of the substituted units been based?

Yes

☐

No

☒

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

Questions or items with item response rates below 90%

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.

The data entry program didn't allow for missing items. Imputation was done only for sole answers in item (a) of question A3.

7.13 What imputation methods have been used?

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.).

(Multiple choices possible)

- ☐ **None**
- ☐ **Deductive imputation**

An exact value can be derived as a known function of certain characteristics (e.g. the value received for a family allowance is a known function of certain characteristics like income class, age of children, etc. As soon as those characteristics are known, it becomes possible to calculate the value of a family allowance without error.)

Deterministic imputation

Deterministic imputation leads to estimators with no random component, that is, if the imputation were to be re-conducted, the outcome would be the same.

- ☐ Mean/Median
- ☐ Mean/Median by class
- ☐ Regression-based
- ☐ Donor
- ☐ Other (please specify): Insert Text

Random imputation

Random imputation leads to estimators with a random component, that is, if the imputation were re-conducted, it would lead to a different result.

- ☒ **Hot-deck** (for sole answers in item a of question A3)
- ☐ Cold-deck
- ☐ Simulated residuals
- ☐ Other (please specify): Insert text

☐ **Multiple imputation**

Multiple imputation methods offer the possibility of deriving variance estimators by taking imputation into account. In multiple imputation each missing value is replaced (instead of a single value) with a set of plausible values that represent the uncertainty of the right value to impute. The incorporation of imputation variance can be easily achieved based on the variability of estimates among the multiply imputed data sets.

7.14 What was the overall imputation rate for the main indicators?

For each of the target indicators that you listed in your response to question 6, please report the proportion of observations that are imputed values. Moreover, if applicable, please report the share of the estimate that is contributed by the imputed values.

Main target indicator	Imp. Rate (% of observations)	Imp. Rate (share of estimate)
Proportion of households having access to the Internet at home (item 'Yes' in variable A2 of the 2013 model questionnaire)	0%	0%

	Proportion of households using a broadband connection (a 'Yes' on option b, c, d, e or f in A3 of the 2013 model questionnaire)	6,7%	Because the imputation was done only for sole answers of item (a) in Q.A3 (variable h_broad) so the overall imputation rate is 6,7%
	Proportion of individuals regularly using the Internet: overall (indiv. who ticked option 1 or 2 in C2 of the 2013 model questionnaire)	0%	0%

7.15	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.).
	Insert text

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.

How were the design weights calculated?

Was balanced sampling used and what are the balancing variables?

Has re-weighting for units non-response been performed?

Has adjustment to external data sources been performed?

Is the sample self-weighted?

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 $H = 90$). 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 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 2nd quarter 2011

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 n_h

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 2012.

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 hij 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 2012.

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 ($\hat{\sigma}_{\hat{\theta}_i}$). 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 percentage points)* for the indicators and subindicators mentioned.

Please note that the accuracy measure used, i.e. the STANDARD DEVIATION, was also addressed in the 2006-2012 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.

Indicator or subindicator - on households and individuals in the general scope (16-74) and related subgroups -		<i>Number of respondents</i>	<i>Estimated proportion (%)</i>	<i>Standard error (% points)</i>
9.1	Proportion of households having access to the Internet at home (item 'Yes' in variable A2 of the 2013 model questionnaire)	2124621	56.3%	0.83
9.2	Proportion of households using a broadband connection (a 'Yes' on option b, c, d, e or f in variable A3 of the 2013 model questionnaire)	2075886	55.0%	0.84
9.3	Proportion of individuals regularly using the Internet: overall (indiv. who ticked option 1 or 2 in variable C2 of the 2013 model questionnaire)	4681529	55.9%	0.92
9.3.1	Proportion of ind. regularly using the Internet: males (as % of all men=4115593)	2457271	59.7%	1.34
9.3.2	Proportion of ind. regularly using the Internet: females (as % of all women=4264271)	2224258	52.2%	1.23
9.3.3	Proportion of ind. regularly using the Internet: age group 16-24 years (as % of all individuals aged 16-24 years=1069041)	952517	89.1%	2.01
9.3.4	Proportion of ind. regularly using the Internet: age group 25-34 years (as % of all individuals aged 25-34 years=1592503)	1295643	81.4%	2.00
9.3.5	Proportion of ind. regularly using the Internet: age group 35-44 years (as % of all individuals aged 35-44 years=1722106)	1210814	70.3%	1.75
9.3.6	Proportion of ind. regularly using the Internet: age group 45-54 years (as % of all individuals aged 45-54 years=1550464)	766215	49.4%	2.00
9.3.7	Proportion of ind. regularly using the Internet: age group 55-64 years (as % of all individuals aged 55-64 years=1346422)	362220	26.9%	1.70
9.3.8	Proportion of ind. regularly using the Internet: age group 65-74 years (as % of all individuals aged 65-74 years=1099328)	94119	8.6%	1.06
9.3.9	Proportion of ind. regularly using the Internet: low educational level (as % of all individuals with low education=2859614)	625655	21.9%	1.40
9.3.10	Proportion of ind. regularly using the Internet: medium educat. level (as % of all individuals with medium education=3365308)	2160353	64.2%	1.42
9.3.11	Proportion of ind. regularly using the Internet: high educational level (as % of all individuals with high education=2154942)	1895522	88.0%	1.06

9.3.12	Proportion of ind. regularly using the Internet: students (as % of all students=768029)	718833	93.6%	1.77
9.3.13	Proportion of ind. regularly using the Internet: employees or self-employed (as % of all employees or self-employed=3471556)	2530614	72.9%	1.29
9.3.14	Proportion of ind. regularly using the Internet: unemployed (as % of all unemployed=1459340)	856958	58.7%	2.44
9.3.15	Proportion of ind. regularly using the Internet: retired, other inactive (as % of all retired and other inactive=2680939)	575123	21.5%	1.15
9.4	Proportion of individuals having downloaded official forms in the last 12 months (individuals who ticked item b in variable D1 of the 2013 model questionnaire)	1593614	19.0%	0.79
9.5	Proportion of individuals having ordered goods or services for private use over the internet in the last 12 months (individuals who ticked option 1 and 2 in variable E1 of the 2013 model questionnaire)	2058535	24.6%	0.86

9.6	<p>Calculation of the standard error</p> <p>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.</p>
	a) Name and brief description of the applied estimation approach
	<p><input checked="" type="checkbox"/> Analytic Method</p> <p>Linearization Method</p> <p><input type="checkbox"/> Taylor linearization</p> <p><input type="checkbox"/> Linearization based on influence functions</p> <p><input checked="" type="checkbox"/> Other, please specify: The variance estimation considers PPS of PSU's with replacement. As the clusters (one or more unified city blocks) are used as primary sampling units (PSUs) in the sample design, the variance procedure estimates the variance from the variation among the PSUs.</p> <p>Replication Methods</p> <p><input type="checkbox"/> Jackknife</p> <p><input type="checkbox"/> Bootstrap</p> <p><input type="checkbox"/> Balanced repeated Replication / Balanced half-samples</p> <p><input type="checkbox"/> Random Groups</p> <p><input type="checkbox"/> Other, please specify: <input type="text" value="Insert Text"/></p> <p><input type="checkbox"/> Other, please specify: <input type="text" value="Insert Text"/></p>
	<p>b) Basic formula</p> <p>Let $w_{hijk} (>0)$ stand for the survey weight attached to the sample individual k ($k = 1, \dots, n_{hij}$, 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.</p> <p>Estimation of survey characteristics</p> <p>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</p>

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 y 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:

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

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:

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

- 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:

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

where:

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

c) What tools were used?

- ☐ CLAN
- ☐ GENESEES
- ☐ SUDAAN
- ☐ POULPE
- ☐ CALJACK
- ☐ BASCULA
- ☐ ReGenesees

Other, please specify: **SPSS**

d) Do the methods take into account the effect of:

☒ **unit non-response?**

The variance estimator $\hat{V}(\hat{\theta})$ has to be adjusted to take unit non-response into account. Different methods can be used: methods based on the assumption that respondents are missing at random or completely at random within e.g. strata or constructed response homogeneity groups, methods using the two-phase approach, etc. See section 3.4 for more information.

If yes, please indicate: **The variance estimation procedure adds the final stratum variance estimates to compute the overall variance estimate.**

☒ **imputation?**

Imputation variance can be estimated if multiple imputation is used.

Replication and analytic methods can be used to incorporate imputation into variance estimation.

Deville and Särndal (1994) proposed a method for the regression imputed Horvitz-Thompson estimator.

See section 3.4 for more information.

If yes, please indicate: **The imputed values were not taken into account in the procedure of variance estimation.**

☒ **coverage errors (over-coverage, multiple listings)?**

Methodology of domain estimation can be used. Target population has to be defined as a domain of the frame population.

The related loss of precision can be quantified.

See section 3.4 for more information.

If yes, please indicate: **domain defined by gender and age groups was used in the calibration process in order to reduce the bias due to coverage errors**

☐ **implicit stratification?**

One way to consider implicit stratification is to define explicit strata, from which each of an independent sample is supposed to have been selected.

Other methods using analytic formulae are available. See section 3.4.

If yes, please indicate:

☐ **rotating samples?**

In case of rotating sample schemes, the overlap of samples between e.g. successive quarters reduces the precision of the average of estimates from e.g. quarterly samples and increases the precision for e.g. the quarter-to-quarter estimates of change.

See section 3.7.

If yes, please indicate:

☐ **calibration?**

Methods to account for the effect of calibration on variance should be used. Deville and Särndal method (1992) is presented in section 3.4.

If yes, please indicate:

e) 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

9.7	Comments on reliability and representativeness of results and completeness of dataset			
<p>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% points for the overall proportions and shall not exceed 5% points 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).</p>				
Indicators and breakdowns in sections 9.1 to 9.5 above:				
Other indicators and breakdowns:				
FOR ALL INDICATORS AND BREAKDOWNS in transmitted aggregate data:				
If applicable, please tick:				
If significant sampling errors were found, were data cells suppressed in the transmitted aggregate 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 aggregate dataset and flagged as unreliable?	Yes		No	

10. Closing remarks

10.1	Problems encountered and lessons to be learnt 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.)
10.2	Other comments, if any

11. Annexes

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

11.1	Questionnaire in national language - Yes
11.2	Questionnaire in English (if available) -No
11.3	Interviewer instructions (if available) - Yes
11.4	National reports on methodology (if available)
11.5	Analysis of key results, backed up by tables and graphs (if available)
11.6	Other annexes 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"> • ... • ...