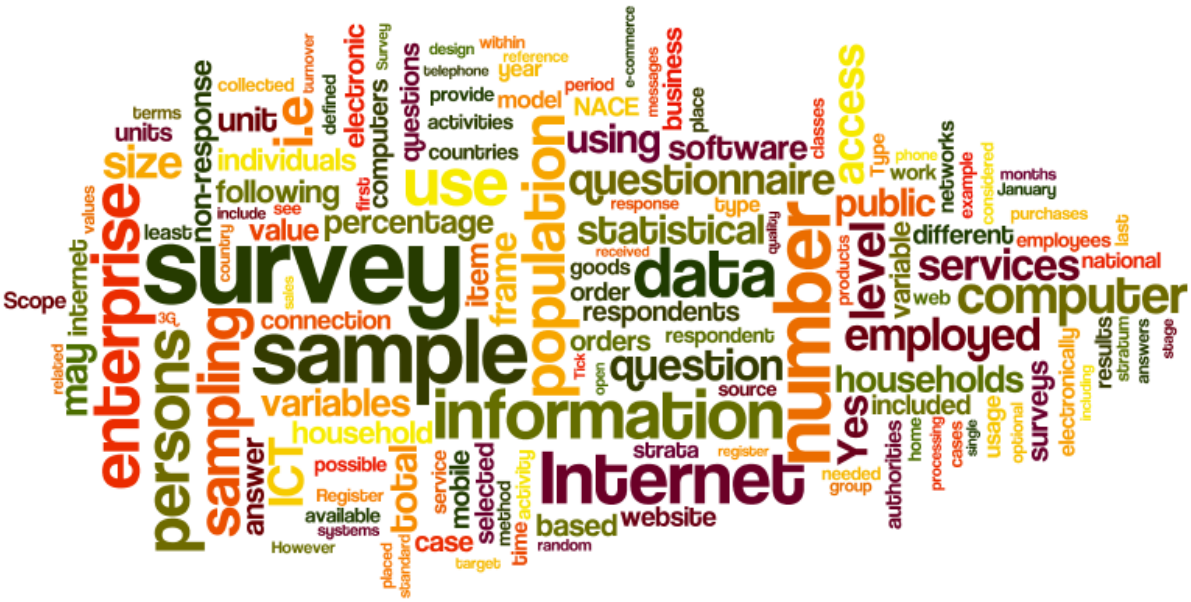
Methodological Manual



for **statistics** on the

Information Society

Survey year 2014, version 1.0

Methodological manual for statistics on the Information Society

Implementation of Article 5 of Regulation (EC) No 808/2004 of the European Parliament and of the Council of 21 April 2004 concerning Community statistics on the Information Society (OJ L143, 30.04.2004, p. 49[1]). The latest amendment of (EC) No 808/2004 is Regulation (EC) No 1006/2009 of the European Parliament and of the Council of 16 September 2009 (OJ L286, 31.10.2009, p. 31[2]).

Published by Eurostat

Survey year 2014

Version 1.0, July 2014

This publication is an update of the previous methodological manuals and has been prepared by Unit G4 "Innovation and Information Society" of Eurostat with the important contribution of the Task Forces supporting the surveys on "ICT usage and e-Commerce in Enterprises" and "ICT usage in households and by individuals". The opinions expressed are those of the authors alone and do not necessarily reflect the position of the European Commission.

Contacts

Part I: Konstantinos Giannakouris (Eurostat, Unit G4),

+352 4301 34877

konstantinos.giannakouris@ec.europa.eu

Part II: Heidi Seybert (Eurostat, Unit G4),

+352 4301 37416

Heidi.Seybert@ec.europa.eu

[1] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2004:143:0049:0055:EN:PDF
http://eur-lex.europa.eu/smartapi/cgi/sga_doc?smartapi!celexapi!prod!CELEXnumdoc&lg=EN&numdoc=32004R0808&model=guicheti

[2] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:286:0031:0035:EN:PDF
http://eur-lex.europa.eu/smartapi/cgi/sga_doc?smartapi!celexapi!prod!CELEXnumdoc&lg=en&model=guicheti&numdoc=32009R1006

TABLE OF CONTENTS

Preface

Information and Communications Technologies (ICT) account for a significant part of the EU productivity and growth and are transforming our societies and economies in a profound and unprecedented way. Official statistics are indispensable for an informed understanding of the implications of the transformations underway. In this aspect the selection of the appropriate set of variables is crucial. The measurement of the Information Society by statistics on society, on business processes and productivity is an area for continuous revision and improvement.

Since 2002 the European Commission has established the annual Information Society surveys to benchmark the ICT-driven development in enterprises and by individuals. Eurostat coordinates two surveys to be carried out at national level, one on "ICT usage and e-Commerce in enterprises" and one on "ICT usage in households and by individuals". The surveys are developed in close collaboration with the Member States and the OECD and are adapted to the changing needs of users and policy makers. These surveys are based on Model Questionnaires and the accompanying methodological guidelines for their implementation. In 2004 the European Parliament and the Council adopted Regulation (EC) No 808/2004 covering the above mentioned surveys. That regulation is a framework regulation hence it allows flexibility to adapt the surveys to newly evolving needs by users and decision makers. Annual implementing measures such as the Commission Regulation[3] (EU) No 859/2013 of 5 September 2013 implementing regulation (EC) No 808/2004 of the European Parliament and of the Council concerning Community Statistics on the Information Society are forming the basis for the Eurostat annual Model Questionnaires and the respective surveys. This ensures harmonized data for all EU-28 Member States. The annual implementing Regulation is EEA relevant.

This methodological manual refers to statistical data used for monitoring the i2010 strategy; a benchmarking framework for the Information Society policy covering the period 2011-2015 was adopted in 2009. Additionally, statistical data support the follow-up of the "Digital agenda for Europe[4]" Commission communication[5] (19/5/2010) to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. The latter communication is an essential part of the Europe 2020 strategy launched in March 2010.

The first manual, which was related to the 2006 survey, was compiled in a joint effort of Eurostat, the National Statistical Institutes (NSI) and the OECD. The current version is drawn up as a tool for helping NSI to translate the two Eurostat Model Questionnaires into national languages and to implement the national surveys following a harmonised methodology. It is a hands-on tool with recommended guidelines and does not replace statistical handbooks. Its main focus is on explanatory notes to the variables and attributes of the Model Questionnaires for the year 2014. As the original manual, it does however include a set of recommendations to NSI to be taken into account when carrying out the survey to ensure the production of high quality and comparable results as foreseen in the regulatory framework. Annexes with the Model Questionnaires, the transmission formats and the reporting templates as well as the legal background documents are included.

The concept of this methodological manual is to use it as a rolling document. Any proposal to the authors to improve the manual is welcome, also from interested readers outside the European Statistical System, in view of reaching harmonized Information Society statistics at global level.

Albrecht Wirthmann Heidi Seybert Konstantinos Giannakouris

Information Society Statistics

[3] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:238:0005:0020:EN:PDF
http://new.eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2013.238.01.0005.01.ENG
[4] http://ec.europa.eu/digital-agenda/en/scoreboard
[5] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:0245:FIN:EN:PDF

Part I

Enterprise survey

I.1. STATISTICAL PRODUCT

The statistical product is the clear and precise definition of the statistical information to be produced. It distinguishes itself from the production methodology. The production methodology is the way or method of doing, while the statistical product is its direct result. Different statistical methodologies can produce the same statistical product, being only different ways of doing the same thing. This means that as long as we guarantee that two figures concern the same statistical product, for instance for two different countries, they are comparable. This way, this distinction between the statistical product and the statistical methodology helps us to focus on those elements more important to assure comparability between the several national statistics and produce new ones at the EU level, i.e. the statistical product, while leaving to the discretion of the national statistical institutes the choice of the better statistical methodology to be applied in its own country taking into consideration its own specificities.

The elements that make up the statistical product, at an input level, are the statistical unit, the target population and the observation variables, and at the output level, the periodicity and the summary measures, aggregate variables and tabulation. Covering all the elements of the statistical product, the statistical concepts and the nomenclatures are the additional needed elements to ensure harmonization and comparability of statistics.

I.1.1. Statistical unit

The statistical unit is the base type of the elements of a group (also called population) that we want to observe or analyse. The basic statistical operations of classification, aggregation and ordering are done on the statistical unit.

The choice of the statistical unit is a matter of both the data collection process (namely the operational restrictions associated to collecting data from each type of statistical unit) and the conceptual framework chosen to observe and analyse the phenomenon. The statistical unit is the bearer of statistical characteristics or attributes, which we ultimately want to measure.

There are several types of statistical units, according to its usage. An observation unit represents an identifiable entity, about which data can be obtained. During the collection of data, this is the unit for which data is recorded. It should be noted that this may, or may not be, the same as the reporting unit. The reporting unit is the unit that reports to the survey authority. It reports information for the observation unit(s). In certain cases it may be different from the observation unit. A reporting unit is a unit that supplies the data for a given survey instance.

The observation statistical unit in the "Survey on ICT usage and e-Commerce in Enterprises" is the **enterprise**, as defined in the Council Regulation[6] (EEC) No 696/93 of 15 March 1993 on the statistical units for the observation and analysis of the production system in the Community defines a list of statistical units:

"The enterprise is the smallest combination of legal units that is an organizational unit producing goods or services, which benefits from a certain degree of autonomy in decision-making, especially for the allocation of its current resources. An enterprise carries out one or more activities at one or more locations. An enterprise may be a sole legal unit."

The enterprise thus defined is an economic entity which can therefore, under certain circumstances, correspond to a grouping of several legal units. Some legal units, in fact, perform activities exclusively for other legal units and their existence can only be explained by administrative factors (e. g. tax reasons), without them being of any economic significance. A large proportion of the legal units with no persons employed also belong to this category. In many cases, the activities of these legal units should be seen as ancillary activities of the parent legal unit they serve, to which they belong and to which they must be attached to form an enterprise used for economic analysis.

However, the definition of the enterprise as the appropriate statistical unit poses some limitations. Some enterprises, especially the larger ones, are composed of several local units (establishment). Because of this, a geographical breakdown of the results (although still possible using the local of the main headquarters of the enterprise) is of limited use. Nevertheless, ICT usage is not easily attributable to the different establishments of an enterprise, and for this reason the enterprise is the statistical unit adopted; for this reason and also for reducing the response burden on enterprises, the geographical breakdown within countries is not considered relevant.

I.1.2. Target population

A population is a collection of objects of the same class, which in statistical terms means a group of elements of the same statistical unit. There are two types of populations to be considered when producing statistics: the target population and the frame population.

The **target population** is the population of interest. It is defined by delimiting clearly the group of statistical elements for which some information concerning the all group we want to know. That delimitation is based on one or more attributes of the statistical unit. For example, for the enterprise, some commonly used attributes to delimit the target population are the size (e.g. number of persons employed), the economic activity and its location. An example of a target population could be "enterprises with 10 or more persons employed, classified in Divisions 41-43 of NACE (Construction), localised in the EU".

The **frame population** is an operationalization of the target population, consisting ideally of a complete list of the elements of the target population. Although a target population can be easily defined, in practice a list of all its elements is needed for its complete or partial (in case a sample is used) observation, and that can be very difficult to obtain. That list should be complete and include every element of the target population only once. However, most of the time, it will suffer from both

[6] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31993R0696:EN:HTML

under-coverage and over-coverage. Usually files of statistical elements (registers) are maintained and updated, containing lists of statistical elements and also information on some attributes, usually used for delimiting target populations. Frame populations are usually extracted from those registers. The frame population will be further explained in I.2.

The target population of the "Survey on ICT usage and e-Commerce in Enterprises" is the group of enterprises delimited by the following attributes:

- **Economic activity:**

NACE Rev.2 has been implemented for the first time in the 2009 survey. During that survey, both classifications NACE Rev. 1.1 and NACE Rev. 2 had been used to report the data. The results for the surveys after 2010 will be reported following NACE Rev. 2 only.

Enterprises classified in the following categories of NACE Rev. 2:

- Section C – "Manufacturing";
- Section D, E – "Electricity, gas, steam and air conditioning supply; Water supply; sewerage, waste management and remediation activities";
- Section F – "Construction";
- Section G – "Wholesale and retail trade; repair of motor vehicles and motorcycles;
- Section H – "Transportation and storage";
- Section I – "Accommodation and food service activities";
- Section J – "Information and communication";
- Section L – "Real estate activities";
- Divisions 69-74 – "Professional, scientific and technical activities " (without veterinary activities);
- Section N – "Administrative and support service activities";
- Group 95.1 – "Repair of computers and communication equipment"

The enterprises are classified in one of these categories according to their **principle** economic activity.

- **Enterprise size**

Enterprises with 10 or more persons employed;

Optionally the target population can be extended to enterprises with number of persons employed between 1 and 9.

Please note that the number of persons **employed** is defined in Commission Regulation[7] (EC) No 250/2009 of 11 March 2009 (p.38-39, Code: 16 11 0; Number of persons employed) and should not be confused with the number of **employees** or with **FTE's**. In the following, from the statistical definition point of view, even though the word "employees" is used it always refers to "persons employed".

- **Geographic scope**

Enterprises located in any part of the territory of the country.

[7] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:086:0001:0169:en:PDF

I.1.3. **Periodicity**

The periodicity is annual, meaning the data is collected and compiled once per year. However, some variables can be observed with a lower frequency, e.g. variables which tend to be stable over time, variables for the biennial benchmarking indicators.

An annual survey should be a compromise or trade-off between response and collection burden and the need for relevant and recent information on a "fast moving" study domain like the Information Society. In particular, variables for the biennial benchmarking indicators aim at keeping the burden on enterprises as stable as possible over the years.

This periodicity is laid down in paragraph c) of Annex I of the Regulation 808/2004 and in section 5 of Annex I of the Commission Regulation[8] (EU) No 859/2013 of 5 September 2013.

I.1.4. **Observation variables**

The attributes of the statistical unit are what in the end we want to observe and the observation variables hold numerical measures of those attributes. Attributes and observation variables shouldn't be confused. An attribute is some property of the statistical unit and each attribute may have one or more observation variables which hold numerical measures of that attribute.

For example, for the statistical unit "enterprise" an attribute is the "percentage of the total purchases that resulted from orders placed electronically". We want to measure this attribute and can just use a variable to hold the percentage of the purchases via computer networks. But collecting the percentage of enterprises' purchases ordered via computer networks might not be feasible. In this case we can use another variable which is feasible, like a categorical variable with 7 different ranges of percentages of purchases, which information enterprises are willing and able to give (see alternative questions in module E-Commerce). In this case two different observation variables can be used to measure the same attribute.

There are many ways of classifying observation variables, but the most relevant one for the ICT usage surveys is the distinction between qualitative and quantitative variables.

Qualitative variables concern non-numerical information. They serve merely as labels or names for identifying special attributes of the statistical unit. An example is the "use (Yes or No) of computers" by the enterprise. Sometimes qualitative variables can be rendered numeric by coding the non-numeric values. An important type of qualitative variables is that of binary variables (dichotomous variables). Binary variables assume only two different values, which usually are rendered numeric by attributing the values "0" and "1" in a meaningful way.

Quantitative variables contain information as to how much or how many. Hence they are always numeric. An example is the number of persons employed that used computers with access to the internet.

The observation variables in the "Survey on ICT usage and e-Commerce" are mainly qualitative, namely binary.

[8] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:238:0005:0020:EN:PDF

For the 2014 ICT survey in enterprises the observation variables to be collected refer to the following (scope and filter questions apply as in the model questionnaire):

Question number in Model Questionnaire	Variable	Optional	Related to benchmarking indicators
A1	computer usage		
A2	(optional) persons employed or percentage of the total number of persons employed who use computers for business purposes.	X	
B1	employment of ICT specialists		
B2a	provision of any type of training to develop ICT related skills for ICT specialists, during the previous calendar year		
B2b	provision of any type of training to develop ICT related skills for other persons employed, during the previous calendar year.		
B3	recruitment or attempted recruitment of ICT specialists, during the previous calendar year		
B4	existence of vacancies for ICT specialists that were difficult to fill, during the previous calendar year		
C1	access to the internet		
C2	persons employed or percentage of the total number of persons employed who use computers with access to the internet for business purposes		X
C3	internet connection: DSL or any other type of fixed broadband connection,		X
C4(a)-(e)	internet connection: maximum contracted download speed of the fastest internet connection in Mbit/s; ([0;<2], [2;<10], [10;<30], [30;<100], [>=100])		
C5a	internet connection: mobile broadband connection via a portable device using mobile telephone networks (so-called 3G or 4G)		X
C5a1	(optional) internet connection: mobile broadband connection via portable computer using mobile telephone networks (so-called 3G or 4G)	X	
C5a2	(optional) internet connection: mobile broadband connection via other portable devices like smartphones using mobile telephone networks (so-called 3G or 4G)	X	
C5b	internet connection: other mobile connection using mobile telephone networks		
C6	persons employed or percentage of the total number of persons employed using a portable device provided by the enterprise which allows internet connection via mobile telephone networks, for business purposes		X
C7	use of a website		
C8a	(optional) provision of the following facility: description of goods or services, price lists	X	
C8b	provision of the following facility: online ordering or reservation or booking		X
C8c	(optional) provision of the following facility: possibility for visitors to customise or design online goods or services,	X	
C8d	(optional) provision of the following facility: tracking or status of placed orders	X	
C8e	(optional) provision of the following facility: personalised content in the website for regular/repeated visitors	X	
C8f	(optional) provision of the following facility: links or references to the enterprise's social media profiles,	X	
C8g	(optional) provision of the following facility: privacy policy statement, privacy seal, or certification of website safety	X	
C8h	(optional) provision of the following facility: advertisement of open job positions or online job application	X	
C8i	provision of the following facility: electronic submission of complaints		

C9a	(optional) use of social networks, not solely used for posting paid adverts	X	
C9b	(optional) use of enterprise's blogs or microblogs, not solely used for posting paid adverts	X	
C9c	(optional) use of multimedia content-sharing websites, not solely used for posting paid adverts	X	
C9d	(optional) use of wiki-based knowledge sharing tools, not solely used for posting paid adverts	X	
C10	(optional) provision of remote access to the enterprise's e-mail system, documents or applications	X	
C11	(optional) use of paid adverts on the internet	X	
D1	use of cloud computing services excluding free of charge services		
D2a	use of e-mail as a cloud computing service		
D2b	use of office software as a cloud computing service		
D2c	hosting the enterprise's database(s) as a cloud computing service		
D2d	storing files as a cloud computing service		
D2e	use of finance or accounting software applications as a cloud computing service		
D2f	use of Customer Relationship Management (CRM) as a cloud computing service		
D2g	use of computing power for running the enterprise's own software as a cloud computing service		
D3a	use of cloud computing services delivered from shared servers of service providers		
D3b	use of cloud computing services delivered from servers of service providers exclusively reserved for the enterprise		
D4a	factors limiting the use of cloud computing services: risk of a security breach		
D4b	factors limiting the use of cloud computing services: problems accessing data or software		
D4c	factors limiting the use of cloud computing services: difficulties in unsubscribing or changing service provider (including concerns with data portability)		
D4d	factors limiting the use of cloud computing services: uncertainty about the location of the data		
D4e	factors limiting the use of cloud computing services: uncertainty about applicable law, jurisdiction, dispute resolution mechanism		
D4f	factors limiting the use of cloud computing services: high costs of buying cloud computing services		
D4g	factors limiting the use of cloud computing services: insufficient knowledge of cloud computing		
D5a	(optional) the degree that benefits related to reduction of ICT related costs were realised, on a scale of: "to a high degree", "to some degree", "to a limited degree", "not at all"	X	
D5b	(optional) the degree that benefits related to flexibility due to scaling cloud computing services up or down were realised, on a scale of: "to a high degree", "to some degree", "to a limited degree", "not at all"	X	
D5c	(optional) the degree that benefits were realised related to easy and quick deployment of solutions based on cloud computing , on a scale of: "to a high degree", "to some degree", "to a limited degree", "not at all".	X	
D6a	(optional) factors preventing the use of cloud computing services: risk of a security breach	X	
D6b	(optional) factors preventing the use of cloud computing services: uncertainty about the location of the data	X	
D6c	(optional) factors preventing the use of cloud computing services: uncertainty about applicable law, jurisdiction, dispute resolution mechanism	X	
D6d	(optional) factors preventing the use of cloud computing services: high costs of buying cloud computing services	X	

D6e	(optional) factors preventing the use of cloud computing services: insufficient knowledge of cloud computing	X	
E1	use of an ERP (Enterprise Resource Planning) software package to share information among different functional areas		X
E2a	use of any software application for managing information about customers (Customer Relationship Management — CRM software) that allows it to capture, store and make available to other business functions information about the enterprise's customers		
E2b	use of any software application for managing information about customers (Customer Relationship Management — CRM software) that allows the analysis of information about customers for marketing purposes		X
F1a	use of EDI-type (Electronic Data Interchange – EDI-type) messages for sending payment instructions to financial institutions		X
F1b	use of EDI-type (Electronic Data Interchange – EDI-type) messages for sending or receiving data to/from public authorities		X
G1	sharing electronically Supply Chain Management (SCM) information with suppliers or customers in order to coordinate the availability and delivery of products or services to the final consumer, without the information being typed manually		X
G2a	sharing electronically Supply Chain Management information with suppliers or customers via websites (enterprise's website, website of business partners or web portals)		
G2b	sharing electronically Supply Chain Management information with suppliers or customers via electronic exchange of information suitable for automated processing		
H1a	use of Radio Frequency Identification (RFID) instruments for person identification or access control		X
H1b	use of Radio Frequency Identification (RFID) instruments as part of the production and service delivery process		X
H1c	use of Radio Frequency Identification (RFID) instruments for product identification after the production process		X
I1	sending invoices to other enterprises or public authorities during the previous calendar year		
I2a	percentage of all invoices sent as eInvoices in a standard structure suitable for automated processing, to other enterprises or public authorities, during the previous calendar years		
I2b	percentage of all invoices sent as invoices in electronic form not suitable for automated processing, to other enterprises or public authorities, during the previous calendar year		
I2c	percentage of all invoices sent only as invoices in paper form, to other enterprises or public authorities, during the previous calendar year		
I3a	percentage of all invoices received as eInvoices, in a standard structure suitable for automated processing, during the previous calendar year		
I3b	percentage of all invoices received as invoices in paper form or in electronic form not suitable for automated processing, during the previous calendar year		
J1	receiving orders for goods or services that were placed via a website (web sales), in the previous calendar year		X
J2	value or percentage of the total turnover of e-commerce sales resulting from orders placed via a website, in the previous calendar year		X
J3a	percentage of e-commerce sales to private consumers (B2C) resulting from orders placed via a website, in the previous calendar year		
J3b	percentage of e-commerce sales to other enterprises (B2B) and e-commerce sales to public authorities (B2G) resulting from orders placed via a website, in the previous calendar year		
J4a	use of online payment systems for sales via a website i.e. payment integrated in the ordering transaction		
J4b	use of offline payment systems for sales via a website i.e. payment not included in the ordering transaction		
J5	receiving orders for goods or services via EDI-type messages (EDI-type sales), in the previous		X

	calendar year		
J6	placing orders for goods or services via a website or EDI-type messages, in the previous calendar year		X
J7	value or percentage of the total turnover of e-commerce sales resulting from orders received via EDI-type messages, in the previous calendar year		X
J8	(optional) placing orders for goods or services via a website, in the previous calendar year	X	
J9	(optional) placing orders for goods or services via EDI-type messages, in the previous calendar year	X	X
J10	placing orders for goods or services of a value equal or more than 1 % of the total purchases' value, in the previous calendar year		X
X1	main economic activity of the enterprise, in the previous calendar year		
X2	average number of persons employed, in the previous calendar year		
X3	total turnover (in value terms, excluding VAT), in the previous calendar year		

I.1.5. Summary measures, aggregated variables, indicators and tabulation

Summary measures are numerical values defined by a statistical measure that is used to summarize the values for a specific variable for all statistical units in a specific group. Such measures can take the form of aggregates (e.g. total number of *yes*-answers on a specific question) or indicators (e.g. percentage of *yes*-answers).

Aggregates can be compiled for the total population or for the different subpopulations defined by the background variables (e.g. NACE category or size class) or for subpopulations defined conditionally on the answers of other study variables (e.g. broadband users versus non broadband users).

E.g.:

- number of enterprises with internet access
- number of enterprises with between 10 and 49 persons employed with internet access

To obtain indicators (proportions, percentages, ratios, etc.), the aggregates need to be divided by the corresponding total population or subpopulation.

The denominator of such proportion can for instance be:

- total population of enterprises in the selected strata (e.g. "proportion of enterprises with between 10 and 49 persons employed with access to the internet" = the number of enterprises with between 10 and 49 persons employed with access to the internet divided by the total number of enterprises with between 10 and 49 persons employed).

The detailed transmission format for sending data to Eurostat can be found in the *Transmission Format (Annex I.3.2)*. In that document, aggregates are specified; on the basis of these aggregates, Eurostat calculates and releases indicators.

I.1.6. Explanatory notes

The explanatory notes in this chapter refer to the questions in the 2014 Model Questionnaire (see Model Questionnaire in section I.3.1). The structure of this chapter follows the Model Questionnaire, i.e.

the explanatory notes are grouped per *module* and per *question*. It is recommended to have the Model Questionnaire at hand while reading this section.

I.1.6.1 GENERAL REMARK: "USE", "HAVE" OR "HAVE ANOTHER ENTERPRISE USE FOR YOU"

In many of the questions and the corresponding notes, reference is made to the use of computers, networks, systems, software etc. The term **"use"** does not refer to the ownership of such goods and infrastructure. For example computers may belong to the enterprise, may be leased or shared with another organisation. In the case where computers etc. are used by the responding enterprise, but provided or maintained by another enterprise, it should be considered as used by the responding enterprise.

Additionally, enterprises often buy ICT services or services which relate to the ICT use in the context of the survey (like accounting). In cases where ICT services or other services which include ICT use are provided totally by another enterprise and the responding enterprise itself is not using ICT for that function, then it should not be counted as use of ICT of the responding enterprise.

In some cases there will be a mixture of ICT usage by the responding enterprise and the ICT service provider (another enterprise). In these cases it should be made clear from the question what kind of activities will be counted as ICT usage of the responding enterprise.

Examples on the interpretation of special cases: In some questions the ICT involvement of the enterprise is not about the ICT usage, but rather that the enterprise has or offers some digital solutions to its partners (customers or business partners). One important example is the question about having a website, where the question is about having, not using, and the subject is the solution that is offered to customers etc. regardless of how the website is technically maintained. The main issue about having a website is about control and responsibility of the contents in a separate area in the web. Web-pages of an enterprise can be maintained and designed by a service provider, but if the responding enterprise "owns" the contents, it is its own website. However, not all presence in the web means that the respondent has a website. Presence in certain service catalogues "yellow pages", address lists etc. are not counted as the enterprise's website. E-commerce systems can also be provided and maintained by service providers. Similarly to the case of the website, if it is the responding enterprise's e-commerce facility, the responding enterprise has control over the content of the site and it is the one who is selling and the owner of the products sold, then it is this respondent's e-commerce. Still a special case is internet market places, where enterprises can sell their products. In this case the respondent does not own or control the site, but it is controlling the sale and owning the products to be sold and therefore it is their e-commerce (if e-commerce definition is valid in the situation).

E-invoices are often sent in cooperation with service providers. The interest is in the use of e-invoices even when they are sent by service providers. It should be counted as respondent's use of e-invoicing when the enterprise itself takes part to the process electronically. The respondent "owns" the invoices and is responsible of the content of the invoices.

I.1.6.2 MODULE A: USE OF COMPUTERS

A1. Does your enterprise use computers?

Computers include Personal Computers, portable computers, tablets, other portable devices like Smartphones.

[Scope: all enterprises]

[Type: one single answer needed, i.e. *Tick only one*; binary (*Yes/No*); filter question]

A computer is a programmable device or machine capable of receiving, processing and outputting various types of data, performing calculations and other manipulations on it, under the control of a stored set of instructions. A computer typically consists of one or more processing units, memory units, and associated peripheral input and output devices.

Depending on size and power, computers can be generally divided in the following categories:

- **Supercomputer:** An extremely fast computer that can perform hundreds of millions of instructions per second.
- **Mainframe:** A powerful multi-user computer capable of supporting many hundreds or thousands of users simultaneously.
- **Minicomputer:** A multi-user computer capable of supporting from 10 to hundreds of users simultaneously.
- **Workstation:** A powerful, single-user computer like a personal computer, but it has a more powerful microprocessor and a higher-quality monitor.
- **Personal computer:** A small, single-user computer based on a microprocessor, has a keyboard for entering data, a monitor for displaying information, and a storage device for saving data.

Amongst personal computers one can find:

- **Desktop computers:** A desktop computer can be defined as a computer that usually remains fixed in one place.
- **Nettops:** Nettops are a classification of low cost desktop computer systems that use low power processors (designed for carrying out less demanding tasks, such as word processing, internet browsing and working with internet-based applications) instead of the usage of traditional desktop components. These energy efficient systems are designed for basic computing such as surfing on the internet, accessing web-based applications, document processing, audio/video playback etc. at very low cost and power consumption. The hardware specifications and processing power are usually reduced and hence nettops are less appropriate for running complex or resource intensive applications.
- **Portable computers (laptop):** A portable computer (or laptop) can be defined as a computer that is small enough to carry. A portable computer can usually carry out the same tasks as a desktop computer. Includes notebooks, netbooks, ultrabooks (smaller, lighter and more portable laptops), ultra mobile PC-UMPC, tablet PC etc. and also other portable devices like Smartphones, PDA phones that are considered as portable computers and are included in a residual category (handheld).
- A **tablet computer**, or simply tablet, is a type of portable computer integrated into a flat touch screen, operated by touching the screen rather than using a physical keyboard. It uses

onscreen virtual keyboard, stylus pen, or a digital pen. In the literature, the terms "tablet computer" and "tablet PC" are used interchangeably; the term "tablet" is proposed in the model questionnaire. The market of tablet computers is expending rapidly. The operating systems that are in principle used on tablets are the iOs (iPad), Android and Windows and tablets allow wireless connection to the internet via Wi-Fi networks and/or a mobile connection via telephone networks (integrated SIM card supporting mobile broadband connection to the internet). Tablets may not replace desktop or laptop computers but may be the next major computing platform for enterprise applications supporting various business processes such as sales, digital delivery, payments and marketing.

- **Other portable (handheld) computers:** A portable handheld computer is a computer that can be used while you are holding it and stored in a small bag or for instance in a pocket. **Personal Digital Assistants** (PDAs) are another term for handheld computers. A PDA uses a pen rather than a keyboard. PDA usually combine a telephone hence the term mostly used is PDA phones. **Smartphones:** Mobile phones that offer more advanced computing ability and connectivity than a contemporary basic phone. Smartphones and feature phones may be thought of as handheld computers integrated within a mobile telephone. Feature phones are able to run applications (not open source) based on platforms such as Java ME. A smartphone allows the user to install and run more advanced applications based on a specific platform (mostly open source operating systems). Smartphones run complete operating system software providing a platform for application developers. A smartphone can be considered as a Personal Pocket Computer (PPC) with mobile phone functions, because these devices are mainly computers, although much smaller than a desktop computer. In the scope of the question are only smartphones if advanced applications are used. Smartphones used only for making calls are thus not in the scope of this question.

All the above types of computers are in the scope of the question. It is noted that in order for a device to qualify as a computer we should consider two issues: a) the capacity/capability i.e. what can a person do with the device, and b) the usage/purpose of the device. Concerning the former point, the device should have computational capacity and capabilities (e.g. can I write a document?) and for the latter point the device should not be used solely as e.g. Electronic Point of Sales (EPOS) or a game box. It is reminded that *electronic tills, terminals, PCs that are solely used for monetary transactions – Electronic Point of Sale (EPOS) and computer controlled machinery* are excluded. From the usage/purpose point of view, "thin clients", qualify as computers for the survey because they may be used for accessing the internet and eventually for using "cloud computing" services hence for a wide range of applications.

A2. How many persons employed use computers for business purposes? *(Optional)*

If you can't provide this value,

Please indicate an estimate of the percentage of the total number of persons employed who use computers for business purposes. *(Optional)*

[Scope: enterprises using computers, i.e. A1 = Yes]

[Type: numerical, in absolute or percentage values; optional]

The concept of computer to be used in this question is the same as the one for question A1.

This question refers not to access only, but to effective use of computers by the personnel in the enterprise activities for business purposes.

The concept of "person employed" to be used in this variable is the one described in the explanatory note of variable X2 ("Average Number of Persons Employed").

This variable can be collected in absolute or in percentage values. The variable in absolute values might be more appropriate for smaller enterprises, while the one in percentage terms for bigger ones. When the number of persons employed is small, it's easier for the respondent just to identify which ones work with computers, instead of having to calculate the percentages afterwards. The collection in percentage terms may introduce measurement errors, because of the rounding that is necessary to provide an answer in percentages made of integer numbers between 1 and 100. The collection in percentage terms can also introduce other measurement errors as it induces the respondent to make an educated guess of its value introducing the bias usually associated to this type of answer. On the other hand, for bigger enterprises, with a high number of persons employed, the identification of every worker using (and not just having access) a computer can be burdensome or not possible and result in non-response. In this case the collection of this information in percentage terms inducing to an educated guess may be preferable.

Eurostat in its Model Questionnaire (annex I.3.1) recommends using a combination of both, giving to the respondent the opportunity to choose either to answer in absolute values or percentage values.

Independently on how this variable is collected, in absolute values or percentage values, the background variable "number of persons employed" is needed for grossing up the overall percentage of persons employed using computers. In order to compute the overall percentage both the total number of persons employed and the total number of persons using computers need to be grossed up. If collected in percentage terms, the corresponding absolute value needs to be calculated afterwards multiplying the answer with the background variable to gross up the number of persons employed using computers.

I.1.6.3 MODULE B: ICT SPECIALISTS AND SKILLS

B1. Does your enterprise employ ICT specialists?

Definition ICT specialists: ICT specialists are employees for whom ICT is the main job. For example, to develop, operate or maintain ICT systems or applications.

[Scope: enterprises using computers, i.e. A1 = Yes]

[Type: one single answer needed, i.e. *Tick only one*; binary (*Yes/No*)]

The purpose of this variable is to measure the incidence of ICT specialists' skills. E-skills are of two categories. E-skills for ICT specialists refer to skills for developing, operating, maintaining ICT systems or applications. Additionally, ICT specialists have the relevant skills to specify, design, install, support, manage, evaluate or perform research activities.

E-skills for other persons employed include skills on the application of systems to support own work, use of generic software tools and use of specialised tools supporting business functions within the enterprise.

Use of the term ICT specialists: In general, "specialists" are considered professionals in an ICT field e.g. database specialists, ICT support specialists, etc. The Task Force discussed using alternatively the term "ICT professionals" that would strictly correspond to the occupations' classification ISCO 08 (25 Information and communications technology professionals that are

broken down to 251 Software and applications developers and analysts, and 252 Database and network professionals).

However, the scope needs to be wider than the strict definition of ISCO 08 "ICT professionals". It was confirmed that from the users' point of view a broader scope should be covered. (See methodological manual 2012). Therefore, it is reminded that in the context of the survey the clarification "ICT is their main job" is important, moreover, because respondents would not answer the questions having in mind the specific classification. The Task Force recommended using the term "ICT specialists" that should include the following ISCO 08 codes and groups (occupations):

133 Information and communications technology service managers

25 Information and communications technology professionals

251 Software and applications developers and analysts

252 Database and network professionals

35 Information and communications technician

351 Information and communications technology operations and user support technicians

352 Telecommunications and broadcasting technicians

7422 Information and communications technology installers and servicers

(http://www.ilo.org/public/english/bureau/stat/isco/isco08/index.htm).

It is noted that in the questions on ICT specialists, we refer to persons employed by the responding enterprise, excluding those who provide their services as ICT specialists to the responding enterprise on behalf of other enterprise(s), and are either employed by the other enterprise(s) or are self-employed (outsourcing). This is in line with the definition of the "persons employed" in the Commission Regulation (EC) No 250/2009 of 11 March 2009 (p.38-39, Code: 16 11 0; Number of persons employed) that excludes "... *manpower supplied to the responding enterprise by other enterprises, persons carrying out repair and maintenance work in the enquiry [responding] unit on behalf of other enterprises...*".

B2. Did your enterprise provide any type of training to develop ICT related skills of the persons employed, during 2013?

a) Training for ICT specialists

Tick No if your enterprise didn't employ ICT specialists during 2013

b) Training for other persons employed

[Scope: enterprises using computers, i.e. A1 = Yes]

[Type: one single answer needed, i.e. *Tick only one*; binary (*Yes/No*)]

The purpose of this question is to identify those enterprises providing training to their employees either internally or using external trainers. The training refers to ICT specialist skills (if the enterprise didn't employ ICT specialists respondents should tick "No"). It is noted that response (b) refers to training aiming at developing or upgrading ICT related skills for other than ICT specialists.

B3. Did your enterprise recruit or try to recruit ICT specialists, during 2013?

[Scope: enterprises using computers, i.e. A1 = Yes]

[Type: one single answer needed, i.e. *Tick only one*; binary (*Yes/No*); filter question]

The aim of the current and the next question is to identify mismatch and shortages[9] of ICT specialists in the labour market in case the enterprise recruited or tried to recruit personnel for jobs requiring ICT specialist skills. If B3 is "No" respondents should continue with question C1.

B4. During 2013, did your enterprise have vacancies for ICT specialists that were difficult to fill?

[Scope: enterprises that recruited or tried to recruit ICT specialists during 2013, i.e. B3 = Yes]

[Type: one single answer needed, i.e. *Tick only one*; binary (*Yes/No*)]

The aim of this question is to identify mismatch and shortages in the labour market concerning ICT specialist skills considering that question B3 is "*Yes*".

For enterprises having had hard-to-fill vacancies for jobs requiring ICT specialist skills, the following reasons are possible (requested in detail in the 2007 survey but not here).

a) Lack or too low number of applicants with ICT specialist skills

Enterprise tried to recruit personnel with ICT specialist skills and found difficulties in doing so, because there was a lack of candidates to the open vacancy.

b) ICT specialists with the required skills not available or not entirely suitable

Enterprise tried to recruit personnel with ICT specialist skills and found difficulties in doing so, because it was difficult to find a candidate with the right knowledge in the ICT field the enterprise needs (cases where the qualifications of the candidates are not appropriate for the enterprise). This might be the result of lack of schools providing the proper training or the training is not good enough. One important element here is that we are referring to formal education in this item. The difference between a) and b) is that the former refers to quantity (not having people applying to the vacancy) and the latter refers to quality (there are people applying, but the qualifications are not appropriate for the enterprise). This item does not include the work experience as part of the skill of the candidates. Work experience is covered in the next item.

c) Lack of work experience in the field of ICT

Enterprise had difficulties in recruiting a person for a job vacancy, not because it lacked candidates (covered in item a) or it lacked candidates with the right education for the job, but because there were not enough candidates with the required work experience required by the job.

d) Salary requests too high

[9] Mismatch between the job seekers' competencies and the required or expected by the employers. Shortages indicate an insufficient number of skilled people in the labour market due to not enough job seekers, too low pay rates offered for them, low unemployment

Enterprise found difficulties in recruiting personnel with ICT specialist's skills, because either the remuneration required by the candidates was more than what the enterprise could afford, or it was difficult to find a candidate who would accept what the enterprise was offering.

I.1.6.4 MODULE C: ACCESS AND USE OF THE INTERNET

C1. Does your enterprise have access to the internet?

[Scope: enterprises using computers, i.e. A1 = Yes]

[Type: one single answer needed, i.e. *Tick only one*; binary (*Yes/No*); filter question]

This variable refers to the access to the internet regardless of property and purposes of the connection. The term "internet access" means having an external connection to the internet through an "internet service provider" (ISP).

An important categorisation is "fixed connections" and "mobile connections". Access via fixed line refers to the connection between two stationary points (including fixed wireless connections, e.g. public Wi-Fi, WiMax), while mobile connections are over mobile telephone networks. Another breakdown of the internet access is between temporary (dial-up on demand) and permanent connection (always on-line e.g. DSL). The technologies used to connect to the internet using "fixed connections" are "physical wire based" or "wireless" (e.g. WiFi, Bluetooth technologies). So we can divide the connections into "wire based connections" and "wireless connections".

The internet is a world-wide open computer mediated system (network) that transmits data by packet switching using a standardised internet Protocol (IP) and many other related communication protocols (internet protocol suite).

The internet is made up of thousands of smaller commercial, academic, private and government networks. It carries (provides access) to various information and communication services, such as electronic mail and the interlinked web pages and other documents of the World Wide Web. Internet refers to the following internet Protocol (IP) based networks and systems: www; Extranet over the internet; EDI over the internet; internet accessed by mobile phones; internet e-mail.

The internet protocol suite is the set of communications protocols that implement the protocol stack on which the internet runs. It is sometimes called the TCP/IP protocol suite, after the two most important protocols in it: the Transmission Control Protocol (TCP) and the internet Protocol (IP), which were also the first ones to be defined.

An "internet service provider" (ISP) is a company that provides access to the internet and related services. An ISP has the equipment and the telecommunication line access required to have a point-of-presence on the internet for the geographic area served. The larger ISPs have their own high-speed leased lines so that they are less dependent on the telecommunication providers and can provide better service to their customers. Most telecommunications operators are ISPs.

Generally, an ISP charges a monthly access fee to the consumer that varies according to maximum speed and data volume. The consumer then has unlimited access time to the internet, although the speed at which data is transferred varies widely and depends on the type of the internet connection.

C2. How many persons employed use computers with access to the internet for business purposes?

Or

Please indicate an estimate of the percentage of the total number of persons employed who use computers with access to the internet for business purposes.

[Scope: enterprises with internet access, i.e. C1 = Yes]

[Type: numerical, in absolute or percentage values]

This variable refers to persons employed who have access to the World Wide Web (www) from their workstation or from a computer to which he/she has free access. The same definition for a computer as in A1 question applies: *Computers include Personal Computers, portable computers, tablets, other portable devices like Smartphones.*

The criterion in this variable to identify internet access is the access to the World Wide Web (www). The reason for applying this specific internet service is that from the several services that can be run in the internet the www is the most common one, together with e-mail. A mere access to the internet through the e-mail service is not considered in this variable, because e-mail is already very common in Europe and it doesn't represent effectively the potential in terms of access to information as the www.

Just like with persons employed using computers (question A2) this variable can be collected in **absolute** or in **percentage** values. Eurostat in its Model Questionnaire (annex I.3.1) recommends using a combination of both, giving to the respondent the opportunity to choose to answer in absolute or percentage values.

Independently on how this variable is collected, in absolute or percentage values, the background variable "Average number of persons employed, during *the previous year*" (X2) is needed for grossing up the overall percentage of persons employed using computers connected to the World Wide Web.

Use of fixed broadband connection to the internet for business purposes

C3. Does your enterprise use DSL or any other type of fixed broadband connection to the internet? (e.g. ADSL, SDSL, VDSL, fiber optics technology (FTTH), cable technology, etc.)

[Scope: enterprises with internet access, i.e. C1 = Yes]

[Type: one single answer needed per item, i.e. *Tick only one per item*; binary (*Yes/No*); filter question]

By "connection" it is meant the type of the "last mile" connection of the enterprise (e.g. enterprise's computer-mediated network) to the network of the "internet service provider" (ISP). The last mile is the final segment between an ISP's infrastructure and a subscriber location for delivering communications connectivity to the enterprise.

Type of fixed broadband internet connection:

- DSL connection e.g. xDSL, ADSL, SDSL, VDSL, etc.

xDSL Digital Subscriber Line (DSL) is a family of technologies that provides digital data transmission over the wires of a local telephone network. DSL is widely understood to mean Asymmetric Digital Subscriber Line (ADSL), the most commonly installed technical varieties of DSL. DSL service is delivered simultaneously with regular telephone on the same telephone line as it uses a higher frequency band that is separated by filtering. A DSL line can carry both data and voice signals and the data part of the line is continuously connected.

The "Asymmetric Digital Subscriber Line" (ADSL), where more bandwidth is allocated to download than upload, and the "High-bit-rate/Data Digital Subscriber Line" (HDSL) are considered dominant DSL technologies. "Symmetric Digital Subscriber Line" (SDSL) refers to either a DSL technology that offers symmetric bandwidth for upload and download or to a particular DSL variant where data is only supported on a single line and that does not support analogue calls. "Very-high-bit-rate Digital Subscriber Line" (VDSL) is a DSL technology that offers faster data transmission. VDSL is capable of supporting e.g. high-definition television, telephone services (Voice over IP) and internet access over a single connection.

- **Fibre optics technology (FTTH), cable technology etc.**

Other high capacity "speed" fixed (wired or wireless) connection includes the following types of internet connections:

- Cable modem "cable TV network connection";
- High capacity leased lines "Frame Relay, ATM, Digital Multiplex";
- Ethernet LANs connection;
- Optical fibre connection;
- Fixed Wireless Access (FWA) connections, e.g. Satellite connection, public Wi-Fi connection, WiMax.

Concerning the Wi-Fi, essentially we are not referring to the Wi-Fi of the enterprises (that maybe xDSL, cable or fibre optics) but to public Wi-Fi, Hotspots, Hotzones that in different countries have different names. For example in Luxembourg it is called HotCity and the coverage is very large. National examples (i.e. WiFi or WiMax with their national brandnames) would help respondents tick correctly.

Cable modem uses modems attached to cable television networks (cable TV lines) for permanent "fixed" access to the internet. The term cable internet (or simply cable) refers to the delivery of internet service over this infrastructure. A cable modem is a device that enables you to hook up your PC to a local cable TV line. It is considered as one of the high capacity "speed" permanent "fixed" internet connections.

A leased line is a telephone line that has been leased for private use. In some contexts, it's called a dedicated line. A leased line is usually contrasted with a switched line or dial-up line. Leased lines are usually available at speeds of 64k, 128k, 256k, 512k, 2Mb and provided to the customer on X.21 presentation. Frame relay protocol and T-1 and T-3 (in Europe called E1 and E3) lines are used for the internet connection via leased lines. Higher speeds are available on alternative interfaces.

High capacity leased line is a permanent telephone connection between two points set up by a telecommunications common carrier. Typically, leased lines are used by businesses to connect geographically distant offices. Unlike normal dial-up connections, a leased line is always active. Because the connection doesn't carry anybody else's communications, the carrier can assure a

given level of quality. For example, a T-1 channel is a type of leased line that provides a maximum transmission speed of 1.544 Mbit/s. You can divide the connection into different lines for data and voice communication or use the channel for one high speed data circuit. Dividing the connection is called multiplexing. Increasingly, leased lines are being used by companies, and even individuals, for internet access because they afford faster data transfer rates and are cost-effective if the internet is used heavily.

Fixed wireless internet connection (FWA) is a technology which uses radio-frequency, infrared, microwave, or other types of electromagnetic or acoustic waves in place of wires, cables, or fibre optics to transmit signals or data (provide internet access) between stationary (fixed) points. It includes e.g. a satellite internet connection (long range wireless transmission) or public Wi-Fi (medium range wireless transmission).

Wi-Fi (or Wi-fi, WiFi, Wifi, wifi), short for "Wireless Fidelity", is a set of Ethernet standards for wireless local area networks (WLAN) currently based on the IEEE 802.11 specifications. New standards beyond the 802.11 specifications, such as 802.16 are currently in the works, they offer many enhancements, anywhere from longer range to greater transfer speeds. Wi-Fi was intended to be used for wireless devices and LANs, but is now often used for internet access (one of the main international standards for wireless broadband internet access and networking, with widespread use in business, homes and public spaces). It is based on radio signals with a frequency of 2.4 GHz and theoretically capable of speeds of over 54 Mbit/s. It enables a person with a wireless-enabled computer or personal digital assistant to connect to the internet when in proximity of an access point called a hotspot. Under this category it's not included the internet connection via high capacity "speed" mobile telephone networks.

- Information on **ISDN connection or dial-up access over normal telephone line is not required in the model questionnaire 2014.**

C4. What is the maximum contracted download speed of the fastest fixed internet connection of your enterprise?

[Scope: enterprises with internet access with fixed broadband connection, i.e. C3 = Yes]

[Type: one single answer needed, i.e. *Tick only one*; binary (*Yes/No*)]

By maximum contracted download speed it is meant the maximum theoretical speed according to the contractual obligations of the "internet service provider" (ISP) at which data can be downloaded. The five options offered are measured in Mbit/s (Mb/s or Mbps) that stands for *megabits per second* and is a measure of bandwidth (the total information flow over a given time) on a telecommunications medium (Mbps is not to be confused with MBps (megabytes per second. Often the problem is that speed tests and ISP's use the measure bits per second while download agents/programs use bytes per second. It is noted that 1 Byte = 8 bits).

The five options offered are: a) less than 2 Mbit/s; b) at least 2 but less than 10 Mbit/s; c) at least 10 but less than 30 Mbit/s; d) at least 30 but less than 100 Mbit/s; e) at least 100 Mbit/s.

Note that the actual bandwidth and download speeds experienced are dependent on a combination of factors including the ISP used, the equipment and the software used, the internet traffic and the destination server, hence may differ from the contracted download speed which is requested in question C4. Respondents are not requested to run a speed test on their equipment in order to provide a response to the question C4.

Indicative table of data transmission technologies (external connection to the internet) and speed. This table shows a certain number of data transmission technologies and technically possible download speeds which do not necessarily correspond to every country.

(1)	(2)					
Speed	Data Transmission Technology (in broad terms)					
(Mbps)	Narrowband *ISDN/Dialup*	DSL *ADSL*	DSL *VDSL*	Cable *DOCSIS 2.0*	Cable *DOCSIS 3.0*	Fibre optics technology *FTTH/B*
120					120 Mbit/s	120 Mbit/s
100						
50			50 Mbit/s			
40				40 Mbit/s		
30						
25		25 Mbit/s				
20					20 Mbit/s	
10			10 Mbit/s			
5						5 Mbit/s
2						
1						
<= 0,5	144 kbit/s	0,5 Mbit/s		0,5 Mbit/s		

Acronyms of the above table: ISDN: Integrated Services Digital Network; DSL: Digital Subscriber Line; VDSL: Very High Speed DSL; DOCSISData Over Cable Service Interface Specification; FTTH: Fiber To The Home, FTTB: Fiber To The Building.

Use of a mobile connection to the internet for business purposes

(Scope: enterprises with access to the internet)

- A mobile connection to the internet means the usage of portable devices connecting to the internet through mobile telephone networks for business purposes. Enterprises provide portable devices and pay for all or at least up to a limit, the subscription and the use costs.

It is noted that enterprises are not expected to provide the portable devices for the explicit purpose of business use. In principle, these are portable devices which are provided allowing the mobile connection to the internet via mobile telephone networks. In the scope of the survey are enterprises that provide portable devices as compensation or as part of the employee's equipment provided that the conditions in the introduction of the module (subscription and costs payment) are fulfilled and the portable devices allow the use of mobile connection to the internet via mobile telephone networks. Additionally, with "business use" we mean the purpose of the use related to work.

Mobile and wireless connection to the internet: The questions C5 and C6 clearly refer to "mobile ([10]) connection to the internet" through mobile telephone networks (either within or outside the premises of the enterprise) which should not be confused with "wireless connection to the internet" (e.g. Wi-Fi, Bluetooth, to a certain extent WiMAX). The exclusive use of wireless connection to the internet is excluded either within the premises of the enterprise or within the range of a hot-spot or hot-zone outside the premises of the enterprise.

From the practical point of view, portable devices can switch between mobile (telephone networks) and wireless connections (Wi-Fi, Bluetooth, and WiMax). However, the intention is to measure the take up of the "mobile use of the internet" by the enterprises if the portable devices connect to the internet over mobile telephone networks, regardless if the devices may connect to wireless networks and regardless of the technology used (3G or more advanced, but also less than 3G) for accessing the mobile telephone networks.

Portable devices: The portable devices that allow a mobile connection to the internet refer to relevant computer devices as defined in A1.

"Other portable devices like Smartphones" qualify as computers and were developed to provide functionalities beyond those of cellular phones (cell phones or mobile phones) which served only one purpose: voice communications and SMS/MMS. Smartphones and similar devices are equipped with customized software, increased bandwidth efficiency, internet access, digital cameras, portable music players, GPS functions and many more - business or entertainment - options despite their relatively small size.

However, growth in demand for portable devices requiring processors that are more powerful, abundant memory and larger screens has placed in the centre of interest low-power portable devices. Netbooks, notebooks, and tablets equipped with built-in (at least 3G technology) card modem or external dongles (USB modem), have considerably extended the family of "mobile connectivity" devices to compact size "portable computers".

The use of handsets exclusively for voice/SMS/MMS communications (e.g. enterprises requiring that persons employed are available to take business calls at all times and anywhere) are not in the scope of questions C5 and C6. Additionally, M2M (machine-to-machine) communications should not be considered in the scope of the mobile use of the internet.

In certain cases, a portable computer can be connected to a handset (e.g. via Bluetooth, cable etc.) and the handset can be connected to the internet through a mobile connection (via the mobile telephone network). We consider that the portable computer is the device connected to the internet. In this case, the handset is the equivalent to a USB modem or a card modem. Moreover, the portable computer is the device used for accessing the business software application (or any application) via the internet and not the handset.

[10] The network type referring to the "Mobile broadband technology" is the Wireless Wide Area Network (WWAN). It provides high-speed internet access through portable devices and coverage of the national territory in each country. It is considered to provide full telephone network coverage. The exclusive use of the Wireless Local Area Networks is not in the scope of questions C5 and C6.

C5. Does your enterprise use any of the following types of mobile connection (via mobile telephone networks) to the internet?

[Scope: enterprises with access to the internet, i.e. C1 = Yes]

[Type: one single answer, i.e. Tick only one; binary (Yes/No); multiple items may be expected]

a) Contracted **mobile broadband** connection via a portable device using mobile telephone networks (so called 3G or 4G) e.g. via a portable computer or other portable devices like Smartphones

A connection via **high capacity "speed" mobile telephone networks** is an access to the internet using a long range wireless transmission of the 3rd / 4th generation (3G, 4G) mobile network technologies based on the CDMA (Code Division Multiple Access) as UMTS (Universal Mobile Telephone System - Wideband "W"-CDMA); CDMA2000x; CDMA 2000 1x EV-DO; CDMA 2000 1x EV-DV).

3G (or 3-G) is short for third-generation mobile telephone technology. The services associated with 3G provide the ability to transfer both voice data (a telephone call) and non-voice data (such as downloading information, exchanging email, and instant messaging). It includes high-speed mobile networks (e.g. CDMA2000 1X, WCDMA, CDMA2000 1xEV-DO, etc.).

4G is the fourth generation of cellular wireless standards (mobile telephone technology). It is a successor of the 3G and 2G families of standards. The ITU-R organization specified the International Mobile Telecommunications Advanced requirements for 4G standards, setting peak speed requirements for 4G service at 100 Mbit/s for high mobility communication (such as from trains and cars) and 1 Gbit/s for low mobility communication (such as pedestrians and stationary users).

Universal Mobile Telecommunications System (UMTS) is one of the third-generation (3G) mobile phone technologies. It uses W-CDMA as the underlying standard, is standardized by the 3GPP, and represents the European answer to the ITU IMT-2000 requirements for 3G Cellular radio systems. It presently delivers packet switched data transmission speeds up to 384 Kbit/s and up to 2 Mbit/s when fully implemented.

CDMA2000 1x is an IMT-2000 3G mobile network technology, based on CDMA that delivers packet switched data transmission speeds of up to 144 kbps. It is also referred to as 1XRTT.

CDMA2000 1xEV-DO is an IMT-2000 3G mobile network technology, based on CDMA that delivers packet switched data transmission speeds of up to 2.4 Mbps.

a1) Mobile broadband connection via portable computer using mobile telephone networks (so called 3G or 4G)

e.g. notebook, netbook, laptop, tablet, etc. *(Optional)*

a2) Mobile broadband connection via other portable devices like Smartphone, PDA phone using mobile telephone networks (so called 3G or 4G) *(Optional)*

b) Other mobile connection

Using e.g. GSM, GPRS, EDGE

General packet radio service (GPRS) is a packet oriented mobile data service available to users of the 2G cellular communication systems global system for mobile communications (GSM), as well as in the 3G systems. In 2G systems, GPRS provides data rates of 56-114 Kbit/s. Only 2G systems are in the scope of B2e.

GPRS data transfer is typically charged per megabyte of traffic transferred, while data communication via traditional circuit switching is billed per minute of connection time, independent of whether the user actually is using the capacity or is in an idle state. GPRS is a best-effort packet switched service, as opposed to circuit switching, where a certain quality of service (QoS) is guaranteed during the connection for non-mobile users.

2G cellular systems combined with GPRS are often described as 2.5G that is a technology between the second (2G) and third (3G) generations of mobile telephony. It provides moderate speed data transfer, by using unused time division multiple access (TDMA) channels in, for example, the GSM system. Originally there was some thought to extend GPRS to cover other standards, but instead those networks are being converted to use the GSM standard, so that GSM is the only kind of network where GPRS is in use. GPRS is integrated into GSM Release 97 and newer releases. It was originally standardized by European Telecommunications Standards Institute (ETSI), but now by the 3rd Generation Partnership Project (3GPP).

Enhanced Data rates for Global Evolution (EDGE) are an intermediate technology that brings second-generation GSM closer to third-generation capacity for handling data speeds up to 384 Kbits/s. The standard is based on the GSM standard and uses TDMA multiplexing technology.

C6. How many persons employed use a portable device provided by the enterprise, that allows internet connection via mobile telephone networks, for business purposes? *(e.g. portable computers, tablets or other portable devices like Smartphones)*

Or

Please indicate an estimate of the percentage of the total number of persons employed who use a portable device provided by the enterprise, that allows internet connection via mobile telephone networks, for business purposes? *(e.g. portable computers, tablets or other portable devices like Smartphones)*

[Scope: enterprises with access to the internet, i.e. C1 = Yes]

[Type: numerical, in absolute or percentage values]

It is noted that in principle there should be a correspondence between the positive answers in question C5 (enterprises using mobile connection to the internet) and persons employed in question C6.

The concept of "persons employed" to be used in this variable is the one described in the explanatory note of variable X2 ("Average Number of Persons Employed").

This variable can be collected in absolute or in percentage values. Eurostat in its Model Questionnaire (annex I.3.1) recommends using a combination of both, giving to the respondent the opportunity to choose either to answer in absolute values or percentage values.

The variable in absolute values might be more appropriate for smaller enterprises, while the one in percentage terms for bigger ones. When the number of persons employed is small it's easier for the respondent just to identify which ones work with portable devices, instead of having to calculate the percentages afterwards. The collection in percentage terms might even introduce measurement errors in those cases, because of the rounding that is necessary to provide an answer in percentages made of integer numbers between 1 and 100. The collection in percentage terms can also introduce other measurement errors as it induces the respondent to make an educated guess of its value introducing the bias usually associated to this type of answer. On the

other hand, for bigger enterprises with a high number of persons employed, the identification of every person employed having a portable device that allows internet connection via mobile telephone networks can be burdensome or not possible and result in non-response. In this case the collection of this information in percentage terms inducing to an educated guess may be preferable.

Independently on how this variable is collected, in absolute values or percentage values, the background variable "Average number of persons employed, *during the previous year*" is needed for grossing up the overall percentage of persons employed having portable devices to access the internet via mobile telephone networks. In order to compute the overall percentage both the total number of persons employed and the total number of persons having portable devices need to be grossed up. If collected in percentage terms, the corresponding absolute value needs to be calculated afterwards multiplying the answer with the background variable to gross up the number of persons employed using portable devices to access the internet via mobile telephone networks.

Devices provided when need arises or upon request e.g. meetings away from office premises, business trips, participation to conferences, provision of maintenance and on-site support to customers etc.: The question refers to persons (not to devices), hence it is suggested to request for an estimate of the full time equivalent of persons provided with the devices. In the case that an enterprise provides portable devices (not permanently but only when needed or requested by persons employed), the alternative question would be to ask the number of (person-) days away from office in "January 2014" assuming that when employees leave the office (for business) they are provided with a portable device. Divide by 22 (conventionally 22 working days in a month) and this would be the FTE of persons employed who were provided with a portable device. The fact that the enterprise has 5 laptops (or 20 or 100) is not significant; the number of days provided is the important issue and then divided by 22 to get approximately the FTE.

Example: In January 2014, there were approx. 50 days of meetings away from the enterprise; divided by 22 and rounded it would give two FTE of persons who were provided with a portable device. It is noted that we no longer use in the questions the "January" as a reference.

Consistency between C5 and C6: In case the responding enterprises answers "yes" to any of the C5 options it is expected that C6 <>0 i.e. cannot be a real zero. In case the responding enterprise declares in C6 the number or the percentage of persons employed that are provided with a portable device (hence C6 <> 0) it is expected that there is a positive answer in at least one option in C5. In principle, it is expected to have a positive answer in C5 (a) (resp. C5 (a1) or c5 (a2)), but in rare cases that need to be verified, it could be that C5 (a) is "no", but C5 (b) is "yes". A mobile internet connection for the enterprise's owner(s) should be included in C5 and C6. Cases with C6 <>0 and C5 with no response should be checked as in certain cases corresponded to the use of private portable devices. Concluding, there should be some cross-checking of answers to questions C5 and C6 in order to ensure consistency between them.

Use of a website

C7. Does your enterprise have a website?

[Scope: enterprises with internet access, i.e. C1 = Yes]

[Type: one single answer needed, i.e. *Tick only one*; binary (*Yes/No*); filter question]

This observation variable doesn't refer specifically to the ownership of the website, but to the use of a website by the enterprise to present its "business". It includes not only the existence of a

website which is located on servers belonging to the enterprise or are located at one of the enterprise's sites, but also third party websites (e.g. one of the group of enterprises to which it belongs i.e. website of the parent company or holding company).

Retailer enterprises having a presence on the web through the company that holds the rights (royalties in the case of franchising) for the brand e.g. the parent company Bosch and the enterprise is a retailer of Bosch garden tools; Levis jeans and the retailer shops etc. The level of web presence can vary from simply an "enterprise locator" on a map (in particular for retailers e.g. find your closest Bosch tools dealer) and indirectly advertisement of the products/services to more sophisticated functionalities. As long as the presence of such enterprises on the web is through other websites for the specific products or services (not through yellow pages or telephone directories) it is considered that the enterprise has a website. Enterprises using an account on certain and specific social media (e.g. social networks as on Facebook etc.) are considered to have a website. In the latter case an enterprise is expected to have control over the contents and in general manage the site.

Enterprises on e-marketplaces where they have the possibility to advertise themselves, quote prices for ad hoc services etc. are not enterprises that are considered to have a website.

However, it doesn't include <u>any</u> presence on the web of the enterprise. That would be too broad, as it would include the presence of the enterprise (e.g. its name or its contact information) in directories and online yellow pages. These cases are not included in this variable.

Finally, this variable includes any type of website, independently of its sophistication or services provided.

It is noted that an enterprise may have e-commerce web sales and still not have a website as the sales are through e-marketplaces.

C8. Does the website have any of the following? *(C8(b) and C8(i) mandatory, C8(a, c, d, e, f, g, h) optional)*

[Scope: enterprises with a website, i.e. C7 = Yes]

[Type: one single answer needed per item, i.e. *Tick only one per item*; binary (*Yes/No*)]

For the enterprises having a website, this question focuses on the measurement of its specific expected use. This information is of particular importance, because the existence of a website provides the *potential* for better performance by the enterprise. The purpose of this question is to *complement* the measurement of *e-business*.

a) Description of goods or services, price lists *(optional)*

It includes the provision of lists of products or services offered by the enterprise to its clients which might include also the characteristics of these products and their prices. The information about the products or services may be detailed or not. The information may be static or dynamic (extracted online from a database and as such always updated).

b) Online ordering or reservation or booking, e.g. shopping cart

This item refers to a facility which allows the user to order products or services with no additional contact offline or via e-mail necessary (for the ordering). It includes also websites which allow the reservation of hotel rooms or the booking of flights. It does not include a link in the website which directs the user to an e-mail application which requires the user to send the order via e-mail.

Payment may or may not be included in the ordering facility, e.g. payment may be made on reception of the product also by other means than electronic payment.

Carrying out via online banking a transaction in general does not qualify as online ordering in case the enterprise is in the banking sector; specific cases however, e.g. when buying shares (with a commission to be paid to the bank), qualify as online orders in this sector.

c) Possibility for visitors to customise or design the products *(optional)*

This item refers to the existence of an interactive interface in the website where users can choose from several possible characteristics of the products (colour, etc.) or services and see online in the website the impact, for instance, on the price. The interface might also include the possibility for the user to visualise online the appearance of the product with the options that were selected.

d) Tracking or status of orders placed *(optional)*

This facility aims to keep the customer informed on the progress of the ordering process, like the arrival or departure of the ordered good, its location, the date and time, and the status.

e) Personalised content in the website for regular/repeated visitors *(optional)*

This facility consists on the ability of the website to recognise the user from previous visits (login/password) and adapt the content of the pages accordingly.

f) Links or references to the enterprise's social media profiles *(optional)*

This facility refers to links to social media for which the enterprise maintains a profile/an account.

Enterprises having a website may maintain a link to several social media (e.g. Facebook) with the solely purpose that the customer clicks and shares (term used on Facebook when a user posts a link with or without comments on her "timeline") the enterprise's information with others (e.g. as "friends" on Facebook). This information may appear as "shared information" on a customer's "timeline" (term used on Facebook).

There are two cases:

- The enterprise has an account/profile on the specific social, hence C8(f) should be "yes" even if the customer just "shares on her timeline" information about the enterprise (link, with/without comments)
- b) The enterprise does not have an account on the specific social media hence C8(f) should be "no" for that reason.

g) A privacy policy statement, a privacy seal or a website safety certificate *(optional)*

This item refers to information that is dealing with the use of personal information - particularly personal information collected via the website - by the website owner. It also refers to information related to measures taken to guarantee secure handling of financial information. Specific labels are usually visible on the website, for instance for making safe payments.

h) Advertisement of open job positions or online job application *(optional)*

This item includes both cases where just simple information on job vacancies is provided in the website as well as those where the website provides also an online facility for candidates to apply for the jobs.

i) Electronic submission of complaints (i.e. via e-mail, web form, etc.)

Websites may offer various electronic communication options to consumers e.g. e-mail, multi-purpose web forms, social media, blogs, etc. With this formulation enterprises are requested to

answer if their website provides the "possibility" or the "facility" to complain electronically (e.g. reference to an e-mail would qualify enterprises to answer "yes"); not a facility exclusively designed or with the only purpose to be used for submitting complaints.

Use of Social media

Enterprises using social media are considered those that have a user profile, an account or a user license depending on the requirements and the type of the social media.

Introduction: The concepts behind **social media** and the use of the relevant web2.0 tools are essentially the use of technology enabling communication, networking, social bonding and sharing e.g. blogging (opposed to personal webPages), using wikis (e.g. Wikipedia compared to Britannica online), using and participating in content management applications and web services (compared to publishing), etc.

More particularly, the term enterprise 2.0 is described as "… *the use of emergent social software platforms within companies, or between companies and their partners or customers*". Essentially, enterprise 2.0 is about developing new forms of collaboration and information management within the enterprises as well as helping employees, customers and suppliers to collaborate, to share, to innovate and to organize knowledge and experiences. Enterprise 2.0 software is developed in a socially calibrated business, aiming to reduce transactions and introduce values related to behaviour. It is considered as an evolution towards "co-evolution" of the current "competitive" environment at the "individual" level and to a certain degree at the "teams" level, within an enterprise and among enterprises.

Enterprises in the scope: The scope of the special topic comprises enterprises that at least use a profile or an account for business use, or a user license depending on the type of the social media. (e.g. Facebook requires to have profile, Twitter an account, Yammer a license, etc.). Enterprises that use social media involving customers, partners or use it within the enterprise, either testing or piloting or as part of their operations are in the scope of the survey.

Enterprises exclusively paying for posting adverts (banner ads) either directly to the owner of the social media (e.g. Facebook, Twitter, etc.) or indirectly through enterprises that provide online services for automatically posting advert messages on other social media (e.g. SoKule, etc.) are excluded from the scope.

Enterprises that have accounts, profiles or licences that are not used are not in the scope. Enterprises for which employees or other persons use social media accounts other than the enterprise's own account are excluded. Enterprises sometimes pay e.g. celebrities to mention specific products in their tweets or Facebook pages for example.

Disclaimer: References to third-party brands, products and trademarks are for the sake of clarification and are not intended to promote the use of such products

C9. Does your enterprise use any of the following social media? (not solely used for paid adverts)

[Scope: enterprises with access to the internet, i.e. C1 = Yes]

[Type: one single answer per item needed, i.e. Tick only one; binary (Yes/No); multiple items may be expected]

During the preparation of the national questionnaires countries are invited to review the examples for the (a) to (d) responses according to the "national" situation on the use of social media. It is acknowledged that the use of certain social media would require that more than one response should be ticked. In particular, the evolution of social media in the future (e.g. integration of Yammer® and Confluence®) as well as the ability of certain of those to include additional features and functionalities requires that respondents should tick more than one response e.g. the use of Yammer would qualify for a) Social network, b) microblogs and d) wiki based knowledge tools. It is noted that Google® Talk, Skype®, etc. are not considered as social media as in principle they provide the means for voice over the internet, for one-to-one communication or many-to-many conferencing among predefined contacts.

a) Social networks (e.g. Facebook®, LinkedIn®, Xing®, Viadeo®, Yammer®, etc)

Social networks or websites are applications based on internet technologies that enable users to connect by creating personal information profiles, share interest and/or activities, share ideas, invite others to have access to their profile and create communities of people with common interests.

b) Enterprise's blog or microblogs (e.g. Twitter®, Present.ly®, etc.)

A weblog or blog is a website or a part of a website that is updated frequently: up to several times a day. A blog can, either be owned by individuals, interest groups of individuals or corporate. In the current context it is the blog of the enterprise and not other blogs to which employees contribute, even if they do so to "promote" the enterprise they work for. An update can have different names, sometimes specific for the website, e.g. called an entry or a post. In most cases messages consist of text only but they can also contain images (e.g. photos), audio (podcast) or videos (vlog). Readers can, sometimes anonymously, respond, share, comment or link to the entry online. Blogs can be used either within an enterprise (corporate blog) or for communicating with external parties such as (potential) customers, business partners or other organisations.

Microblogging refers to the posting of very short updates about oneself messages (on specific microblogs websites that serve this purpose). It is in contrast to long-form blogging, where there are usually at least a few hundred words. Microblog posts usually involve a few hundred characters or less. For example, in the context of microblogging services, "Tweets" (Twitter®) are text-based posts of up to 140 characters displayed both on the user's profile page and in the "timeline" of the user's "followers". Microblogs are often used for sharing hyperlinks to other websites, including those containing videos or images/photos.

c) Multimedia content sharing websites (e.g. YouTube®, Flickr®, Picassa®, SlideShare®, etc.)

Multimedia content communities offer the possibility of sharing media content between users. Photo and video services/Podcasting: A podcast (or non-streamed webcast) is a series of digital media files (either audio or video in various file format e.g. .aiff, .wav, .midi etc., for the former and .mov, .avi etc for the latter) that are released episodically. The mode of delivery differentiates podcasting from other means of accessing media files over the internet, such as direct download, or streamed webcasting. Presentation sharing websites offer the possibility to share presentations, documents and professional videos over the internet (share publicly or privately among colleagues, clients, intranets, networks, etc.). These websites offer the possibility to upload, update and access presentations and/or documents. Very often, presentation sharing websites are linked to blogs and other social networking services or websites.

d) Wiki based knowledge sharing tools

A wiki is a website that allows the creation and editing of any number of interlinked web pages via a web browser using a simplified markup language or a WYSIWYG text editor. Wikis are typically powered by wiki software and are often used collaboratively by multiple users who can contribute to the website or edit webpages. A wiki can be open to a global audience or can be restricted to a selected network or community. Examples include community websites, corporate intranets, and knowledge management systems, Wikipedia-like sites, wikia-like sites, etc.

Other use of the internet

C10. Do any persons employed have remote access to the enterprise's e-mail system, documents or applications?

[Scope: enterprises with access to the internet, i.e. C1 = Yes]

[Type: one single answer needed, **i.e**. ***tick only one*****; binary (*Yes/No*); optional]**

Question C10 refers to the ability of an enterprise to provide remote access to the enterprise's e-mail system, documents or applications through an internet connection. It is assumed that potentially the persons employed may use remote access (not necessarily teleworking) to the enterprise's resources without being physically present. The respective data should provide information on the readiness of enterprises enabling employees to work outside the premises of the enterprise and having no need to physically travel or commute

It is noted that the question does not specify if the e-mail system (documents or applications) is a "service" delivered by the enterprise or free-of-charge by the internet provider or simply from the cloud, or a "system" that is hosted by the enterprise or by another enterprise. Therefore, as long as the enterprise's e-mail is used for **business purposes** (i.e. recognised by the enterprise, the customers, suppliers, partners) respondents should tick the "Yes" box even if the e-mail system is provided by a third party (e.g. e-mail accounts provided by national internet providers, Yahoo, Gmail, etc.).

C11. Does your enterprise pay to advertise on the internet? *(e.g. adverts on search engines, on social media, on other websites, etc.)*

[Scope: enterprises with access to the internet, i.e. C1 = Yes]

[Type: one single answer needed, **i.e**. ***tick only one*****; binary (*Yes/No*); optional]**

The question refers to enterprises using paid adverts (ads) on the internet aiming at advertising their goods or services. For example, enterprises may pay for having their ads at the top of results for relevant search terms of various search engines, on social media banners, etc.

I.1.6.5 MODULE D: USE OF CLOUD COMPUTING SERVICES

[Scope: enterprises with access to the internet, i.e. C1 = "Yes"]

Cloud computing refers to ICT services that are used over the internet to access software, computing power, storage capacity etc.;

where the services have all of the following characteristics:

- are delivered from servers of service providers

- can be easily scaled up or down (e.g. number of users or change of storage capacity)

- can be used on-demand by the user, at least after the initial set up (without human interaction with the service provider)

- are paid for, either per user, by capacity used, or they are pre-paid

Cloud computing may include connections via Virtual Private Networks (VPN)

Purpose: "**Module D: Use of cloud computing services**" was proposed as the special topic for the 2014 ICT survey in enterprises. The purpose of the respective questions is to provide information on the use of a service model for the ICT provision i.e. software, computing power, storage capacity. Cloud computing services should not be considered as outsourcing of business functions. The business model of cloud computing relies on economies of scale.

Introduction: For the convenience of the reader and in order to allow a better understanding of cloud computing, the following information is excerpted from NIST Special Publication 800-145. Overall, the characteristics for cloud computing services are those used in the context of the ICT usage and e-commerce survey for 2014.

Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model is composed of five essential characteristics, three service models, and four deployment models.

Essential Characteristics:

On-demand self-service. A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service's provider.

Broad network access. Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, tablets, laptops, and workstations).

Resource pooling. The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. There is a sense of location independence in that the customer generally has no control or knowledge over the exact location of the provided resources but may be able to specify location at a higher level of abstraction (e.g., country, state, or datacenter). Examples of resources include storage, processing, memory, and network bandwidth.

Rapid elasticity. Capabilities can be rapidly and elastically provisioned, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be appropriated in any quantity at any time.

Measured Service. Cloud systems automatically control and optimize resource use by leveraging a metering capability[11] at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service.

[11] Typically this is done on a pay-per-use or charge-per-use basis.

Service Models:

Cloud Software as a Service (SaaS). The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure[12]. The applications are accessible from various client devices through a thin client interface such as a web browser (e.g., web-based email), or a program interface. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.

Cloud Platform as a Service (PaaS). The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or -acquired applications created using programming languages and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly application hosting environment configurations.

Cloud Infrastructure as a Service (IaaS). The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, deployed applications; and possibly limited control of select networking components (e.g., host firewalls).

Deployment Models:

Private cloud. The cloud infrastructure is provisioned for exclusive use by a single organization comprising multiple consumers (e.g., business units). It may be owned, managed, and operated by the organization, a third party, or some combination of them, and it may exist on or off premises.

Community cloud. The cloud infrastructure is provisioned for exclusive use by a specific community of consumers from organizations that have shared concerns (e.g., mission, security requirements, policy, and compliance considerations). It may be owned, managed, and operated by one or more of the organizations in the community, a third party, or some combination of them, and it may exist on or off premises.

Public cloud. The cloud infrastructure is provisioned for open use by the general public. It may be owned, managed, and operated by a business, academic, or government organization, or some combination of them. It exists on the premises of the cloud provider.

Hybrid cloud. The cloud infrastructure is a composition of two or more distinct cloud infrastructures (private, community, or public) that remain unique entities, but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load balancing between clouds).

[12] A cloud infrastructure is the collection of hardware and software that enables the five essential characteristics of cloud computing. The cloud infrastructure can be viewed as containing both a physical layer and an abstraction layer. The physical layer consists of the hardware resources that are necessary to support the cloud services being provided, and typically includes server, storage and network components. The abstraction layer consists of the software deployed across the physical layer, which manifests the essential cloud characteristics. Conceptually the abstraction layer sits above the physical layer.

Enterprises in the scope: The scope of the special topic comprises **enterprises with access to the internet**. Enterprises that are in the process of evaluating, testing or piloting the use of cloud computing services, or enterprises that are using free of charge cloud computing services **are not considered as users of cloud computing in the context of this survey**. Questions D1, D2, D3, D4 and D5 (optional) are addressed to users of cloud computing services. Question D6 (optional) is addressed to enterprises not using cloud computing services (D1="No"). It is noted that in the model questionnaire we refer to all deployment models of cloud computing.

Routing and filter question: Routing requires adaptation according to the choices of the optional questions that will be included in the national questionnaires.

8 paths will be possible (C1: access to the internet = Yes)

No optional questions

1) First path without any optional: D1 (Yes), D2, D3, D4, E1 …

2) Second path without any optional: D1 (No), D1 filter instruction go to E1

Only D5 optional

3) Third path with only D5 optional: D1(Yes), D2, D3, D4, D5, E1

4) Fourth path with only D5 optional: D1(No), D1 filter instruction go to E1

Only D6 optional

5) Fifth path with only D6 optional: D1(Yes), D2, D3, D4, instruction go to E1 after the D4

6) Sixth path with only D6 optional: D1(No), D1 filter instruction go to D6

Both optional D5, D6

7) Seventh path with both D5, D6 optional: D1 (Yes), D2, D3, D4, D5, instruction go to E1 after the D5

8) Eighth path with both D5, D6 optional: D1 (No), D1 filter instruction go to D6

Disclaimer: References to third-party brands, products and trademarks are for the sake of clarification and are not intended to promote the use of such products

D1. Does your enterprise buy any cloud computing services used over the internet?

(Please refer to the definition of cloud computing above, exclude free of charge services)

[Scope: enterprises with access to the internet, i.e. C1 = "Yes"]

[Type: one single answer needed, i.e. Tick only one; binary (Yes/No); filter question]

The question aims at measuring the take up of using IT cloud computing services, either private or public. Users requested information in order to approximate the total "volume" of cloud computing services including both deployment models. It is noted that respondents are requested to identify their own model(s) of cloud computing in question D3, after identifying the services that they use. The difference lies in the way that the services are **"delivered from"** i.e. a) shared servers of service providers or b) servers of service providers exclusively reserved for the enterprise or both (a) and (b) (multiple responses are allowed).

Private cloud computing implies that dedicated servers are used in order to store data exclusively for an enterprise and depending on the case, the exclusive use of customised software on these

servers. It is additionally noted that in the context of the ICT survey in enterprises, "community cloud computing" may be considered in a wider context as "private".

The use of connections via VPN (Virtual Private Network) means the use of secured communication over the internet and includes enterprises using cloud computing services provided that all the other conditions (box above) are fulfilled.

In all cases, respondents should consider that all criteria of the definition should be fulfilled. particularly concerning the first condition that **services are delivered from servers of service providers**; hence **a)** enterprises providing cloud computing services (as their main economic activity), and **b)** enterprises using their own cloud (having perhaps a different economic activity) should answer "No" in question D1.

D2. Does your enterprise buy any of the following cloud computing services used over the Internet? *(Please refer to the definition of cloud computing above, exclude free of charge services)*

[Scope: enterprises using cloud computing services, i.e. D1 = Yes]

[Type: one single answer per item needed, i.e. Tick only one; binary (Yes/No); multiple items may be expected]

It is noted that for enterprises to qualify for using the following cloud computing services the relevant conditions that are mentioned in the introduction of the module must **all be fulfilled to the extent that they are applicable**. E.g. for enterprises using e-mail as a cloud computing service they must use the servers of the service providers, adding e-mail addresses and increasing mailbox storage capacity can be easily done, and the service is priced according to the number of e-mail addresses and the mailbox storage capacity that is used (or pre-paid for a specific number of e-mail addresses and the respective mailbox storage capacity).

a) E-mail (as a cloud computing service)

For example *Gmail Enterprise®, Microsoft® Exchange Online/Office 365, etc.*

b) Office software (e.g. word processors, spreadsheets, etc.) (as a cloud computing service). For example *Microsoft® Office Cloud, etc.*

c) Hosting the enterprise's database(s) (as a cloud computing service).

For example *EnterpriseDB®, LongJump®, Elastra®, etc.* This response includes as well the respective database functionalities to store, search, retrieve etc. information.

d) Storage of files (as a cloud computing service) For example *Dropbox®, Amazon® S3, EMC Mozy®, Acronis® Online, Diino®, etc.*

This response refers to the storage of any type of files that are physically stored on some media; storage functionality is supplied as service over the internet (e.g. documents, images, sounds, presentations, etc.). From the service point of view this response includes storing backup files and restoring them if needed (disaster recovery).

e) Finance or accounting software applications (as a cloud computing service) For example *StepStone®, Hubwoo®, SAP® Business ByDesign, Twinfield®, Concur®, Netsuite®, etc.*

f) Customer Relationship Management (CRM, software application for managing information about customers) (as a cloud computing service) For example *salesforce.com®, Oracle® CRM On Demand, etc.*

g) Computing power to run the enterprise's own software (as a cloud computing service) For example *Amazon® EC2, Flexiscale®, Joyen®, etc.*

This response essentially refers to enterprises using computing power (as a cloud computing service) for running the enterprise's software application. The service may be provided either as Software as a service (SaaS) or as Infrastructure (hardware/software) as a Service (IaaS). For the former the enterprise (end-users) uses the software application that has been developed and it is accessible and used through a browser, for the latter the enterprise additionally maintains control of the software environment e.g. add new modules for the software application, put whatever software on the cloud, etc.). To a certain extent this response may overlap with some of the above, however, it covers a wider scope as it concerns enterprises running their software using the computing power of the cloud. These enterprises do not have to buy servers, maintain them etc.

The above list of ICT cloud computing services is not exhaustive but it is expected to cover the majority of the existing services. Cloud computing services may refer to basic functions (storage of files, hosting of enterprise's database) or more sophisticated functions e.g. CRM. The aim is to combine the above requested information in order to evaluate the level of dependence and integration in the enterprises' operations. Services (a), (b) and to a certain extent (d) indicate a low dependence. Service (c) indicates medium/high dependence. Services (e), (f) and (g) indicate a high degree of dependence on cloud computing and integration of the relevant services in the enterprises' business functions.

For information, other cloud computing services that are not mentioned in the above list may be *System and network management* often called Management-as-a-Service MaaS (e.g. Service-Now.com®, ZenDesk®, etc.); paid use of *social media*; *security application* as service over the internet (e.g. Google®/Postini, Symantec®/Messagelabs); *Content management* (e.g. Clickability®, CrownPeak®, OmniUpdate®); *Personnel, Human Resources-HRl, Human Capital Management-HCM or Talent Management* (e.g. Taleo®, Successfactors®); *Business intelligence and analytics* (e.g. IBM® Smart Analytics Cloud, SAP® BusinessObjects On Demand); *application development* (e.g. force.com®, Microsoft® Azure, Servoy) and/or online testing etc.

Simplified examples, provided only for illustration purposes, for the three main service models that are used in the definition of NIST are:

SaaS is something like Gmail where an enterprise gets its software application (in this case e-mail) via web browser access. With **IaaS** an enterprise has additionally control over the software as with Amazon web Services[13]. With **PaaS**, an enterprise may get to create apps within a certain environment. An example would be Facebook, which lets you create apps to plug into the Facebook platform. Service providers offer many different cloud computing solutions that are difficult to strictly categorise according to the definition.

Validation rules should apply for checking the existence of CRM or other characteristics that are requested in other modules of the questionnaire.

[13] Amazon Web Services offers a complete set of infrastructure and application services that enable an enterprise to run virtually everything in the cloud: from websites and mobile apps, to big data projects and enterprise applications

D3. Does your enterprise buy any cloud computing services delivered from: *(Please refer to the definition of cloud computing above, exclude free of charge services)*

[Scope: enterprises using cloud computing services, i.e. D1 = Yes]

[Type: one single answer per item needed, i.e. Tick only one; binary (Yes/No); multiple items may be expected]

a) **shared servers** of service providers

b) **servers** of service providers **exclusively reserved for your enterprise**

Response (b) clearly refers to infrastructure exclusively reserved for an enterprise and according to the above "a single organisation comprising multiple consumers (e.g. business units)". It is reminded that the difference with public cloud computing services is that the **cloud infrastructure is provisioned for exclusive use** i.e. dedicated servers used by the enterprise. As mentioned above, in the context of the ICT survey in enterprises, "community cloud computing" may be considered in a wider context as "private".

D4. Do any of the following factors limit your enterprise from using cloud computing services? *(Please refer to the definition of cloud computing above, exclude free of charge services)*

[Scope: enterprises using cloud computing services, i.e. D1 = Yes]

[Type: one single answer per item needed, i.e. Tick only one; binary (Yes/No); multiple items may be expected]

As source for the explanations below we used material from "Quantitative Estimates of the Demand for Cloud Computing in Europe and the Likely Barriers to Take-up, SMART 2011/0045, D2-Interim Report, IDC 24/2/2012.

The question is only addressed to users of cloud computing services

a) Risk of a security breach

Response refers to failures that may be due to service providers' lack of provisions around enterprise's system/data protection or privacy. System/data protection and privacy refer to ensuring that enterprises' data are processed according to the enterprise's standards and used exclusively by the enterprise for the purposes of the enterprise.

In a wider context service providers should take all possible measures defining and transparently applying procedures related to possible security breaches to system and services intended for the enterprise. The issue should as well be seen in the context of "accountability" for service providers. Accountability means providing clear rules about cloud service providers' accountability and liability (especially around security), regardless of the country of origin.

Many cloud service providers (typically, almost all those offering public cloud services) offer a "best efforts" service. In other words, they use their track records in availability and resilience to possible security breaches as evidence that they will continue to provide services continuously and reliably to their customers.

Their argument is that their services are typically far more reliable and secure than their customers' own systems. There is a lot of justification in this argument – the market determines that an unreliable, insecure cloud service would rapidly lose all its customers and be unable to secure new customers. Nevertheless, there are a number of customers who would like to see more

direct guarantees from their service providers, and who may already use the cloud but will not move certain types of systems to the cloud until they receive the accountability and liability guarantees they require.

b) Problems accessing data or software

Response refers to failures due to hardware failures (connectivity problems or downtime e.g. unavailability of the system failing to provide or perform its function), software errors or human errors. It implies limited or no access to the data or the software according to the service providers' provisions, commitments and service agreements. In this context data availability refers to data that are present and ready for use; obtainable at all times.

c) Difficulties in unsubscribing or changing service provider (including concerns with data portability)

One of the key issues of cloud computing is the ability to adopt and remove the service either at the point of use (to scale up and down as needed) or to make choices to use new services or change service provider. However, migration between providers of cloud services may still be a challenge. There are portability issues referring to moving data among different cloud computing application programs, computing environments provided as a service in the context of the current survey. Moreover, enterprises may become dependent on their current service provider because the latter hosts on the cloud servers the corporate and customers' data. From that point of view one may consider that there is a trade-off between the opportunities related to speed, cost, flexibility and scalability of resources and services from one cloud service provider versus the flexibility and choice of alternative services and cloud solutions from another provider. Data portability may become an important issue considering the volume of data that enterprises may store in the cloud (in the context of the current survey).

d) Uncertainty about the location of the data

Cloud services are often hosted in one country and consumed in others. Enterprises may consider a risk the issue of where the data are located (data residency). The data centres used by cloud computing service providers may be scattered around the globe for a number of reasons. For example, operation cost and particular requirements for the operation of the hardware that they use (requirements for electricity, low temperatures for cooling off the systems, water supply etc.).

e) Uncertainty about applicable law, jurisdiction, dispute resolution mechanism

Linked to the response (d) or for other reasons that may result to a dispute between the service provider and the customer, there is a question of where legal jurisdiction over these services resides. These issues need to be addressed at multiple levels – locally within countries, at the EU level, and at an international level since data from EU countries is often stored outside the EU, for example either as a working copy or as a back-up copy of the enterprises' data.

f) High cost of buying cloud computing services

One of the advantages of buying services on the cloud is that operation costs can be easily scaled up or down (e.g. number of users or change of storage capacity), can be used on-demand by the user and as pay-as-you-use services. It is expected that these services are less expensive than the ones provisioned from enterprises' own data centres taking into account the purchase of hardware, the cost for licencing and versioning of software, the maintenance cost and the cost for backup, ICT planning and management etc. However, the initial start-up may require higher investment due to a transition period (both systems running in parallel) and in general raised costs in short term. Moreover, running costs can be disproportionally high in case an enterprise uses continuously available open instances (on-demand computing power and storage) under a regular

pricing per hour. Another pricing model proposed by service providers may include a subscription to cloud computing services (pre-paid at better prices) with the drawback to be committed to providers for years without a way to change/leave without an additional costly investment.

g) Insufficient knowledge of cloud computing

Similar to the use of other ICT, lack of knowledge or expertise may hinder the take-up of the relevant technologies. In particular for small/medium enterprises knowledge or expertise on technical aspects of cloud computing is important. Additionally, the use of cloud computing services may require special ICT management skills (e.g. evaluation of the ICT needs, new ICT management tools for controlling consumption) and legal support (e.g. legal jurisdiction) even before the decision for buying cloud computing services is taken.

D5. To what degree were any of the following benefits realised from using cloud computing services? *(Please refer to the definition of cloud computing above, exclude free of charge services) OPTIONAL*

[Scope: enterprises using cloud computing services, i.e. D1 = Yes]

[Type: one single answer per item needed, i.e. Tick only one; Scale (To a high degree/To some degree/To a limited degree/Not at all**); multiple items may be expected]**

The responses refer to the experiences of the enterprises using cloud computing services with regard to realised benefits. A scale is proposed that corresponds to the degree of realisation of the benefits in the following three responses. The scale essentially corresponds to a three-point ordinal scale for High-Medium-Low (with the addition of "Not at all").

a) Reduction of ICT related costs

b) Flexibility due to scaling cloud computing services up or down

c) Easy and quick deployment of solutions based on cloud computing

D5 a) is related to the reduction of costs and expenses (thinking of what the alternatives would be), response b) is related to addressing the varying needs of enterprises and response c) is related to easy and quick deployment (similar to "plug and play" without worrying about settings and configuration elements).

For response b) the issue is if and to what degree cloud computing services offered adequate (increased or decreased) resources e.g. computational power, licences, platforms for development, etc. when an enterprise needed them. Of course "scaling of cloud computing services" can be reflected in a) and c) but on the one hand a) means reduction is expenses and on the other hand c) means "plug and play" solutions.

D6. Do any of the following factors prevent your enterprise from using cloud computing services? *(Please refer to the definition of cloud computing above, exclude free of charge services) (optional)*

[Scope: enterprises with access to the internet, i.e. C1 = "Yes" AND enterprises <u>not</u> using cloud computing services, i.e. D1 = No]

[Type: one single answer per item needed, i.e. Tick only one; binary (Yes/No); multiple items may be expected; optional]

The question is only addressed to non-users of cloud computing i.e. D1="No"

The following responses refer to a restricted selection of possible factors that may be considered as preventing the use of cloud computing services (according to the introduction) for enterprises that are not users (in the context of the ICT survey).

A negative answer to all the possible factors would be acceptable and interpreted as "other factors or reasons not mentioned here". In the latter interpretation it may as well be enterprises that are in the process of evaluating, testing or piloting the use of cloud computing services, or enterprises that are using free of charge cloud computing services, and **are not considered as users of cloud computing services in the context of this survey**.

The respective explanations provided in **D4** should apply.

a) Risk of a security breach

b) Uncertainty about the location of the data

c) Uncertainty about applicable law, jurisdiction, dispute resolution mechanism

d) High costs of buying cloud computing services

e) Insufficient knowledge of cloud computing

I.1.6.6 MODULE E: SHARING INFORMATION WITHIN THE ENTERPRISE

This module measures the internal integration of business functions within the enterprise (scope: enterprises with computers).

Electronic and automated sharing of information between different functions of the enterprise means any of the following:

- data linking between the software applications that support the different functions of the enterprise;
- using a common database or data warehouse accessed by the software applications that support the different functions of the enterprise;
- within this enterprise, sending or receiving electronically information that can be processed in automated way.

The above are expected to be included in the use of one single ERP (Enterprise Resource planning) software application that supports the different business functions of the enterprise.

E1. Does your enterprise use an ERP software package?

[Scope: enterprises using computers, i.e. A1 = Yes]

[Type: one single answer needed per item, i.e. *Tick only one per item;* binary *(Yes/No*)]

An ERP (Enterprise Resource Planning) is a software package used to manage resources by sharing information among different functional areas (e.g. accounting, planning, production, marketing, etc.)

ERP stands for Enterprise Resource Planning and consists of one or of a set of software applications that integrate information and processes across the several business functions of the enterprise. Typically ERP integrates planning, procurement, sales, marketing, customer

relationship, finance and human resources. Enterprises are considered using an ERP software package if they are using all or only one set of software applications (modules).

ERP systems have the following characteristics:

1. ERP systems are software designed for the enterprises server environment, whether traditional or web/html based.
2. ERP systems integrate the majority of a business's processes.
3. ERP systems process a large majority of an organization's transactions.
4. ERP systems use an enterprise-wide database that typically stores each piece of data once.
5. ERP systems allow access to the data in real time.

ERP software can be customised software, custom built for or built within the enterprise, but there is also package software designed to perform this function. However, this ERP package software is usually built is a modular way allowing enterprises to customise it for their specific economic activity or size, by implementing only some of those modules.

The main relevance of the measurement of ERP adoption is the integration of business processes and the consequent impact on the productivity of the enterprise.

Experience in other surveys shows that ERP is easy to ask. However, it is not necessarily easy to understand by the respondents.

It is argued that if respondents do not know what ERP is, then they do not use it (just like for most ICT related questions). However, this opinion is not consensual. It depends on what one specifically means by ERP: a software package or the concept of generalised integration of business processes. In the second case it is possible that an enterprise is in such a situation, but doesn't recognise it as ERP. Mentioning of examples of nationally used ERP software packages might also help to recognise the use of ERP.

This question is intended to measure the use of ERP software packages and not of ERP as a concept of generalised integration of business processes.

Questions in previous model questionnaires attempted to measure the integration of business processes (internal and external) looking at the link between the front-office (in contact with customers and suppliers respectively) and other specific business functions: management of inventory levels, accounting, production or services management and distribution management. In the current model questionnaire question E1 is intended to measure the adoption of an ERP software package.

The difference between these two approaches is that with an ERP package the enterprise is supposedly integrating its business processes, but it doesn't provide information about to what extent and which processes it is integrating. As an ERP package is usually made of several modules, the enterprise might be making a very restrictive use of it. On the other hand the enterprise doesn't necessarily need an ERP package to integrate. If the enterprise is small the use of normal office software can provide this integration.

The adoption of ERP software packages is also important because there are simple (free/open source) solutions which might be a good opportunity for enterprises to increase their internal integration of business processes.

Among other functions ERP is used also for sharing information with the marketing area. In the pretesting project issue was raised if CRM is considered as part of ERP. Very basic CRM-type tools (contact information, information on sales, etc.) are a part of many ERP systems core modules, but a CRM to really manage business interactions and communications is typically a separate piece

of software. CRM packages can be made available from ERP software providers (and these CRM systems can be integrated with ERP). Many enterprises use stand-alone CRM software alongside ERP. Thus the decision if CRM is part of ERP should be left to respondents. But no editing of data is recommended if ERP question is answered with 'No' and CRM with 'Yes'.

E2. Does your enterprise use CRM software to manage:

a) The collection, storing and making available information about customers to various business functions

b) The analysis of information about customers for marketing purposes. (E.g. setting prices, sales promotion, choosing distribution channels, etc.)

[Scope: enterprises using computers, i.e. A1 = Yes]

[Type: one single answer needed per item, i.e. *Tick only one per item;* binary *(Yes/No*)]

CRM (Customer Relationship Management) refers to any software application for managing information about business' customers.

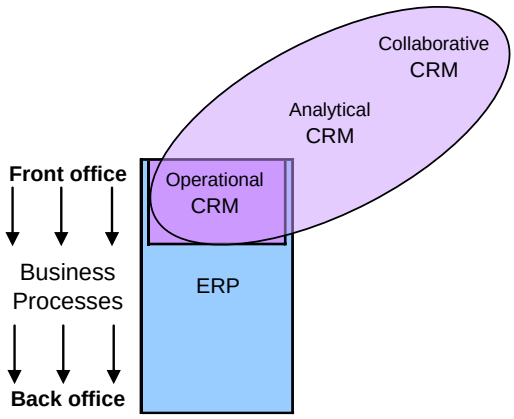
CRM consists of a management methodology which places the customer at the centre of the business activity, based in an intensive use of information technologies to collect, integrate, process and analyse information related to the customers.

CRM can be considered to be composed of three elements:

1. Operational CRM – Integration of the front office business processes that are in contact with the customer.
2. Analytical CRM – Analysis, through data mining, of the information available in the enterprise on its customers. This aims to gather in depth knowledge of the customer and how to answer to its needs.
3. Collaborative CRM – Adoption of a set of practices aiming to facilitate interaction with the customer, such as after sales support. Collaborative CRM means customer service where the customer and the company agent communicate in real-time with the aid of ICT. So web co-browsing solutions (where the agent and the customer browse together on the customer's desktop), chat, instant messaging, and various forms of application or desktop sharing can all be considered "Collaborative CRM".

From these three elements it is evident that CRM is broader than just ICT adoption and use. It includes also the creation of new business processes, like data mining on customer information and after sales support. It can be even understood in a broader sense including the restructuring of all the enterprise's business processes in a customer-centric way.

Therefore, CRM is broader in concept than ERP as it includes more than just integration of business processes. ERP on the other hand is broader in extension as it includes the integration of the majority of the business processes, front office and back office, for the majority of the transactions, and not only those related to customer relationship.



Although CRM is more than adoption of ICT it is based on an intensive use of these technologies. It is believed that the adoption of CRM software improves the marketing functional area of the enterprises, by improving customer service and customer relationship. Improvement comes, for instance, from providing user-friendly mechanisms for receiving complaints, helping identify potential problems before they occur, in general, by facilitating communication with the customer and help track customer interests, purchasing habits and product use. When these technology driven improvements lead to long-term customer satisfaction, they can ensure increased customer loyalty, decreased customer turnover, decrease marketing costs, increase sales revenues and increase profit margins.

Therefore the approach adopted in this variable is to measure the adoption of CRM software tools.

The question is restricted to the first two meanings of CRM, operational and analytical CRM. Collaborative CRM seems not to be a term and it is relevant mostly for manufacturing and distribution sectors. Having both first two meanings assembled in one single question would make it too complex. Therefore, it should be split in two, one focusing on operational CRM and another on analytical CRM (items a) and b) of the question).

Note 1: There is a connection between the question E2a and questions E2b. Enterprise must first collect, store the data before they can be e.g. used for analysis for marketing purposes. Thus logical controls should be foreseen in the process of data validation to prevent inconsistencies.

Note2: There is a connection between questions E2a and D2f. Enterprises may use CRM as a cloud computing services. Thus logical controls should be foreseen in the process of data validation to prevent inconsistencies.

I.1.6.7 MODULE F: USE OF EDI-TYPE MESSAGES SUITABLE FOR AUTOMATED PROCESSING

Electronic transmission of data suitable for automated processing - often called **Electronic Data Interchange** (EDI) - means:

– **sending and/or receiving of messages** (e.g. payment transactions, tax declarations, orders, etc.)

– in an **agreed or standard format** suitable for **automated processing**,

E.g. EDI, EDIFACT, XML, xCBL, cXML, ebXML, ODETTE, TRADACOMS,

– **without the individual message being typed manually**

[please add national examples]

F1. Does your enterprise send or receive EDI-type messages suitable for automated processing for?

[Scope: enterprises using computers, i.e. A1 = Yes]

[Type: one single answer needed, i.e. *Tick only one*; binary (*Yes/No*)]

a) Payment instructions to financial institutions

b) Data to/from public authorities (e.g. tax returns, statistical data, import or export declarations [please add national examples])

Reference to the Machine-2-Machine (M2M) communication aspect of electronic transmission of data suitable for automated processing and national examples would help respondents recognise this form of external ICT integration into their business functions.

a) Payment instructions to financial institutions

A payment instruction, also named payment order, is a business message sent by the enterprise to a bank where it has an account instructing the bank to debit its account by a certain amount and to arrange for its payment to a certain beneficiary. The beneficiary can be a supplier, an employee or any other entity. Home banking/internet banking is outside the scope of the electronic transmission of data suitable for automated processing as it entails the use of web forms and typing manually the individual messages.

b) Data to/from public authorities (e.g. tax returns, statistical data, import or export declarations, [national examples], etc.)

Neither of the responses refers to filling in web forms and typing manually the individual messages.

I.1.6.8 MODULE G: SHARING SUPPLY CHAIN MANAGEMENT INFORMATION ELECTRONICALLY

This module intends to measure the level of integration of business processes of an enterprise with those of its suppliers and/or customers. It focuses on processes related to the supply chain management (SCM). Supply chain management can be defined as the "design, planning, execution, control, and monitoring of supply chain activities with the objective of creating net value, building a competitive infrastructure, leveraging logistics, synchronizing supply with demand and measuring performance. In module G the scope of business processes is studied more narrow – the coordination of availability and delivery of products or services – while excluding only manually typed e-mail messages as carrier of the electronic exchange of information.

Sharing information electronically on the supply chain management means:

- exchanging all types of information with suppliers and/or customers in order to coordinate the availability and delivery of products or services to the final consumer;
- including information on demand forecasts, inventories, production, distribution or product development;
- via computer networks between computers of different enterprises.
- it can be from your enterprise to its suppliers/customers or the other way around.

This information may be exchanged via websites or other means of electronic data transfer but it excludes manually typed e-mail messages.

G1. Does your enterprise share supply chain management information electronically with its suppliers or customers?

[Scope: enterprises using computers, i.e. A1 = Yes]

[Type: one single answer needed, i.e. *Tick only one;* binary *(Yes/No*); filter question]

This question intends to measure the electronic exchange of SCM information with suppliers, customers or both. For example sharing of information with **suppliers** or **customers** may refer to inventory levels, production plans, demand forecasts or progress of deliveries (i.e. distribution of raw materials or finished products).

G2. How does your enterprise share supply chain management information electronically?

[Scope: enterprises that share electronically supply chain management information with their suppliers or customers, i.e. G1 = Yes]

[Type: one single answer needed per item, i.e. *Tick only one per item;* binary *(Yes/No*)]

The items in this question allow measuring the technology used for the sharing of information electronically; but, sharing information via automated data exchange implies a further integration with systems of suppliers and/or customers (agreed format).

a) via websites (yours, those of your business partners) or web portals

b*)* via electronic transmission suitable for automated processing (e.g. EDI-type systems, XML, EDIFACT, etc.)

I.1.6.9 MODULE H: USE OF RADIO FREQUENCY IDENTIFICATION (RFID) TECHNOLOGIES

Radio Frequency identification technologies (RFID) means:

- an automated identification method to store and remotely retrieve data using RFID tags or transponders,

An RFID tag is a device that can be applied to or incorporated into a product or an object and transmits data via radiowaves.

[please add national examples]

Radio Frequency identification (RFID) is an automated identification method, relying on storing and remotely retrieving data using devices called RFID tags or transponders. An RFID tag is an object that can be applied to or incorporated into another object (e.g. a good, a person) for the purpose of identification using radio waves. Some tags can be read from several meters away and beyond the line of sight of the reader.

H1. Does your enterprise make use of Radio Frequency Identification instruments for the following purposes?

[Scope: enterprises with computers, i.e. A1 = Yes]

[Type: one single answer needed per item, i.e*. tick only one per item*; binary (*Yes/No*)]

It should be noted that the use of tags could cover several purposes at the same time.

a) Person identification or access control

RFID tags can be used for building access control. Such tags are widely used in identification badges, replacing earlier magnetic stripe cards. These badges need only to be held within a certain distance of the reader to authenticate the holder.

b) As part of the production and service delivery process (e.g. monitoring and control of industrial production, supply chain and inventory tracking; service , maintenance or asset management, etc.)

RFID can be used in supply chain management and to improve the efficiency of inventory tracking and management. *Business partners are able to share information about goods and end to end through the supply chain and to identify the current location and status of items (O'Reilly "RFID Essentials" by Bill Glover & Himanshu Bhatt)*. RFID combined with mobile computing and web technologies provide an effective way for organizations to identify and manage their assets. Web based management tools allow organizations to monitor their assets and make management decisions from anywhere in the world. Web based applications now mean that third parties, such as manufacturers and contractors can be granted access to update asset data, including for example, inspection history and transfer documentation online ensuring that the end user always has accurate, real-time data.

c) For product identification after the production process (e.g. theft control, counterfeiting, allergen information, etc.)

The purpose of this item is to allow business partners and customers to distinguish RFID-tagged "original" goods from illegally made copies. It also allows for the control on theft of such goods. RFID tags give also opportunity to address consumer's concerns on well-being for example by warning/information on allergen contents.

The responses are meant to measure the specific purposes of the enterprises using RFID a) for the personnel, staff identification b) product identification during the production and service delivery comprising logistics and transportation and c) product identification after production and/or service delivery comprising logistics and transportation in the following cases given in the examples:

- producing enterprise theft control in its premises;

- other than the producing enterprise, retail sales enterprises for theft control (products going through security gates);

- the producing enterprise or any other enterprise with retail sales may check the originality of the products in all cases; for example when the retailer buys the product from the producer, when a producer or retailer is requested to identify the product or when a consumer wants to identify the product before she buys the product or after she buys the product;

- potential retailers, other organisations, institutes or consumers can be informed of allergen contents.

It is noted that one of the most common use at the moment (after person identification and access control) is in logistics or transports. These are often after sales and after production in order to follow the products. It often is not on single item but on packages of them or loads of them. In some situations RFID in logistics is about service delivery process b), e.g. when the enterprise is the logistics/transport firm and is following the movement of the things they transport etc. But on other cases where the items are followed by the producer they are basically sold and then followed during the logistics chain. In the latter they should be considered as well under the (b) response.

I.1.6.10 MODULE I: ELECTRONIC INVOICING

The purpose of this module is to provide quantitative information on the use of electronic invoices. An electronic invoice is an electronic transaction document that contains billing information.

There are invoices in paper form and electronic form. Invoices in electronic form are of two types:

- eInvoices in a standard structure suitable for automated processing.

(e.g. EDI, UBL, XML, [please add national examples]). They are exchanged either directly or via service operators or via an electronic banking system.

- Invoices in electronic form not suitable for automated processing.

(e.g. e-mails, e-mail attachment as pdf, images in TIF, JPEG or other format)

E-invoices may be converted automatically into a payment by the client but the definition would exclude direct debiting or bank transfer from the client to the vendor without any exchange of invoice.

The condition for automated processing is that the e-invoice is sent in a standard or at least agreed format. This can be EDI (Electronic Data Exchange) or the more modern UBL (Universal Business Language). In case of UBL, the messages are encoded in XML (Extensible Markup Language). There are XML implementations for EDI, too. EDI-invoices are often sent via service operators.

To help respondents in supplying the quantitative information ***it is indicated that the percentages could be indicated with certain approximation.***

Special Case: Enterprise **A** is a retailer that (possibly) sells over-the-counter to other enterprises. **A** person of enterprise **B** visits enterprise **A** and places an order for goods on behalf of enterprise **B**. The salesperson of Enterprise **A** issues and hands over the invoice to the person from enterprise **B**. Payment and delivery of goods or provision of services are not relevant in this case.

Clarification: Assuming that enterprise **A** responds to our survey we expect that the respondent will consider this case as "invoices only in paper form" (I2 (c)) when answering our questionnaire although the invoice was not literally "sent". It is clarified that in the questions on sending Invoices to other enterprises/public authorities (questions I1, I2 of the 2014 model questionnaire), we refer to enterprises that invoiced other enterprises or public authorities and might have sent the invoices by post (obviously in paper form), electronically (either suitable for automated processing or not) or in rare cases handed over in paper form. Therefore, "sending" invoices to other enterprises or public authorities should be seen in the widest sense, particularly in the context of a business setting (B2B, B2G for enterprises with more than 10 persons employed) that hand-to-hand delivery of an invoice is not common practice.

I1. Did your enterprise send invoices to other enterprises or public authorities during 2013?

[Scope: enterprises with computers i.e. A1 = Yes]

[Type: one single answer needed per item, i.e. *tick only one per item*; binary (*Yes/No*); filter question]

For sending invoices the proposed scope is B2B/B2G. The two segments B2B/B2G and B2C – both of interest to the users should be separated not only because of the terms that have a different meaning and are translated differently in some languages (invoice vs. bills) but as well because the former segment is the leading segment in terms of using standards and in particular making savings on transaction costs for both parties "sending" and "receiving" enterprises, hence having as well a great potential for growth. The filter question makes it clear for the "sending" enterprises and consistent with the terminology of "invoices" in order to have comparable statistics across countries. Moreover, it becomes clearer for the responding enterprises that the questions refer only to transactions between themselves and other enterprises or public authorities i.e. suppliers and enterprises/public authorities buying from them. These transactions are governed by national laws (national accounting standards related to VAT) implying that responding enterprises keep a record for them.

I2. Of all invoices the enterprise sent to other enterprises or public authorities during 2013, what percentage was sent as:

[Scope: enterprises using computers, i.e. A1 = Yes and enterprises that sent invoices to other enterprises or public authorities during 2013, i.e. I1 = Yes]

[Type: numerical, in percentage values]

a) eInvoices in a standard structure suitable for automated processing? (e.g. EDI, UBL, XML, please add national examples)

b) Invoices in electronic form not suitable for automated processing?

(e.g. emails, e-mail attachment as pdf, images in TIF, JPEG or other format)

c) Invoices only in paper form?

Two distinct types of electronic invoices are proposed i.e. "eInvoices in a standard structure suitable for automated processing" and "Invoices in electronic form not suitable for automated processing". The core difference between the two types of electronic invoices – requiring or not requiring manual intervention - is essential to the commercial case for improving efficiency, increase of productivity and reduction of costs through the dematerialisation of financial processes.

Question I2 includes all three types of invoices and the percentages add up to 100%. This option follows the standard constant sum allocation principle. Respondents are essentially requested to provide the figures for at least the two types of invoices as the third one will be the difference from 100%. The advantage is that respondents become aware of the relative distribution of the percentages and may easily correct the figures in the case of estimates. Moreover, non-response is reduced in case that one of the figures is not readily available; respondents are invited to provide an estimate in order to comply with the sum of 100% of invoices.

I3. Of all invoices the enterprise received during 2013, what percentage was received as:

[Scope: enterprises using computers, i.e. A1 = Yes]

[Type: numerical, in percentage values]

a) eInvoices in a standard structure suitable for automated processing? (e.g. EDI, UBL, XML, please add national examples)

b) Invoices in paper form or in electronic form not suitable for automated processing?

(e.g. emails, e-mail attachment as pdf, images in TIF, JPEG or other format)

In question I3 "invoices only in paper form" or "in electronic form not suitable for automated processing" are proposed together. Therefore, "only paper form" are combined with the unstructured electronic invoices.

Both types of invoices (received) sum up to 100%.

I.1.6.11 MODULE J: E-COMMERCE

In order to assure the broadest international comparability of the enterprise ICT usage statistics, the OECD definition of e-commerce is used throughout this module ("Update of the OECD Statistical Definition of e-commerce (DSTI/ICCP/IIS(2009)5/FINAL):

Table 2. The 2009 definition of e-commerce

OECD definition of e-commerce	Guideline for the Interpretation
An **e-commerce transaction** is the sale or purchase of goods or services, conducted over computer networks by methods specifically designed for the purpose of receiving or placing of orders. The goods or services are ordered by those methods, but the payment and the ultimate delivery of the goods or services do not have to be conducted online. An e-commerce transaction can be between enterprises, households, individuals, governments, and other public or private organisations.	**Include:** orders made in Web pages, extranet or EDI.[7] The type is defined by the method of making the order. **Exclude:** orders made by telephone calls, facsimile, or manually typed e-mail.

Table 3. The framework for measurement

Type	Definition of the type of e-commerce
Web e-commerce	Orders made at an online store (webshop) or via web forms on the Internet or extranet regardless of how the web is accessed (computer, laptop, mobile phone etc.)
EDI e-commerce	Orders initiated with EDI. EDI (electronic data interchange) is an e-business tool for exchanging different kinds of business messages. EDI is here used as a generic term for sending or receiving business information in an agreed format which allows its automatic processing (*e.g.* EDIFACT, XML, etc.) and without the individual message being manually typed. "EDI e-commerce" is limited to EDI messages placing an order.

This module covers not only received orders (sales) and placed orders (purchases) which are transmitted via internet, but also those transmitted via other computer networks. A breakdown between internet and other networks is not included in the survey. A functional split between e-commerce sales and e-commerce purchases is used. Since the 2011 survey web sales and EDI-

type sales are separate sub-modules and comprise mandatory reporting variables. For the e-commerce purchases the distinction between web sales and EDI-type sales is optional.

Bookings and reservations are considered similar to orders; for some economic sectors (e.g. hotels) these terms are more commonly used to describe the "order".

One important difference in this module compared with most other variables in the questionnaire is that the reference period is the previous calendar year, instead of the current situation. Flux economic variables like turnover and purchases, the main variables to measure in e-commerce, need to be measured for a longer period instead of just one point in time. In order to keep comparability with the main business statistics a calendar year is taken as reference period.

E-commerce consists of a business process which is conducted via computer networks. The role of the computer network is very important. It doesn't include the use of this network merely as a person-to-person communication tool. For this reason conventional e-mails manually typed are excluded from this definition. However, if the e-mail system is just used for the automated transmission of a message, i.e. computer-to-computer without human intervention, which represents an order for a product or service, then it is considered an e-commerce transaction.

Therefore transmitted orders by a person using e-mails through a computer network are not considered e-commerce. The typical ways in which e-commerce orders are placed are: 1) through a website and 2) through electronic transmission of EDI-type type messages suitable for automated processing, i.e. using EDI, XML, which in principle characterises business-to-business e-commerce.

Guidelines for specific cases of possible interpretation problems

1) The following example intends to clarify the distinction between the EDI-sales and web sales.

The responding enterprise has a website with sales functionalities. The customer chooses the product and the transaction is completed by pressing the "Send" button. The website creates an EDI-type message which is sent to the sales department to prepare electronically the transportation documents and the goods, to the accounting department to prepare the electronic invoice and to the department responsible for the dispatching the goods to prepare the planning of the dispatch of the goods. The same case can be replicated if the customer is not using the enterprise's website but the website of an online shop which produces the same EDI messages.

According to the e-commerce definition[14], this is referred to as web sales because the order was placed via the enterprise's website (or a website of a third party selling on behalf of the enterprise - online store) even if the enterprise received it as an EDI-type message. The reporting enterprise should avoid double counting the sales.

Sales over an extranet are not EDI-type sales but web sales. In e-Commerce it is explicitly mentioned that *"Web sales are sales made via an online store (web shop) or via web forms on*

[14] "The type of e-commerce transaction is defined by the method of making the order. This approach should mitigate the interpretation problems where both types, EDI and Web, are used in the process. An example is a situation where an order is made by the customer through a web application but the information is transmitted to the seller as an EDI-message. Here the type of selling application is however web, EDI is only a business application to transmit information" (DSTI/ICCP/IIS(2009)5/FINAL)

your enterprise's website or extranet, regardless of how the web is accessed (computer, laptop, mobile phone …)."

2) Use of "apps", e.g. on mobile phones for ordering goods and services:

"Apps" is a short name for "web applications". There are two types: The browser based and the client based web applications. The former are web applications accessible over a webpage (html and Javascripts) and run within the web browser. The latter are installed on the device (e.g. smartphone), run without going through a browser but they use web protocols.

In the example below, we try to clarify the issue of e-commerce in relation to apps. Having such apps on a mobile phone, tablet or other device allows you to place an order, buy credit, etc. Then for example in coffee shops, the app is used to pay for a coffee. The app uses an internet connection to link to the coffee shop. So, if the coffee shop completes the questionnaire, should they regard such sales as e-commerce? The actual order is only made when the person walks into the coffee shop and orders a coffee. Therefore, according to the definition about *e-commerce being defined as where the order is made regardless of the payment method*, buying credit in this situation is not considered e-commerce. We should keep in mind what we want to measure: some kind of transaction that has an effect on the economic activity. The OECD definition of e-commerce is based on the way the order is placed (payment and delivery play no role). Accordingly, the criterion to decide whether the online purchase of credit from a coffee shop should be considered as e-commerce is whether this can be regarded as an order. The basic requirement of an order is that it indicates what is ordered. It can be goods (e.g. a table) or a service (e.g. the maintenance of the heater), but its content must be clearly defined. In the example above, the content of the order would be "whatever item is available at the shop at the time of the next visit".

The use of apps to order goods (e.g. a coffee and cake next time I enter the shop), are considered as e-commerce and the enterprise should consider them as e-commerce web sales. The responding enterprise should avoid double counting; the purchase of credit should not be considered as e-commerce while ordering goods/services should be considered as e-commerce web sales.

Concluding, the responding – e.g. coffee shop - enterprise should count "customers placing orders" as e-commerce. The double counting problem may occur if the enterprise reports on the "purchased credit" and the "value of the orders received via web sales". Only the "value of the orders received via web sales (apps) should be reported. *(the issue of e-commerce when purchasing credit and placing orders is currently under discussion in the TF)*

3) The following seven cases are grouped together; some of them raise certain issues related to the application of the e-commerce definition.

Case (1) is the usual case of an enterprise having web sales. Case (2) refers to the use of "methods specifically designed for the purpose of receiving orders" over a website even though the selling enterprise (responding enterprise) retrieves (downloads) the order on its own initiative.

Cases (3), (4), (5), (6) and (7) are not considered as e-commerce websales because the webform or the PDF order form is used as a Word document, text document or an e-mail. The website does not provide the means "to place the order" but in principle to "construct the order" that is eventually placed differently (e-mail, other electronic message via the website or post). In these cases the orders are not made necessarily on webpages, i.e. they can be made by printing a PDF document which is then "re-introduced" into a computer system and because of this break, there is no e-commerce because there is a human intervention in the process, and a "break" in the automation.

Case 1) An e-commerce website can offer the possibility to fill in a webform (online) hence the order is placed using a website and the responding enterprise receives the order accordingly.

Case 2) An e-commerce website can offer the possibility to fill in a webform (online) hence the order is placed using a website and the responding enterprise enters the website and retrieves online the order in any format.

Case 3) An e-commerce website can offer the possibility to fill in a webform (online), the form is "printed" in PDF and sent by the customer to the responding enterprise as an e-mail attachment

Case 4) An e-commerce website can offer the possibility to fill in a webform (online), the form is "printed" in PDF and sent by the customer to the responding enterprise by other than e-mail nor electronically i.e. post

Case 5) An e-commerce website can offer the possibility to fill in a PDF order form (not necessarily online), the form is "saved" in PDF, completed and sent by the customer to the responding enterprise as an e-mail attachment

Case 6) An e-commerce website can offer the possibility to fill in a PDF order form (not necessarily online), the form is "saved" in PDF, completed and sent by the customer to the responding enterprise via the same website as an attachment to a message

Case 7) An e-commerce website can offer the possibility to fill in a PDF order form (not necessarily online), the form is "saved" in PDF, completed and sent by the customer to the responding enterprise by other than e-mail nor electronically i.e. post

Web sales

Web sales are sales made via an online store (web shop) or via web forms on your enterprise's website or extranet, regardless of how the web is accessed (computer, laptop, mobile phone …).

J1. During 2013, did your enterprise receive orders for goods or services placed via a website? *(Excluding manually typed e-mails)*

[Scope: enterprises using computers, i.e. A1 = Yes]

[Type: one single answer needed, i.e. *Tick only one;* binary *(Yes/No*); filter question]

J2. Please state the value of the turnover resulting from orders received that were placed via a website (in monetary terms, excluding VAT), in 2013.

Or

Please indicate an estimate of the percentage of the total turnover resulting from orders received that were placed via a website, in 2013.

[Scope: enterprises received orders placed via a website, i.e. J1 = Yes]

[Type: numerical]

Please state the value of the turnover resulting from orders received that were placed via a website (in monetary terms, excluding VAT), in 2013.	☐☐☐ %

For turnover the SBS definition is to be used. See background variable X3 in this document.

Collecting electronic orders in percentage values has the advantage of allowing us to obtain an estimate from respondents who don't have in the enterprise any records which can provide this value. Therefore, unless specific records of all these transactions are kept in a centralised form, it may be difficult for the enterprise to provide the value of orders received that were placed via a website. Another way of coping with this problem is to allow enterprises to estimate the web sales turnover in monetary terms by indicating that "an estimation would suffice".

The disadvantage of collecting this variable in percentage values is that for the majority of enterprises this is still a small value. For most of the enterprises it is actually less than 1%. Therefore it is possible to give the answer in decimals (0.xx).

Also, even when making educated estimates the respondents have the tendency to round it to multiples of 10% and 5%. For very small numbers there is evidence that respondents round significantly around 1%. Therefore, when collected in this form, a large proportion of the values collected are 1%.

Another disadvantage of this method is that for big enterprises a value of less than 1% can still be a significant amount. When answers from big enterprises are involved this can bring a significant instability to the results.

Please state the value of the turnover resulting from orders received that were placed via a website (in monetary terms, excluding VAT), in 2013.	*(National Currency)*

Therefore, the preferable way to collect orders received that were placed via a website is in absolute values. Unfortunately, for a large number of enterprises this value is not available. For this reason, currently the recommended method is to give the alternative to the enterprise to answer in precise absolute values if it has the information or, in case such information is not available, ask for an estimation, either in monetary terms or as a percentage of the total turnover.

Please state the value of the turnover resulting from orders received that were placed via a website (in monetary terms, excluding VAT), in 2013 If you can't provide this value, **Please indicate an estimate of the percentage of the total turnover resulting from orders received that were placed via a website, in 2013.**	*(National Currency)* %

This intends first to reflect what the practice in several national surveys is already: some countries ask for the absolute value, while others ask for the percentage value. Second, it intends to provide an alternative to avoid the bias that might exist from asking in each of this ways: underestimation in the case of the absolute value and overestimation in the case of percentage values, by giving the enterprises the option of using the one for which they can provide more accurate answers.

The absolute or percentage values collected can be grossed-up to absolute or the overall percentage. Eurostat preferred practice is to get the grossed-up value of the percentage, i.e. the overall percentage of turnover resulting from orders received that were placed via a website.

When grossing-up, the values collected as percentages need to be weighted by the total turnover (variable X3). Please refer to section I.2.3.3 for more information on the weighting.

J3. Please provide a percentage breakdown of the turnover from orders received that were placed via a website in 2013 by type of customer. *(Estimates in percentage of the monetary values, excluding VAT)*

a) B2C *(Sales to private consumers)*

b) B2B *(Sales to other enterprises)* ***and B2G*** *(Sales to public authorities)*

c) Total: 100%

[Scope: enterprises received orders placed via a website i.e. J1 = Yes]

[Type: Numerical; expression in percentage]

For respondents having received orders that were placed via a website, a percentage breakdown of the turnover by type of customer is requested. The two most important types of e-commerce occur in business-to-consumer (B2C) and business-to-business (B2B) markets.

a) B2C (Sales to private consumers)

The term B2C stands for business-to-consumer and refers to electronic commerce transactions between enterprises and the individuals as the end consumer. Business-to-consumer electronic commerce typically takes the form of websites that offer the possibility to individuals to place order for products.

b) B2B (Sales to other enterprises) and B2G (Sales to public authorities)

The term B2B stands for business-to-business and refers to electronic commerce transactions between enterprises (different from transactions between enterprises and other groups, like consumers (individuals) and the government). B2B refers to commercial transactions between the responding enterprise and other enterprises (e.g. manufacturer and a wholesaler, a wholesaler and a retailer). Business-to-business electronic commerce typically takes the form of processes between trading partners and is performed in higher volumes than business-to-consumer applications (e.g. use of e-marketplaces or via the respondent's websites using login/password procedures). B2B e-market places connect buyers and sellers through a hub where on-line transactions can be executed. The distinction of e-commerce to B2B on the one hand and B2C on the other hand is important because B2B transactions have advantages for enterprises e.g. reduction of product cycle times, lowering stock levels and increasing trade volumes.

B2G stands for business-to-government and includes the electronic commercial transactions between the responding enterprise and public authorities that are conducted via a website.

J4. Which of the following means of payment are accepted for sales via a website?

[Scope: enterprises received orders placed via a website i.e. J1 = Yes]

[Type: one single answer per item needed, i.e. Tick only one; binary (Yes/No); Multiple items may be expected]

E-commerce only refers to receiving online orders, not to online payment or shipping. This question provides information on the integration of the related processes of payments, i.e. is the payment made online or offline.

a) Online payment, i.e. payment integrated in the ordering transaction (e.g. credit, debit card, direct debit authorisation, via 3rd party accounts)

b) Offline payment, i.e. payment process is not included in the ordering transaction (e.g. cash on delivery, bank transfer, cheque payment, other not online payment)

Online payments that are not an **integrated** part of the online ordering transaction are not included in answer option a. This means that if the payment is done via e-banking after the process of online ordering has been completed, this is not regarded as online payment as meant here. For example, if a web shop sends an invoice to a customer after the order has been placed online, and the customer pays this invoice/bill online via e-banking, this should be seen as **offline** payment here (i.e. answer option b).

EDI-type sales

EDI-type sales are sales made via EDI-type messages (EDI: Electronic Data interchange) meaning:

- In an agreed or standard format which allows their automated processing (e.g. EDIFACT, UBL, XML,)
- Without the individual messages being typed manually

J5. During 2013, did your enterprise receive orders for goods or services placed via EDI-type messages?

[Scope: enterprises using computers, i.e. A1 = Yes]

[Type: one single answer needed, i.e. *Tick only one;* binary *(Yes/No*); filter question]

J6. Please state the value of the turnover resulted from orders received that were placed via EDI-type messages (in monetary terms, excluding VAT), in 2013

Or

Please indicate an estimate of the percentage of the total turnover resulted from orders received that were placed via EDI-type messages, in 2013.

[Scope: enterprises that have received orders that were placed via EDI-type messages, i.e. J5 = Yes]

[Type: numerical]

For turnover the SBS definition is to be used. See background variable X3 in this document.

Note: Clarification on sales over the **extranet**: Web-sales or EDI-sales

Enterprises (respondents) may receive orders placed over the **extranet** and in principle concern business-to-business transactions. It may not always be sufficiently clear whether the orders received over an **extranet** should be considered as web-sales or EDI-type sales. In case that it is not clear, the level of automation for placing the order should be considered and clarifications may be required from the respondents. In fact respondents should know how their customers place orders besides just EDI-type messages.

The OECD definition of e-commerce is based on the way the order is placed (payment and delivery play no role). In case the order is placed via forms/web forms (may or may not include a procedure to "log on" on the supplier's site for authentication purposes) it is clearly web-sales regardless of the fact that an exchange of EDI-type messages follows. However, some enterprises only know that they transmit EDI-type messages for placing their orders and respectively their suppliers (respondents) receive and process automatically EDI type messages over the extranet. The issue to be clarified with the enterprise (respondent) should be exactly on the operation of their partner when placing the order.

It is not sufficient to know that manual "log on" authentication procedures are used before the exchange of EDI-type messages. The exchange of EDI-type messages in some cases comprises the use of authentication procedures ("log on" security procedures) before engaging in transactions which may be either automated (computer-to-computer) or manual (individual-to-computer). The "log on" procedure may take place either once in time or every time before engaging in transactions (a "log on" procedure may require for example login/password, a static IP address and a specific port). A manual "log on" procedure for the identification of the purchaser on the extranet of the supplier (respondent) is not sufficient for clearly describing web- or EDI- sales.

Only the description of the actual operation for placing the order would provide the necessary information for web- or EDI-sales and in certain cases the NSI has to take a decision. However, it is suggested that "web-sales" are completely excluded before declaring the sales as "EDI-sales".

E-commerce purchases

E-commerce purchases are purchases made via any of the following ways:

- via an online store (web shop) or via web forms on a website or an extranet of another enterprise, or
- via EDI-type messages (EDI: Electronic Data Interchange) which means messages in an agreed or standard format suitable for automated processing (e.g. EDIFACT, UBL, XML etc.) without the individual messages being typed manually.

For purchases, the SBS definition is to be used:

Purchases of goods and services include the value of all goods and services purchased during the accounting period for resale or consumption in the production process, excluding capital goods the consumption of which is registered as consumption of fixed capital. The goods and services concerned may be either resold with or without further transformation, completely used up in the production process or, finally, be stocked.

Included in these purchases are the materials that enter directly into the goods produced (raw materials, intermediary products, components), plus non-capitalised small tools and equipment. Also included is the value of ancillary materials (lubricants, water, packaging, maintenance and repair materials, office materials) as well as energy products. Included in this variable are the purchases of materials made for the production of capital goods by the unit.

Services paid for during the reference period are also included regardless of whether they are industrial or non-industrial. In this figure are payments for all work carried out by third parties on behalf of the unit including current repairs and maintenance, installation work and technical studies. Amounts paid for the installation of capital goods and the value of capitalised goods are excluded.

Also included are payments made for non-industrial services such as legal and accountancy fees, patents and licence fees (where they are not capitalised), insurance premiums, costs of meetings of shareholders and governing bodies, contributions to business and professional associations, postal, telephone, electronic communication, telegraph and fax charges, transport services for goods and personnel, advertising costs, commissions (where they are not included in wages and salaries), rents, bank charges (excluding interest payments) and all other business services provided by third parties. Included are services which are transformed and capitalised by the unit as capitalised production.

Expenditure classified as financial expenditure or extraordinary expenditure in company accounts is excluded from the total purchases of goods and services.

Purchases of goods and services are valued at the purchase price, i.e. the price the purchaser actually pays for the products, including any taxes less subsidies on the products bought excluding however value added type taxes.

All other taxes and duties on the products are therefore not deducted from the valuation of the purchases of goods and services. The treatment of taxes on production is not relevant in the valuation of these purchases.

For the statistics on activities defined in Section 3 of Annexes I to IV, of Regulation (EC) No 295/2008 except for the enterprises with an activity classified in NACE Rev.2 Section K, expenditure classified as financial expenditure or extra-ordinary expenditure in company accounts is excluded from the total purchases of goods and services.

For the statistics on activities of NACE Rev.2 groups 65.1 and 65.2, the total purchases of goods and services is defined as the gross value of reinsurance services received plus total of commissions as referred to in Article 64 of Council Directive 91/674 plus any other external expense on goods and services (excluding personnel costs).

For the statistics on activities defined in Section 3 of Annex VI of Regulation (EC) No 295/2008, the total purchases of goods and services is defined as commissions payable plus other administrative expenses plus other operating charges.

Source: Commission Regulation (EC) 250/2009 (variable 13 11 0); see also that Regulation of further details of the link to company accounts.

J7. During 2013, did your enterprise place orders for goods or services via a website or EDI-type messages? *(excluding manually typed e-mails)*

[Scope: enterprises using computers, i.e. A1 = Yes]

[Type: one single answer needed, i.e. *Tick only one*; binary (*Yes/No*); filter question]

E-commerce purchases are purchases made via any of the following ways:

-via an online store (web shop) or via web forms on a website or an extranet of another enterprise, or

-via EDI-type messages (EDI: Electronic Data Interchange) which means messages in an agreed or standard format suitable for automated processing (e.g. EDIFACT, UBL, XML etc.) without the individual messages being typed manually.

[Purchases of goods or services include the value of all goods and services purchased during the accounting period for resale or consumption in the production process, excluding capital goods the consumption of which is registered as consumption of fixed capital.]

For more information please refer to the introduction of e-commerce at the beginning of this module.

J8. During 2013, did your enterprise place orders for products or services via a website?

[Scope: enterprises sent orders for products or services via a website or EDI-type message, i.e. J7 = Yes]

[Type: one single answer needed, i.e. *Tick only one*; binary (*Yes/No*); optional question]

J9. During 2013, did your enterprise place orders for products or services via EDI-type messages?

[Scope: enterprises sent orders for products or services via a website or EDI-type message, i.e. J7 = Yes]

[Type: one single answer needed, i.e. *Tick only one*; binary (*Yes/No*); optional question]

J10. During 2013, was the value of the orders that your enterprise placed electronically equal or more than 1% of the total purchases' value? (in monetary terms, excluding VAT)

[Scope: enterprises sent orders for products or services via a website or EDI-type message, i.e. J7 = Yes]

[Type: one single answer needed, i.e. *Tick only one;* binary *(Yes/No*)]

I.1.6.12 MODULE X: BACKGROUND INFORMATION

The background variables have several purposes. Firstly, they are used to breakdown the results from the main characteristics. That is the case for the "Main economic activity of the enterprise" and "Average number of persons employed".

Secondly, they are needed to weight the percentages of turnover and purchases from e-commerce. The background variables "Total purchases of goods and services" and "Total turnover" are used for that. The variable "Average number of persons employed" is similarly used to weight the percentage of persons employed using computers, the percentage of persons employed using

computers with access to the www, etc. The number of persons employed is also used to weight the qualitative variables.[15]

Thirdly, the background variables are used in the sampling design. Namely, the "Main economic activity" and the "Average number of persons employed" are used to stratify the sample.

The background variables described so far may be collected through the ICT survey questionnaire or obtained from alternative sources. The alternative sources are mainly the registers and one main business survey, usually used to produce the structural business statistics. It is very important that the background information is at least consistent with the structural business statistics.

X1. Main economic activity of the enterprise

[Scope: all enterprises]

[Type: categorical]

The main (or principal) economic activity is identified as the activity which contributes most to the total value added of the enterprise. The principal activity so identified does not necessarily account for 50% or more of the enterprise's total value added. The classification of principal activity is determined by reference to NACE, first at the highest level of classification and then at more detailed levels ("top-down" method).

The nomenclature NACE Rev. 2 is available in Eurostat's RAMON database:

http://ec.europa.eu/eurostat/ramon/nomenclatures/NACERev2

The main economic activity of the enterprise should be classified by NACE Rev. 2 at its highest level of detail (4 digits). Nevertheless, only the following level of detail which is used in the breakdown is strictly necessary.

NACE Rev. 2 groupings		
Aggregates for calculation of **national** NACE Rev. 2 aggregates :		
1	10 – 18	Manufacture of products based on: food, beverages, tobacco, textile, leather, wood, pulp and paper; publishing and printing
2	19 – 23	Manufacture of coke, refined petroleum products, chemical products, basic pharmaceutical products, rubber and plastics, other non-metallic mineral products.
3	24 – 25	Manufacture of basic metals and fabricated metal products excluding machines and equipments
4	26 – 33	Manufacture of computers, electric and optical products, electrical equipment, machinery and equipment n.e.c, motor vehicles, other transport equipment, furniture, other manufacturing, repair and installation of machinery and equipment
5	35 – 39	Production and distribution of electricity, gas, steam and air conditioning; water supply, sewerage, waste management and remediation activities

[15] For more information on the weighting refer to chapter I.2.3.3.

6	41 – 43	Construction
7	45 – 47	Wholesale and retail trade; repair of motor vehicles and motorcycles
8	49 – 53	Transport and storage
9	55	Accommodation
10	58 – 63	Information and communication
11	68	Real estate activities
12	69 – 74	Professional, scientific and technical activities, except veterinary activities
13	77 – 82	Administrative and support service activities
14	26.1 – 26.4 + 26.8 + 46.5 + 58.2 + 61 + 62.01 + 62.02 + 62.03 + 62.09 + 63.1 + 95.1	Manufacture of electronic components and boards, consumer electronics, magnetic and optical media; wholesale of information and communication equipment; software publishing; telecommunications; computer programming, consultancy and facilities management activities, other information technology and computer service activities; data processing, hosting and related activities, web portals; repair of computers and communication equipment or ICT sector
Aggregates for calculation of **European** NACE Rev. 2 aggregates:		
1a	10 – 12	Manufacture of beverages, food and tobacco products
1b	13 – 15	Manufacture of textiles, wearing apparel, leather and related products
1c	16 – 18	Manufacture of wood and products of wood and cork, except furniture; articles of straw and plaiting materials; paper and paper products; printing and reproduction of recorded media
4a	26	Manufacture of computer, electronic and optical products
4b	27 – 28	Manufacture of electrical equipment, machinery and equipment n.e.c.
4c	29 – 30	Manufacture of motor vehicles, trailers and semi-trailers, other transport equipment
4d	31 – 33	Manufacture of furniture and other manufacturing; repair and installation of machinery and equipment
7a	45	Wholesale and retail trade and repair of motor vehicles and motorcycles
7b	46	Wholesale trade, except of motor vehicles and motorcycles
7c	47	Retail trade, except of motor vehicles and motorcycles
9a	55 – 56	Accommodation; food and beverage service activities
10a	58 – 60	Publishing activities; motion picture, video and television programme production, sound recording and music publishing activities; programming and broadcasting activities
10b	61	Telecommunications
10c	62 – 63	Computer programming, consultancy and related activities, information service activities
13a	77 – 78 + 80 – 82	Activities for : rental and leasing, employment, security and investigation , services to buildings and landscape, office administrative, office support and other business support
13b	79	Travel agency, tour operator and other reservation service and related activities
14a	95.1	Repair of computers and communication equipment

The NACE Rev. 2 categories are grouped together for dissemination purposes into several aggregates organised in 5 hierarchal levels. At the first level there are 2 categories distinguishing "Manufacturing, Energy and Construction" and "Non-financial services". At the second level, activities are grouped at the Section level of NACE, making 11 categories. The content of these groupings is described below. See also the Transmission format in annex I.3.2.

NACE Rev. 2 groupings

Level 0	Level 1	Level 2	Level 3	Level 4	Level 5
Total	Manufacturing, energy and construction	Section C	10 – 33	10 – 18	10 – 12
					13 – 15
					16 – 18
				19 – 23	
				24 – 25	
				26 - 33	26
					27 – 28
					29 – 30
					31 – 33
		Section D,E		35 – 39	
		Section F		41 – 43	
	Non-financial services	Section G		45 – 47	45
					46
					47
		Section H		49 – 53	
		Section I		55	55 – 56
		Section J		58 – 63	58 – 60
					61
					62 – 63
		Section L		68	
		Section M		69 – 74	
		Section N		77 – 82	77 – 78 + 80 – 82
					79
	ICT sector	26.1 – 26.4 + 26.8 + 46.5 + 58.2 + 61 + 62.01 + 62.02 + 62.03 + 62.09 + 63.1 + 95.1			95.1*

* Since 2010, 95.1 is included in the overall total. Additionally, the NACE Rev 2 economic activity I56 (Section I, Division 56: food and beverage service activities) breakdown is not requested to be reported separately (Data transmission format document). However, NACE I56 is to be included (together with I55) in the totals (i.e 10_C10_S951_XK).

For the breakdowns coloured in green, Member States have to collect data with a sufficient high quality to allow the publication of aggregates at national level.For the breakdowns coloured in blue, Member States are required to collect data with a sufficient quality to allow the publication of European aggregates only. On a voluntary basis – where applicable included in a grant agreement - they may also collect data of better quality, so that the publication of national aggregates is feasible.

***X2. Average number of persons employed*, during 2013**

[Scope: all enterprises]

[Type: numerical]

For purposes of general harmonisation of enterprise ICT usage statistics and the more general field of business statistics, the concept of persons employed used here is the one from the Structural Business Statistics (SBS):

The number of persons employed is defined as the total number of persons who work in the observation unit (inclusive of working proprietors, partners working regularly in the unit and unpaid family workers), as well as persons who work outside the unit who belong to it and are paid by it (e.g. sales representatives, delivery personnel, repair and maintenance teams). It includes persons absent for a short period (e.g. sick leave, paid leave or special leave), and also those on strike, but not those absent for an indefinite period. It also includes part-time workers who are regarded as such under the laws of the country concerned and who are on the payroll, as well as seasonal workers, apprentices and home workers on the payroll.

The number of persons employed excludes manpower supplied to the unit by other enterprises, persons carrying out repair and maintenance work in the enquiry unit on behalf of other enterprises, as well as those on compulsory military service.

Unpaid family workers refer to persons who live with the proprietor of the unit and work regularly for the unit, but do not have a contract of service and do not receive a fixed sum for the work they perform. This is limited to those persons who are not included on the payroll of another unit as their principal occupation.

Note: In order to check the comparability of data, it is necessary to indicate whether voluntary workers have been included under this heading or not.

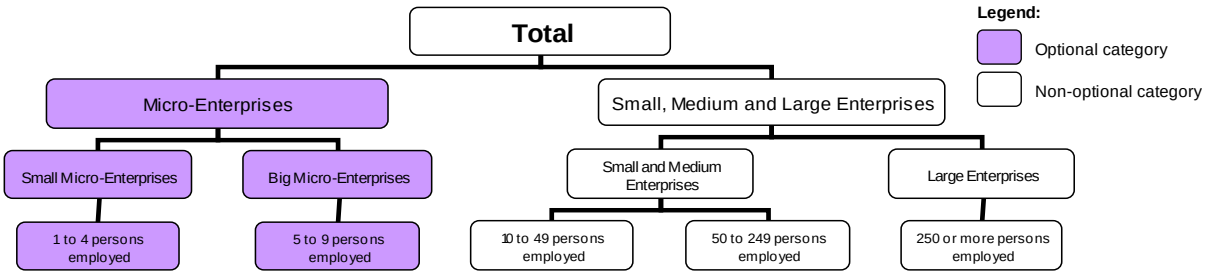
The number of persons employed is a headcount and is measured as an annual average using at least data for each quarter of the year except for the statistics on the activities defined in Section 3 of Annex V, VI and VII of Regulation (EC) No 295/2008 for which the calculation can be done on the basis of data with a lower frequency.

Source: Commission Regulation (EC) 250/2009 (variable code 16 11 0)

The number of persons employed should be measured as the yearly average during the previous calendar year. For comparability reasons, the number of persons employed should not be confused with the number of employees (that excludes unpaid workers) or the number of employees in full time equivalent units (fte's).

The average number of persons employed is coded into 5 size categories. For the tabulation of the results, the size categories are grouped in a hierarchical way into 3 levels. The first level distinguishes micro-enterprises from the others. The second and third levels further breakdown these two groups.

Size Categories (persons employed)	
Compulsory	
10 to 49	Small enterprises
50 to 249	Medium enterprises
250 or more	Large enterprises
Optional	
Less than 5	Small micro-enterprises
5 to 9	Big micro-enterprises



X3. Total turnover (in value terms, excluding VAT), for 2013

[Scope: all enterprises]

[Type: numerical]

This background variable is needed to weight the percentage of turnover resulting from orders received via computer networks.

As for the number of persons employed and purchases of goods and services, the SBS definition is to be used.

For the statistics on activities defined in Section 3 of Annexes I to IV and VIII of Regulation (EC) No 295/2008 except for the activities classified in NACE Rev.2 Section K turnover comprises the totals invoiced by the observation unit during the reference period, and this corresponds to market sales of goods or services supplied to third parties. The sales of goods include the goods produced by the enterprise as well as the merchandise purchased by a retailer or land and other property held for resale (if land and other property were initially purchased for investment purposes they should not be included in turnover). The rendering of services typically involves the performances by the enterprise of a contractually agreed task over an agreed period of time. The revenue of the long-term contracts (e.g. building contracts) should be recognised by reference to the stage of completion of the contract and not the finished contract method. Goods produced for own consumption or investment should be excluded from turnover.

Turnover includes all duties and taxes on the goods or services invoiced by the unit with the exception of the value added type taxes (VAT). VAT are collected in stages by the enterprise and fully borne by the final purchaser.

It also includes all other charges (transport, packaging, etc.) passed on to the customer, even if these charges are listed separately in the invoice. Reduction in prices, rebates and discounts as well as the value of returned packing must be deducted.

Income classified as other operating income, financial income and extra-ordinary income in company accounts according to the 4th Accounting Directive and revenue from the use by others of enterprise assets yielding interest, royalties and dividends and other income according to IAS/IFRS is excluded from turnover. Operating subsidies received from public authorities or the institutions of the European Union are also excluded.

For the statistics on activities defined in Section 3 of Annex V of Regulation (EC) No 295/2008, the corresponding title of this characteristic is "Gross premiums written". This characteristic is defined in Article 35 of Directive 91/674/EEC[16]. Note: For the layout of the profit and loss account (technical account): Article 34, I. 1(a) of Directive 91/674/EEC for non-life insurance business and Article 34, II. 1(a) of Directive 91/674/EEC for life insurance business.

For the statistics on activities defined in Section 3 of Annex VII of Regulation (EC) No 295/2008, the corresponding title of this characteristic is "Total pension contributions". This characteristic shall comprise all pension contributions, due during the financial year, in respect of pension contracts, such as all mandatory contributions, other regular contributions, voluntary additional contributions, incoming transfers, other contributions.

Source: Commission Regulation (EC) 250/2009 (variable code 12 11 0); see also further details of the link to company accounts.

I.1.7. Benchmarking indicators linked to 2014 Model Questionnaire

Benchmarking indicator B10 (% of persons employed using computers connected to the internet in their normal work routine) is based on the following calculation:

a) The % of persons employed using computers connected to the internet is calculated by dividing the number of persons employed who use computers with access to the internet for business purposes by the total number of persons employed, multiplied by 100. (MQ2014, C2)

Benchmarking indicator B11 (% of enterprises with broadband access (fixed or mobile)) is based on the following calculation:

a) The % of enterprises with broadband access is calculated by dividing the number of enterprises with broadband access (fixed or mobile) by the total number of enterprises, multiplied by 100. (MQ2014, C3 or C5).

Benchmarking indicator B12 (% of enterprises giving devices for a mobile connection to the internet to their employees) is based on the following calculation:

a) The % of enterprises giving devices for a mobile connection to the internet to their employees is calculated by dividing the number of enterprises providing to persons employed portable devices that allowed an internet connection via mobile telephone networks for business purposes by the total number of enterprises, multiplied by 100. (MQ2014, derived from C6)

Benchmarking indicator B13 (% of persons employed provided by the enterprises with devices for a mobile connection to the internet) is based on the following calculation:

a) The % of persons employed who were provided by the enterprises with devices for a mobile connection to the internet is calculated by dividing the persons employed who were provided with a portable device that allowed an internet connection via mobile telephone networks for business purposes by the total number of persons employed, multiplied by 100. (MQ2014, C6)

Benchmarking indicator D1 (Integration of internal business processes: % of enterprises whose internal business processes are automatically linked) is calculated the following way:

[16] OJ L 374, 31.12.1991, p. 7

a) The % of enterprises which use an ERP software package to manage resources by sharing information among different functional areas of the enterprise is calculated by dividing the number of enterprises which use such an ERP software package to manage resources by sharing information among different functional areas by the total number of enterprises, multiplied by 100. (MQ2014, E1)

Benchmarking indicator D2 (% of enterprises using dedicated applications for employees to access human resources services) is not calculated, relevant question not available in MQ2014.

Benchmarking indicator D3 (% of enterprises electronically exchanging business documents with suppliers and/or customers broken down by type of document) is based on the following calculations:

The % of enterprises which electronically exchange business documents with suppliers and/or customers is calculated by dividing the number of enterprises which send or receive EDI-type messages suitable for automated processing for a) payment instructions to financial institutions and/or b) data to/from public authorities by the total number of enterprises, multiplied by 100. (MQ2014, F1(a) or (b)

Benchmarking indicator D4 (% of enterprises sharing electronically information on supply chain management broken down by business function) is calculated by dividing the number of enterprises sharing supply chain management information electronically with its suppliers or customers by the total number of enterprises, multiplied by 100. (MQ2014, G1)

Benchmarking indicator D5 (% of enterprises sending and / or receiving e-invoices) is not used for benchmarking anymore.

Benchmarking indicator D6 (% of enterprises using key technologies for the internet of things, by purpose) is calculated by dividing the number of enterprises that are using Radio Frequency Identification instruments for either of a) Person identification or access control or b) as part of the production and service delivery process or c) for product identification after the production process by the total number of enterprises multiplied by 100. (MQ 2014, H1(a) or (b) or (c))

Benchmarking indicator D7 (% of enterprises having a website with e-commerce functions) is calculated by dividing the number of enterprises having a website with online ordering or reservation or booking by the total number of enterprises, multiplied by 100. (MQ 2014, C8b)

Benchmarking indicator D8 (% of enterprises using software applications for managing information about clients, like CRM) is based on the following calculations:

a) The % of enterprises using software applications for managing information about clients, like CRM is calculated by dividing the number of enterprises that use CRM software to manage the analysis of information about customers for marketing purposes by the total number of enterprises, multiplied by 100. (MQ2014, E2b).

Benchmarking indicator D9 (enterprises' turnover from e-commerce as % of total turnover) is calculated by dividing the turnover from orders received via either web sales or EDI-type sales by the enterprises' total turnover, multiplied by 100. (MQ2014, J2, J6)

Benchmarking indicator D10 (% of enterprises selling by e-Commerce) is based on the following calculation:

a) The % of enterprises selling by e-Commerce is calculated by dividing the number of enterprises which in 2013 received orders either via a website or EDI-type messages by the total number of enterprises, multiplied by 100. (MQ2014, J1 or J5)

b) The % of enterprises selling by e-Commerce is calculated by dividing the number of enterprises which in 2013 received orders (of at least 1% of turnover) either via a website or EDI-type messages by the total number of enterprises, multiplied by 100. (MQ2014, J2 or J6)

Benchmarking indicator D11 (% of enterprises purchasing by e-Commerce) is based on the following calculation:

a) The % of enterprises purchasing by e-Commerce is calculated by dividing the number of enterprises which during 2013, sent orders for products or services via a website or EDI-type messages by the total number of enterprises, multiplied by 100. (MQ2014, J7)

b) The % of enterprises purchasing by e-Commerce is calculated by dividing the number of enterprises which during 2013, placed electronically orders (of equal or more than 1% of the total purchases' value) by the total number of enterprises, multiplied by 100. (MQ2014, J10)

Benchmarking indicator D12 (% of enterprises doing e-commerce transactions broken down by destination (National, EU, World) is not calculated as the relevant questions are not available in the MQ 2014.

Benchmarking indicator E3 (% of enterprises using the internet for interacting with public authorities broken down by level of sophistication is not calculated as the relevant questions are not available in the MQ 2014.

I.2. PRODUCTION METHODOLOGY

I.2.1. Timetable – Survey period and reference period

It is recommended that the survey period of the annual survey (year **T**) on *ICT Usage and e-Commerce in Enterprises* is the first quarter of calendar year **T**. The variables in the questionnaire are predominantly qualitative (dichotomous questions, Yes/No response or tick-box-questions). The questions refer to the "current" situation (at the time of the survey). The answers to the quantitative questions - for example "the percentage of the total turnover resulting from orders received that were placed via EDI-type messages" - refer to the calendar year **T-1** (previous year). Due to national planning, for some countries the reference period still deviates from the one of the Model Questionnaire; this might affect the comparability.

Some countries have difficulties in collecting monetary data (turnover) right at the beginning of the year. In these cases, the collection should start as soon as possible. Preferably still during the first quarter.

To maximize the level of harmonization between the national surveys in the different Member-States, it is important that this survey period is respected as much as possible.

Table: Overview of the survey period[17] and reference period

Source: Metadata/Quality reports for the survey on ICT Usage and e-Commerce in Enterprises (2013)

Country	Data collection period	Reference period
Belgium	17/03-29/9/2013	as in model questionnaire
Bulgaria	7/06-30/06/2013	as in model questionnaire
Czech Republic	28/02 – 2/09/2013	as in model questionnaire
Denmark	22/02 – 25/06/2013	as in model questionnaire
Germany	March - 12/07/2013	as in model questionnaire
Estonia	December 2012 – May 2013	as in model questionnaire
Ireland	22/01 - September 2013	*
Greece	April 2013 – October 2013	as in model questionnaire
Spain	January - end of March 2013	as in model questionnaire
France	9/01 - 13/05/2013	2012, no reference to January 2013 because questionnaires have been sent in January.
Italy	26/04 – 2/09/2013	as in model questionnaire
Cyprus	30+31/01 - 05/2013	as in model questionnaire
Latvia	12/01 - 10/05/2013	as in model questionnaire.
Lithuania	January - May 2013	as in model questionnaire
Luxembourg	15-18/03 - June 2013	as in model questionnaire
Hungary	4/03 - 31/03/2013	as in model questionnaire

[17] The periods mentioned refer to the general survey; for some countries differences can occur for the survey in the micro-enterprises in case a separate survey is organised.

Malta	February - May 2013	as in model questionnaire
Netherlands	5/06/2013 – 19/11/2013	current situation instead of January 2013
Austria	25/2/2013 – 24/07/2013	as in model questionnaire
Poland	16/04-24/05/2013	as in model questionnaire
Portugal	15/02 – 27/08/2013	as in model questionnaire
Romania	April/May – July 2013	as in model questionnaire
Slovenia	20/02 – 02/08/2013	as in model questionnaire
Slovak Republic	20/03-30/06/2013	as in model questionnaire
Finland	04/02 – 25/06/2013	as in model questionnaire
Sweden	26/03 – 31/08/2013	as in model questionnaire
United Kingdom	4/01 – 28/10/2013	December 2012 instead of Jan 2013
Island	8/04/ - 30/05/2013	as in model questionnaire
Norway	26/02-29/05/2013 (to be continued through summer)	as in model questionnaire
Montenegro	15/04-1/05/2012	as in model questionnaire
Croatia	13/05/2013 – 19/07/2013	as in model questionnaire
Macedonia	February - June 2013	as in model questionnaire
Turkey	11/06-25/07/2012	as in model questionnaire

* Information not available yet

I.2.2. **Frame population**

Normally, the enterprise ICT survey will be carried out in the form of a sample survey. The *frame population* (or *sampling population*) is the list of enterprises from which the sample will be drawn. Ideally, this list of enterprises should be equivalent to the target population as both over-coverage and under-coverage can induce bias and affect the reliability of the survey results.

E.g. over-coverage due to inclusion of enterprises for which the main economic activity is agriculture (outside the scope of the survey) in the sampling frame: if this sector shows a below the average adoption of ICT, the population total may show a downwards bias due to inclusion of this group.

E.g. under-coverage due to the fact that some remote areas are not covered in the register: as such areas usually have a lower penetration of ICT, the overall indicators may be relatively overestimated for such a country.

The sample for the survey should be drawn from the business register in the different Member-States as defined in the Regulation (EC) No 177/2008 of the European Parliament and the Council (EC) setting out a common framework for the harmonisation of the national business registers for statistical purposes. Part of this register is the activity code at the four-digit (class) level of NACE Rev. 2 and the size measured by the number of persons employed of the enterprises.

The choice to comply with the "enterprise" as the appropriate statistical unit and the business register and its enterprise' characteristics as the framework population, is inspired by the wish to keep a certain level of consistency with the Structural Business Statistics. On a macro-level as well as on a micro-level results of the Structural Business Statistics will be combined with data on the ICT usage of enterprises for analysis and for weighting-purposes (for the latter: see paragraph I.2.3.3).

To allow for an assessment, it is important to include the coverage and the sampling frame when reporting metadata (see Chapters 3 and 5 in the reporting template, Annex I.3.3).

The table below presents the sampling frames used in the national statistical institutes for the latest survey.

Table: Overview of the sampling frames

Source: Metadata/Quality reports for the survey on ICT usage and e-commerce in enterprises (2013)

Belgium	The frame population is a sub-population of SBS. Frame population: - Active enterprises with more than 5 persons employed - NACE Rev. 2 sections C to J, L and N, divisions 69-74 and group 95.1. - Enterprises with market activities - Enterprises located in any part of Belgium The number of persons employed was calculated from the number of employees (social security data, average of the first three quarters of 2011) with adjustment for the number of unpaid persons employed on the basis of the legal form.
Bulgaria	Sampling: The frame population consists of all economic active enterprises with 10 or more employees existing in the Business Register and responding to the defined by Eurostat main activity groups by NACE rev.2 as of the end of 2011. Weighting and grossing up: the frame population will be updated with actual data from Business Register as of the end of 2012. Same frame is used for all business surveys (SBS). There was not any deviation from the recommendation stated in model survey.
Czech Republic	The frame population is the same as the one for the Structural Business Survey (SBS). Different latest snapshots on the frame population are used during the statistical process.
Denmark	The frame population corresponds with SBS. The frames used for sampling is also used for grossing up. The frame population were drawn 19 December 2012. Further correction for inactive enterprises immediately before sending out. The frame population in 5.3. has been corrected in the course of the survey. Thus the population send for the datatransmission "pop" is a little smaller than the frame population.
Germany	All active enterprises recorded in the German business register (snapshot 2012) form the frame population. For sampling and grossing up the same frame is used.
Estonia	For the population described in the ICT-survey the same frame is used as for the SBS 2012 survey. Differently from SBS the enterprises which may start the activity in 2013 are included and enterprises which deceased in 2012 are excluded. In addition for national needs the survey will cover active enterprises with 10 and more persons employed according to NACE Rev2 (Sections A, B, P, Q, R; Division 75, 95, 96; Groups from 64.2, 64.3; Classes 64.91, 64.99, 66.11)
Ireland*	The frame population is the same as the one used for the Structural Business Statistics. Yes we use the same frame population throughout the ICT survey (sampling, weighting, grossing up) and this frame population is the one used for the SBS. The source of the frame population is the National Business Registry and the latest snapshot taken was December 2012 for the reference year 2012.
Greece	The sampling frame used for the sample design was based on the Business Register of the EL.STAT. The Business Register is based on the V.A.T. Register of the Ministry of Finance and it is updated through the statistical surveys of the EL.STAT and the Register of the Social Insurance Foundation. This population frame was used for the selection of the sampling units and for computing the inclusion probabilities of the surveyed units. For the grossing up compilation a new updated register, available at that time, was used. This new register was updated using data of Business Surveys conducted by ELSTAT.
Spain	The frame population of the ICT survey is the same as the one used for the SBS: DIRCE. The Central Businesses Directory (DIRCE) collects all Spanish businesses in a single directory. Its basic objective is to enable business-targeted sample surveys to be conducted, and consequently, it registers information such as identity data, location, main activity or number of employees. This information is obtained from administrative sources (Inland Revenue and Social Security) and complemented with data from common statistical operations. Moreover, this directory is annually updated. The same frame population is used in the whole process.
France	Whole French enterprise population taken from French business register, except "Financial and insurance activities". Same frames are used during different stages of the statistical process.
Italy	Frame population deriving from enterprises with 10 and more persons employed that are included as ''operating'' in business register. BR considered for sampling and for Metadata report is the final version one and it is referred to year 2011.

	For grossing up procedure and Quality report we will use the same version of BR.
Cyprus	The frame population is the Business Register. It is updated annually (April-May) using information from administrative resources (VAT Department, Social Insurance Department and Department of the Registrar of Companies) so as to take into account changes in the structure of employment by industry, occupation and size of establishment. After relevant sample surveys are completed, the updated information of sampled enterprises is provided to Business Register. The Register was used during the different stages of the statistical process (sampling, grossing-up). The frame population is used for the Structural Business Statistics too.
Latvia	Frame population consists of units which are economic active at the end of 2011, belongs to defined NACE groups (see 5.3). Frame population consists of: • Individual merchants; • Commercial companies; • Peasant and fishermen's farms. The sampling frame is made from statistical business register. The economical active enterprises with at least 10 employees are included in the sampling frame. The frame population is the same as used for the Structure Business Statistics.
Lithuania	The sampling base for the ICT-Enterprise survey is the Statistical Profile Business Register. Statistical units are economically active enterprises. Enterprises, which do not match the size criteria, are removed from the frame used for grossing up.
Luxembourg	The frame population used for sampling is drawn from the business register in March. Since quantitative data for the last two months of the reference year 2012 are not yet available at this date, we take into account the average number of persons employed from January to October 2012 for the definition of size-classes when the frame population is drawn. All mandatory activities laid out in the model survey are covered. In order to be in line with the SBS frame population (size classes of 10 employees or more), we also take into account concepts such as market activity and economic territory when defining the frame population. For the production of the final results, an updated frame population from the business register is used (covering all of 2012), as well as information from SBS preliminary results for 2012.
Hungary	The frame of the survey embraces the following economic categories according to NACE Rev.2. The target population of the survey consists of the operating enterprises with at least 10 employees The underlying register concerning the target population is the Business Register. All these enterprises build up the frame of the survey, i.e. the sampling frame. The observation units of the survey consist of the enterprises that belong to the above mentioned NACE categories and to the Business Register as well (and indicated as working enterprises).
Malta	The frame population is extracted from the Business Register held and maintained by NSO.
Netherlands	The population described in the ICT-survey is drawn from the same business-register as the survey for the Structural Business Statistics and at the same reference date i.c. 1-12-2012. In principle there is no change in the population frame during the course of the statistical process. In practice however, we may find out that some enterprises that were originally defined in our population are in fact no longer active for some reason. Due to administrative delays they may no longer have existed in most of the reference period without us knowing this in time. In these cases we may adjust the population somewhat a posteriori. However, these adjustments are insignificantly small in number.
Austria	The frame population is the number of enterprises in the National Business Register. SBS use the same source for sampling. The same frame population is used during the whole statistical process (NBR as of January 2013).
Poland	Frame was constructed following Eurostat recommendation. The source of the frame was BJS –Polish register of statistical units and it corresponded to the snapshot taken on the end of December 2012. For financial sector full-scale survey was performed. Only codes for active enterprises were used : 11,12,13,14 (acting enterprises, enterprises under organization, enterprises in liquidation, enterprises declaring insolvent) which had turnover, purchases and persons employed during a period which was of interest.
Portugal	The survey frame population is a subset of the Business Register (BR) and fully respects the recommendations of the model questionnaire regarding the economic activity, the enterprise size and the geographic scope. The BR is updated every year and the survey frame is updated during the data collection. Date of enterprise incorporation < 2012.
Romania	Basis for the sampling frame is the Business Register. The sampling frame consists of: the active enterprises at 31 December 2011 that have as main activity included in NACE aggregation mentioned before; the new active enterprises created in 2012. Information about activity code, size class and region is included in the sampling frame and is used to draw the sample.
Slovenia	The sampling frame was made by using the data from the Business Register of the Republic of Slovenia and the information from some other surveys. Regarding activity, we included NACE Rev.2 sections 10-82 and 95.1 for the enterprise survey and NACE Rev.2 sections from 64-66 for the enterprises of the financial sector.

	Regarding the number of employees, we included units with at least 5 employees.
Slovak Republic	Frame population was based on the same Register of statistical units as in case of structural business statistics to 1. 1. 2013.
Finland	The frame population is drawn from the production database of Structural Business Statistics. Frame for drawing the sample is the same as the one used for grossing up the data (minus some out of scope enterprises).
Sweden	The sampling frame is the Swedish Business Register. The version used was frozen in March 2013.
United Kingdom	Whole UK economy taken from the ONS business register, based on SBS. The same frame is used in all stages of the statistical process, from sampling to the production of the final results.
Island	Frame population is based on pay-as-you-earn tax registers. Statistics Iceland does not have a business register to base on. As mentioned earlier legal units are used as proxies for enterprises, which has been the tradition for business statistics in Iceland.
Norway	The frame population is drawn from the Structural Business Statistics from in the beginning of February 2013. The sample was drawn from all enterprises with at least ten persons employed in the Central Register of Establishments and Enterprises.
Croatia	Sampling frame was designed by using data from the Business Register of the Central Financial Agency (FINA) for 2012. All enterprises with 10 or more employees and any income were included ins the sampling frame.
Macedonia	Frame population include all active enterprise with 10 and more employees in given economic activities according to ICT methodology. Sampling frame is from Statistical Business Register 31.12.2012.

I.2.3. **Sampling design**

The survey should be based on a probability sample from which results representative of the population could be derived, considering the agreed breakdowns defined in the questionnaire.

The sampling design and the resulting sample size (see below) should be appropriate for obtaining accurate, reliable and representative results on the survey characteristics and breakdowns specified in the Regulation and the Model Questionnaire.

This objective should be achieved for the overall proportions as well as for the proportions relating to the different subgroups of the population. The estimated standard error for any indicator/variable (in principle proportions) should not exceed 2 percentage points for the overall population and should not exceed 5 percentage points for the different subgroups of the population.

The aim of such requirement is to ensure the collection of a complete dataset – without empty, confidential or unreliable cells - for these indicators. In case this aim cannot be reached, clarifications should be provided.

Further information: (Eurostat 2006). Handbook on methodological aspects related to sampling designs and weights estimation.

The following table gives an overview of the sampling design applied in the participating countries. Since the 2010 survey, NACE Rev. 2 is exclusively used.

Table: Overview of the sampling design and sampling methods

Source: Metadata/Quality reports for the survey on ICT Usage and e-Commerce in Enterprises (2013)

Belgium	Stratified random sample Population is stratified by: - nace rev.2 : 25 breakdowns - size class: 5-9 ; 10-49 ; 50-249 ; 250+ persons employed - NUTS 1 –level (three Belgian regions) Resulting in 25 * 4 * 3 = 300 strata – 12 strata (12 non-existing combination of the 3 variables : region * nacecat

	* sizeclass) = 288 strata A total of +/- 7700 enterprises was selected (with +/- 2566 enterprises from each NUTS 1 level, and +/- 2550 micro enterprises). No coordination with other surveys.
Bulgaria	The sampling method used is one-stage stratified random sampling. The stratification was done by economic activity (25 categories, incl. NACE categories related to European aggregates) and by number of persons employed. In order to improve the sample quality the defined size classes were broken down into smaller groups as follows: 10 – 19, 20 – 49, 50 – 99, 100 – 149, 150 – 199, 200 – 249, 250+ persons employed. As a result the total number of strata was 175. The number of enterprises per stratum was determined using Neyman allocation method. The enterprises in the 6-th and 7th size groups (with 200+ persons employed) were surveyed exhaustively. Minimum sample size per strata was 5 enterprises. Module Complex Samples in SPSS was used for sample selection. The sample was independent, i.e. there were no procedures used for coordination or non- overlapping with samples of other surveys.
Czech Republic	Stratified random sampling for enterprises with 10-249 and census for 250 and more employees with respect to Business Register. An intent sampling was used for enterprises with large share of e-commerce or with huge turnover according to its size Number of enterprises and number of employees was used for stratification. The sample was designed with no reference to any other survey. Sample size was designed to enable accurate, reliable and representative results.
Denmark	The sample is a stratified random sample made by use of register-based data. Strata was made by a combination of size classes and activity. • Sizeclasses: full-time equivalent employment (10-19, 20-49, 50-99, 100-249, 249+). • Activity groups: Nace, 2-digit level.
Germany	For sampling a two-stage stratified random sample was used. The variables of stratification were the federal state, the economic activity (according to NACE Rev. 2) and the number of employees (including a „0"-stratum, see annex for further information). The final number of strata amounted to 2,240. At first stage the total sample size was allocated to the variables federal state and economic activity using the "Methode der vergleichbaren Präzision" (method of equivalent precision). At second stage the optimal allocation by Neyman-Tschuprow was applied for the variable number of employees in each stratum (federal state x economic activity). For the „0"-stratum in each stratum of the sample (federal state x economic activity) a certain number of enterprises from the "0"-stratum was selected. For most cases this corresponded to an inclusion probability of less than 1%. (Size class 1 to 4 persons employed in table 5.3 and table 6.2 includes enterprises of the „0"-stratum; for the exact number of enterprises in the „0"-stratum, see annex.) No coordination with samples of other surveys took place.
Estonia	The stratified simple random sampling method is used. The sample is stratified by economic activity and by number of persons employed. The enterprises were divided by number of persons employed into following size groups: 10-19, 20-49; 50-99; 100-249; 250+. The Neyman optimal allocation was used for sample allocation and determination of sample size. By number of persons employed, the survey includes all enterprises with 50 and more persons employed and 10 and more persons employed in financial sector. Stratified simple random sampling is used for enterprises 10 to 49 persons employed. Sample was drawn using permanent random numbers. The choice of starting point between 0 and 1 guarantee non-overlap with two large sample surveys: wages survey and SBS survey. Population was stratified according to NACE Rev.2 2-digit level.
Ireland*	We used a combination of matched and unmatched sampling techniques using SAS Enterprises Guide and the use of the random number generator in selecting the sample for the E-Commerce Survey to enterprises. This technique ensures adequate coverage for each NACE industry/employment cell.
Greece	Let h be one of the final strata (Final stratum = Region X Economic Activity X Size Class). The final strata, for the enterprises of the 4th and of the 5th size class, are census strata (take-all), because they comprise the large size units and the element population variance for all variables is high. In each of the final strata, a systematic sample of nh enterprises was selected from the total number of enterprises Nh with equal selection probabilities. In the final strata, the distribution of sampling units was conducted by applying the optimal (Neyman) allocation. The sample size is 4.624 enterprises (sampling fraction 21,4%). The number of enterprises that belong to the census size classes is 443, while the remaining 4.181 enterprises belong to the first three size classes (surveyed on a sample basis). The determination of the sample size was carried out so that at each aggregation of NACE Rev. 2 subsection, the coefficient of variation (CV) of the value estimated variable "number of persons employed who used computers with access to the world wide web at least once a week" does not exceed the value of 5%. Concerning the data processing, the collected data will be inflated, with adequate weights (extrapolation factors) for the estimation of the survey characteristics.
Spain	According to the frame population (extracted from the DIRCE), the statistical units are stratified by crossing the following variables: size, economic activity and location. In every stratum, a systematic selection randomly started is executed, sorting out the enterprises by size and location. The categories of the variables used to stratify are:

	- size: 0 to 2, 3 to 9 (or enterprises with less than 10 employees), and, 10 to 19, 20 to 49, 50 to 99, 100 to 199, 200 to 499, and 500 or more (for enterprises with 10 or more persons employed) - main activity: following the aggregations of NACE classification as followed: 10 a 18, 19 a 23, 24 y 25, 26 a 33, 35 a 39, 41 a 43, 45, 46, 47, 49 a 53, 55, 56, 58 a 63, 68, 69 a 74, 77 + 78 + 80 a 82, 79 + 951 For micro-enterprises the aggregations of NACE classification is as followed: 10 a 39, 41 a 43, 45, 46, 47, 49 a 53, 55, 56, 58 a 63, 68 + 69 a 74 + 77 +78 + 80 a 82, 79 + 951 - Region: all Spanish regions (18) are covered. The total number of initial strata for the general questionnaire is 1.836 (17*6*18), including empty strata for enterprises with 10 or more persons employed and 396 (11*2*18) for enterprises with less than 10 persons employed (following NACE rev.2). The survey is carried out alone.
France	The following strata are used based on activity, reported employment and turnover , held on each statistical source. 10- 19, random sampling 20- 49, random sampling 50 - 249, random sampling 250 - 499, random sampling 500+ Fully enumerated strata Coordination with the sampling of TIC 2012 in order to have half of the sample of the 2012 survey. We add accuracy constraints in the allocation, in order to improve the results for small strata. For national purposes, we exceptionally increased the sample size this year (around 1000 units more).
Italy	Stratified random sample: stratification by economic activities, by size classes (10-49; 50-99; 100-249; 250 and more), by geographical breakdown (21 administrative regions). It was implemented a JALES procedure for negative coordination of the current ICT sample with previous ICT samples and non-overlapping with those of other surveys. We take a census of enterprises included in 250 and more persons employed size class. The sample allocation has been obtained by applying a multivariate and multi-domain technique (Bethel,1989). The solution has been detected in order to be optimal with respect to the domains defined on the basis of NACE Rev2 classification, by imposing stronger constraints on the expected coefficients of variations for the couple variable-domain than those required by ICT Regulation.
Cyprus	The sampling method used was stratified random sampling. Two variables were used for stratification, Nace group and size. There were 16 Nace groups and 3 size groups. Due to the small number of medium and large enterprises, it has been decided to use all those enterprises in the sample. Small enterprises were selected systematically from a list ordered in ascending order by enterprise size. The sample size in each Nace group is determined using the method for estimating proportions. The size of the sample, i.e. 1835 enterprises, could satisfactorily serve the desirable and acceptable confidence intervals limits of the survey results, taking always into consideration time and cost constraints. The total number of all financial enterprises falling in the Nace groups listed, for the size classes 10-49, 50-249, 250+ is small (160 enterprises), thus all enterprises are used.
Latvia	Stratified simple random sampling is used as sampling design. Enterprises are stratified by NACE groups and number of employed groups (10-19, 20-49,50-99, 100-249, 250+). The number of strata is 151.Optimal sample size was calculated for ICT use survey. In each NACE group sample allocation is computed using iterative procedure for Neyman optimal allocation, taking into account the proportion 0.5 for each enterprise and maximum of allowed standard error 3% (for individual NACE groups 2%). "Big" enterprises with number of employee 100 and higher are sampled with sampling fraction 100%. The procedures for the coordination are applied.
Lithuania	The sample is a stratified random sample made by use of register-based data and previous year sample data. The number of persons employed and economy activity were used to stratify the population. There were 25 NACE Rev. 2 groups and 3 enterprise size groups (10- 49, 50 - 249, and 250 +). All enterprises from small groups (less than 20 units) by NACE Rev 2 activities and number of persons employed groups were included into the sample. Within every NACE Rev 2 group the strata of enterprises were formed by number of persons employed groups. Then the Neyman optimal allocation (with variable number of persons employed) was used for determination of the sample size for each stratum specified. In each stratum simple random sampling was used. The final number of strata is 72. The sample was designed with no reference to other surveys
Luxembourg	The frame population was stratified using the following criteria: - 3 size classes (i.e. 10-49, 50-249, 250+ employees) - 15 NACE categories. These categories are mainly based on the aggregates listed in the model questionnaire for possible calculation of national NACE Rev. 2 aggregates, the only difference being that division 56 and group 95.1 are included as separate categories: C10_18, C19_23 C24_25, D26_36, D35_39, F41_43, G45_47, H49_53, I55, I56, J58_63, L68, M69_74, N77_82, S951 The method used for sampling was a stratified random sample, with varying sampling rates depending on size class: - For the two size classes 50-249 and 250+, the sampling rate was 100% (i.e. a census); - For the size class 10-49, the sampling rate was generally fixed 60%, the only exception being I56 with a sampling rate of 30%, as it is not calculated at the national level. The sampling rate of 60% was chosen after simulating the effect of various sampling rates on the coefficients of variation of the 2011 and 2012 data, taking into account historical response rates. Simulations were performed

	on the coefficients of variation for the following indicators: cuse, iacc, web, ade, awsell, aebuy.
Hungary	The sample selection is being done with a stratified sampling scheme. The dimensions used for stratification are (1) NACE Rev.2. activity groups, (2) size of enterprises (according to number of employees), (3) regional breakdown. The strata are formed in the following way. 1. According to the Eurostat recommendation concerning NACE Rev. 2. activity categories we concentrate on those sections, divisions, groups, and classes which are asked by Eurostat. 2. Concerning the number of employees the following categories are being treated separately: 30 (10-19 employees), 40 (20-49 employees), 50 (50-99 employees), 61 (100-149 employees), 62 (150-199 employees), 70 (200-249 employees), 81 (250-299 employees), 82 (300-499 employees), 91 (500-999 employees), 93 (1000-1999 employees), 94 (2000-4999 employees), 95 (5000- employees). 3. Regional dimension of strata construction will consist of the capital and 7 regions. The distribution of the sample among the above defined strata is being done by the following. In cooperation with the informatics department the number of enterprises belonging to the sample frame in all the single strata is known. This information is available for the time period January 2012 and January 2013 (in the latter case provisional data are available). The final sample sizes are determined by the well-known Neyman-type optimal allocation scheme. For this procedure we make use of the information available for previous time-periods concerning the stratum level statistics for different variables. We conduct the rotation during the sample selection process in the following way. A random number is being assigned to all the enterprises in the sampling frame. After that we add 1 or 2 to the value of some enterprises in order to give preference to some of them. Then the process of sample selection is being done by ordering the enterprises by the assigned values and then those possessing the biggest values are chosen. The underlying concept of the distortion process can be described in this way. We intend to include those enterprises first, which were in the sample in 2012 but not between 2009 and 2011. Then, those enterprises will be chosen, which were not included in the sample neither in 2012 nor between 2009 and 2011. Lastly, those enterprises which were in the sample between 2009 and 2011 but not in 2012.
Malta	All enterprises fulfilling the defined NACE and size criteria are selected.
Netherlands	The target-variables (of year t-1) used to allocate the sample were: - the number of enterprises receiving orders online (a qualitative variable) - the number of fte's spent on developing own account software (a quantitative variable). The population/sample was divided into 61 NACE-aggregates and 10 groups of size classes resulting in 610 cells/strata. Ca. 12 000 enterprises were allocated in a way that on a national level the relative confidence intervals for the target variables were as small as possible. Enterprises with 250 or more employed persons were integrally included in the sample. In principle the Neyman allocation was used to allocate the 12 000 enterprises over the different strata. The allocation is therefore based on the variance of the target variable in the relevant strata. The most recent allocation was determined in the autumn of 2011, based on the data of the 2011 survey. A system to 'spread' the administrative burden for enterprises is used to avoid - if possible - that certain enterprises receive many questionnaires in year t and others none or just a few. So, the overlap with other surveys in the same period (e.g. the structural business statistics survey) is minimized.
Austria	Sampling was carried out as stratified random sampling. Two dimensions were used as stratification variables (26 x 3 = 78 strata): a) Main economic activity b) Size classes (3 strata) 10-49 employed persons 50-249 employed persons 250 and more employed persons A full census was carried out among enterprises with 250 and more employed persons. There were higher inclusion probabilities for the following NACE sections: 26 (small and medium sized enterprises), 29-30 (small and medium sized enterprises), 61 (small and medium sized enterprises), 79 (medium sized enterprises) and 95.1 (small and medium sized enterprises). Other surveys were not considered during the sampling process, but the two previous ICT surveys have an influence. If a unit was in the sample in one of the two previous years, the probability to be in the sample again is decreased compared to the other units in the same strata. This is only valid for not complete sample strata.
Poland	Stratified random sample was used. The variables for stratification were the economy activity (NACE rev2, 25 categories) and size of enterprise (as number of employed persons, 3 categories). - the categories of size were: 10-49, 50-249, above 249 employees - the categories for economic acitivty (NACE rev2) The sample enables to obtain estimates with sufficient accuracy for additional group defined as NACE rev2: 26.1-26.4, 26.8, 46.5, 58.2, 61, 62.01, 62.02, 62.03, 62.09, 63.1, 95.1 All large enterprises (having more than 249 employees) were included into sample (3052 units). All units from the last year sample that existed in 2012 and were not present in 2011 year sample were included into the sample (6764 units). Also stratas in the frame contained small number of units were integrally included into the sample (721 units). The calculation of sample sized was based on precision requirements and was equal to 17409 units. No coordination with other survey samples was done.

Portugal	Stratified sample according to the following variables and breakdowns: 1) Economic activity (NACE Rev.2 – at two digit level, except when other aggregation specified) 2) Number of employed persons (0; 1-4; 5-9; 10-49; 50-149; 150-249; 250 and more) 3) Turnover (up to 5000000€; [500000€; 25000000€[; 25000000€ and more) Remark: In 2012 some categories of stratification variables (economic activity and people employed) were broken down based on results of a study in 2011. This study analyzes the main indicators, such as Internet access and website. This procedure intends to reduce the heterogeneity observed in some groups of enterprises. Some requirements: - sample size according to the Neyman-Allocation; - definition of at least five enterprises by strata; - a priori CV is 5% for the economic activity strata and for the employed persons strata; - sample selection according to a random sampling; - priority for the enterprises with less respondent load; - for large enterprises (enterprises with 250 and more employed persons or enterprises with a turnover ≥ 25000000€) a census is adopted.
Romania	In order to insure the representativity, Neyman allocation was used for the allocation of the sample into strata. The sample within each stratum was drawn without replacement. The sample for ICT survey it's in coordination with SBS survey; special for such indicators as: turnover, number of employees etc. The stratification variables (qualitative and quantitative), used in the sampling design for each survey, are the following: a) Main activity. We used the aggregation proposed by Eurostat for the model of the Community survey on ICT usage and e-commerce in enterprises; b) Size class breakdown. The size classes of the number of persons employed are: 10 – 49 (small enterprises); 50 – 249 (medium enterprises); 250+ or more (large enterprises).
Slovenia	Sample design was stratified sampling (SRS within each stratum). We used coordinated sampling. Activity group and size class were used as stratification variables.
Slovak Republic	The sample was created on the base of the same methods, which were used in other surveys organized by SO SR (short-term surveys, business surveys). In comparison with other surveys the boundary between exhaustive and sample survey was changed in the direction upwards. In this enterprises survey the sample was organised in enterprises with 1 to 249 employees whereas in STS and SBS the sample concerns the enterprises with 1 to 19 employees. Stratification sampling was realised according to: - NACE, - NUTS2, - Number of employees.
Finland	Stratified random sample was used. The variables for stratification were the size of the enterprise (as number of persons employed; 4 categories) and economic activity as NACE Rev.2 (33 categories). The categories of size were 4: 10-19, 20-49, 50-99 and 100+. All enterprises in the size class 100+ were included into the sample. The categories of economic activity as NACE Rev.2 classes were: 10-12, 13-15, 16-18, 19-23, 24-25, 26.1-26.4+26.8, 26.5-26.7, 27-28, 29-30, 31-33, 35-39, 41-43, 45, 46.1-46.4+46.6-46.9, 46.5, 47, 49-53, 55, 56, 58.1+59-60, 58.2, 61, 62, 63.1, 63.9, 68, 69-74, 77-78+80-82, 79, 95.1, 64.19+64.92, 65.1+65.2, 66.12+66.19. Sample size was given to be around 4300 enterprises and the sample was allocated to stratas. Simple random sampling was used in each stratum by specified proportions. No coordination with other surveys samples was done.
Sweden	The sampling frame for the survey on ICT usage and e-commerce in enterprises 2013 consist of all active enterprises in the Swedish Business Register (BR) classified into the economic activities (based on NACE Rev 2) 10-82 and 95.1 (excluding 75, 64.11, 64.2, 643, 64.91, 64.99, 65.3 and 66.11) with at least ten employees and classified into the institutional sector codes ((INSEKT 2000): 110-130, 212-215, 217, 219, 221, 223-232 and 240. A version of the BR established in March year 2013 has been used. The stratification of the sampling frame is based on economic activity (NACE Rev 2) and on number of employees. The stratification according to NACE is done in order to fit the domains of study, i.e. subsets of the population, in terms of NACE Rev 2. Each economic activity strata is divided into seven size strata and the limits for the different size strata are based on number of employees. Neyman (optimum) allocation was used to decide sample sizes in each stratum except for the two size strata including the largest enterprises (with at least 250 employees). Enterprises in those size strata are completely enumerated in each economic activity strata. A precision for the estimated proportion, in terms of a standard error, is specified for each economic activity strata. The allocation only considers domains that agree with strata (or aggregates of strata), in other words, domains that cut across strata can not be considered in the allocation. In addition, the allocation does not take possible non-response into account. Stratification is done in two steps, by economic activity d (d =1,2,….,D), and within economic activity by size group h (h=1,2,….H). The sample is positively coordinated with the survey on ICT usage in Enterprises 2012.

United Kingdom	The sample was designed using a Neyman Allocation optimisation program. Neyman Allocation is a sampling methodology which minimises the variance (standard error) of an estimate by selecting an appropriate sample from a population. This process determines the optimal number of enterprises that should be selected within each cell (SIC and employment sizeband), to minimise the standard errors for the questions used to optimise the sample. Each ONS survey begins its permanent random number (PRN) selection at different points on the PRN line to limit overlap with other surveys as much as is possible. By knowing where each survey begins its PRN selection we know where to position new surveys on the line. The larger the business the more significant it is to all surveys and therefore the more ONS questionnaires it receives. To minimise burden on the enterprises in the 10-49 employment sizeband, we excluded any businesses selected by the Annual Business Survey. The following strata are used, based on reported employment held on our business register: 10- 49 random sampling 50 - 249 random sampling 250 - 999 random sampling 1000+ fully enumerated strata
Island	A stratified random sample was drawn, based on Nace Rev.2 and the Eurostat size classes. The sample was drawn in accordance with breakdown: 10-18; 19-23; 24-25; 26-33; 35-39; 41-43; 45; 46; 47; 49-53; 55; 58-63; 68; 69-74; 77-82; 26.1-26.4, 26.8, 46.5, 58.2, 61, 62, 63.1, 95.1. Also size classes 10-49 p.e.; 50-149 p.e.; 250+ p.e. The sample size was 849 enterprises. Applying the policy of not drawing the same enterprise into the sample of the same survey for two consecutive years, and with the sample of the survey year 2012 covering 732 out of the 1514 enterprises, the sample of 2013 was thus partly determined by the sample of the survey year 2012. However, standards of minimum coverage for each stratum of Nace class breakdown + size class were also ensured (with a minimum of 9-10 enterprises per stratum, or else census). Some enterprises from the sample of the previous year thus ended up in the sample of this year. 115 enterprises that were in the sample of 2012 were redrawn, using stratified random sampling, or census. Of the remaining enterprises that were not in the sample of the survey year 2012, 47 were drawn out of the sample.
Norway	The sample was drawn from all enterprises with at least ten persons employed in the Central Register of Establishments and Enterprises. It was stratified by industry and size of the enterprise measured by employment. Among enterprises with 10-19 employees the sample was 7.5 per cent, among enterprises with 20-49 employees the sample was 15 per cent, among enterprises with 50-99 employees the sample was 50 per cent, among enterprises with 100-249 employees the sample was 75 per cent and among enterprises with at least 250 employees all enterprises were included in the sample. For industries the percentage are adjusted to be at least 5 units in each stratum if possible.
Croatia	Business Register from Central Financial Agency (FINA) is available in June, for the previous year. Therefore, data collection will be conducted in two stages: frame population for medium and large enterprises was based on previous year sample and these enterprises will be surveyed in first stage. There is very small chance that some of these enterprises will be shut down, so we estimated that last year population is good enough to start the dta collection process. Small enterprises will be surveyed in second stage upon the final update of Business Register, when the actual sampling is done. Reason for this is that certain number of small enterprises will be shut down due to economic crisis, therefore it is not possible to use Business Register from previous year for sampling frame as it would generate incorrect data. Frame population in tables 5.3 and 6.2 will be updated as soon as final data from Business Register is avialable, together with X module variables. Estimated population size is about 4500 enterprises. Stratified random sampling will be used for small enterprises (with 10-49 employees), NACE Rev. 2 economic activity was used as stratification variable (26 strata in total). In each particular stratum we used random sample selection of units. Census approach was used for medium and large enterprises (50 and more employees). None of the procedures for the coordination or non-overlapping with samples of other surveys were applied.
Macedonia	All enterprise with 50 and more employees are included in the sample, and stratified random sample with square root allocation is used for the group 10-49 divided by 28 groups of economic activities.

I.2.3.1 STRATIFICATION

The recommendation is to use a stratified sample of enterprises with the aim to form groups of units characterised, in terms of the variables collected in the survey, by maximum homogeneity within the group and maximum heterogeneity between the groups.

The background variables X1 - *"Main Economic Activity"* (in terms of NACE) and X2 - *"Average Number of Persons Employed"* (in terms of size classes) should be used for the stratification of the sample. These variables, according to the Council Regulation (EC) No 177/2008 on business

registers for statistical purposes, are presented in the sampling frame and can, therefore, be used to stratify the sample *à priori*.

The purpose of the stratification by main economic activity and size class is to assure *à priori*, accurate results for the breakdowns. In fact, if the sample is not stratified by these variables, the number of enterprises that casually end up in some NACE category, size class, or region might be too small to produce accurate results.

For the definition of the categories and level of detail of the stratification variables, the level of dissemination concerning NACE-aggregates and size-class has to be taken into account (please refer to the explanatory notes of the background variables in chapter I.1.6.12). The minimum stratification of the frame population has to be consistent with this level of dissemination, meaning that the most detailed level of categories of all the breakdowns foreseen have to be defined by adding the aggregated results of the different strata (cells) distinguished in the stratification.

Given the level of dissemination, the minimum level of stratification concerning NACE-aggregates and size-class is defined in the following table. For reasons of estimation efficiency, additional stratification by size-class and NACE-aggregates may be useful. Under certain conditions the efficiency of the estimator can be improved by defining more detailed (homogeneous) groups of NACE-aggregates or size-classes. Efficiency means here a higher reliability of the estimates for the same overall sample size.

In the figure below, the "**x**" in bold represent the minimum level of stratification for the mandatory scope of the enterprise survey (81 strata). The "x", in normal text, represents the minimum level of stratification for the optional scope of the survey (an additional 54 strata)

For providing results that are of a sufficiently high quality after aggregation over all EU countries, strata on the lowest level of the NACE Rev. 2 groupings (e.g. 10-12) in the figure below might not be needed, in particular for the smaller Member States. In case results on national level are to be released, this level of detail in the stratification is however required.

				Total				
				Micro-Enterprises		Small, Medium and Large Enterprises		
				Small Micro-Enterprises	Big Micro-Enterprises	Small and Medium Enterprises		Large enterprises
						Small Enterprises	Medium Enterprises	
				1 to 4 persons employed	5 to 9 persons employed	10 to 49 persons employed	50 to 249 persons employed	250 or more persons employed
Total NACE Rev. 2	Manufacturing, energy and construction	10 – 18	10 – 12	x	x	**x**	**x**	**x**
			13 – 15	x	x	**x**	**x**	**x**
			16 – 18	x	x	**x**	**x**	**x**
		19 – 23		x	x	**x**	**x**	**x**
		24 – 25		x	x	**x**	**x**	**x**
		26 – 33	26	x	x	**x**	**x**	**x**
			27 – 28	x	x	**x**	**x**	**x**
			29 – 30	x	x	**x**	**x**	**x**
			31 – 33	x	x	**x**	**x**	**x**
		35 – 39		x	x	**x**	**x**	**x**
		41 – 43		x	x	**x**	**x**	**x**
	Non-financial services	45 – 47	45	x	x	**x**	**x**	**x**
			46	x	x	**x**	**x**	**x**
			47	x	x	**x**	**x**	**x**
		49 – 53		x	x	**x**	**x**	**x**
		55	55 – 56	x	x	**x**	**x**	**x**
		58 – 63	58 – 60	x	x	**x**	**x**	**x**

			61	x	x	**x**	**x**	**x**
			62 – 63	x	x	**x**	**x**	**x**
		68		x	x	**x**	**x**	**x**
		69 – 74		x	x	**x**	**x**	**x**
		77 – 82	77 – 78 + 80 – 82	x	x	**x**	**x**	**x**
			79	x	x	**x**	**x**	**x**
	ICT sector	26.1 – 26.4 + 26.8 + 46.5 + 58.2 + 61 + 62.01 – 62.03 + 62.09 + 63.1 + 95.01	95.1	x	x	**x**	**x**	**x**

I.2.3.2 SAMPLE SIZE

Calculation of sample sizes should take into account that it has to ensure representative results for all the estimates produced. In particular calculation of sample size should take into account that each statistic has to be tabulated by NACE category and size class.

As budgets are limited, the design of samples involve making trade-offs along various dimensions. Larger samples make it possible to analyse sub-groups in depth but every interview or questionnaire treated increases survey costs.

On the basis of the previous considerations, it is suggested to adopt a mixed view, based on both cost and organisational criteria and on an evaluation of the sample errors of the main estimates on a national level and with reference to each of the territorial domains and to each of the breakdown variables of interest.

The calculation of sample sizes should be based on precision requirements (see I.2.3). On this basis countries should decide on sample design and calculate the sample sizes in order to receive estimates with sufficient accuracy and within possible budgetary constraints.

In practice, the sample size is usually calculated by applying the desirable overall reliability of the estimate to a target-variable. This target variable can be one of special relevance for the survey or one that correlates well with the majority of the variables to be collected. The resulting sample size is set by the dispersion of this target-variable. However, some times for several reasons, e.g. the administrative burden of enterprises, often a maximum number of enterprises to be surveyed is defined. This number of enterprises is allocated to the different strata in such a way that the reliability of the estimates is optimized. An efficient way to allocate a specified number of enterprises to the different strata is the so-called Neyman-allocation, meaning that the number of enterprises is allocated to the relevant strata inversely proportional with the variance of a specified target-variable in these strata.

$$n_h = n \times \frac{N_h \times S_h}{\sum N_h \times S_h}$$

Where: n_h is the number of units in the sample in stratum h;

n is total sample size;

N_h total number of units in the frame population for stratum h;

S_h true standard deviation in stratum h for the relevant variable.

Estimates of the variance of the target variable might come from a pilot-survey or the survey from a previous year.

Additional to the outcome of the Neyman-allocation, a minimum number of enterprises in each stratum can be specified. For larger enterprises one can decide to include them integrally in the survey. However, for qualitative questions this is less crucial than for quantitative variables like for example those referring to the turnover.

More advanced sampling techniques may be used as long as it is possible to calculate the normal statistical variables and it meets the output specified in this manual.

By specifying a maximum number of enterprises in the sample it is useful to anticipate - based on experience with a previous survey or another comparable survey - a response rate. If experience shows that only 50 per cent of the enterprises addressed, actually respond in a usable way, the sample size should be adapted to this response rate, meaning should be doubled.

The table below gives an overview of the sample sizes reported by the national statistical institutes in the Quality Reports for the latest national surveys. Note that due to the adaptation of the sample size to specific national needs (e.g. production of regional estimates), the sample sizes can't always be compared; countries also include in different degrees the optional parts in terms of NACE groupings and employment size classes in their survey. Note that the scope in terms employment size class – inclusion in the survey or not of micro-enterprises - has in the countries an impact on the – average – response rate. For most countries, the final or net sample size was between 3000 and 6000 elements. The response and non-response as well as the quality (in terms of standard errors) are discussed in more detail in the next chapters.

Table: Overview of the sample sizes and response rates[18]

Source: Quality reports for the survey on ICT usage and e-commerce in enterprises (2013)

Country	Scope in terms of size classes of number of persons employed	Gross sample size (1)	Net sample size used for tabulation (2)	Response not usable for tabulation*	Non-response	Net response rate (2)/(1)
Belgium	5+	7.717	3.717	47	3.941	48%
Bulgaria	10+	4.597	4.118	335	144	90%
Czech Republic	10+	7.746	6.362	131	1.253	82%
Denmark	10+	4.224	3.945	162	117	93%
Germany	1+	20.000	9.239	:	10.761	46%
Estonia	10+	3.020	2.740	30	250	91%
Ireland	10+	6.788	2.844	656	3.288	42%
Greece	10+	4.682	3.100	517	1.065	66%
Spain	1+	26.090	18.344	4.048	3.698	70%
France	10+	13.387	10.044	342	3.001	75%
Croatia	10+	4.500	3.431	22	1.047	76%
Italy	10+	32.378	19.114	1.456	11.758	59%
Cyprus	10+	1.835	1.492	151	192	81%
Latvia	10+	3.804	3.562	94	148	94%
Lithuania	10+	2.965	2.924	27	14	99%

[18] Sampling based on NACE Rev. 2

Luxembourg	10+	2.289	1.869	15	405	82%
Hungary	10+	6.953	6.231	107	615	90%
Malta	10+	1.316	1.134	:	182	86%
Netherlands	10+	11.107	7.896	290	2.921	71%
Austria	10+	5.148	3.196	52	1.900	62%
Poland	10+	18.571	14.399	3.943	229	78%
Portugal	1+	6.700	5.711	379	610	85%
Romania	10+	10.998	9.684	541	773	88%
Slovenia	5+	2.804	2.313	45	446	82%
Slovak Republic	1+	4.281	3.899	24	358	91%
Finland	10+	4.226	3.067	100	1.059	73%
Sweden	10+	4.200	3.419	11	770	81%
United Kingdom	10+	7.813	5.813	1.160	840	74%
Island	10+	843	670	:	:	79%
Norway	10+	4.306	4.155	:	151	96%
Macedonia	10+	2.093	1.406	23	664	67%

I.2.3.3 WEIGHTING – GROSSING UP METHODS

The grossing up method, or weighting procedure, to be adopted for the production of figures for the total target population is determined in the first place by the sampling design used. The weighting factors are calculated taking into account in particular the probability of selection of each unit in the sample.

In this chapter, the explanation of the several weighting methods for the several types of variables will assume the selection of a stratified random sample, which is the method recommended in this manual for the enterprise survey. The formulas have to be adapted if a different sampling design is used in the survey.

In the second place, the grossing up method is determined by the type of variable collected and the statistics produced with those variables. In the enterprise ICT usage survey there are two types of observation variables:

1) Qualitative / Binary variables, e.g. *"Did your enterprise have access to the internet?"*; and
2) Quantitative variables, e.g. *"Please indicate the percentage of the number of persons employed who used computers with access to the www."*.

The enterprise ICT survey is dominated by qualitative questions and most results are published as percentages of the number of enterprises (total number of enterprises, enterprises with access to the internet, etc.). This leads to conclusions or observations like "x percent of all enterprises have access to the internet" or "y percent of all enterprises with access to the internet sell online". To produce these results the observations are **weighted by the number of enterprises** in the stratum to which they belong.

When presenting overall results, these conclusions or observations are usually dominated by the results of the small and medium-sized enterprises because they are the largest group expressed in number of enterprises. However, the majority of the labour force may well work in bigger enterprises, where ICT usage is qualitatively and quantitatively different from the others. For this reason, when looking especially at more employment related issues, the results should also be additionally **weighted by the number of persons employed**. In this case the ICT usage of larger enterprises gain influence on the overall result. Weighting by employed persons leads to

conclusions or observations like "x percent of the employed persons work at enterprises with access to the internet".

The quantitative questions are in general collected in percentage terms, e.g. percentage of persons employed using computers, percentage of turnover resulted from orders received via computer networks. In order to produce overall averages of these percentages for the total population they also have to be additionally **weighted** by the base variable of those percentages, i.e. by the **number of persons employed** if the variable is percentage of persons employed, by the **total turnover** if the variable is percentage of turnover, etc. Quantitative variables in the e-commerce module related to sales/purchases should be weighted by total turnover/total purchases.

When the quantitative variables are collected in absolute terms and not in percentages, the variables don't need to be additionally weighted by the base variable of the percentages.

In the remaining of this chapter the two weighting methods are explained in detail.

Basic weighting by number of enterprises

Assuming that a stratified random sampling is used, the estimator of a total in the population based on the sample is:

$$Y = \sum Y_h \,, \qquad (1)$$

$$Y_h = \frac{N_h}{n_h} \sum_{i \in h} y_{hi} \,, \qquad (2)$$

where Y is the estimated total value of variable *y* for the total population

Y_h is the estimated value of variable *y* for the total population in stratum *h*;

N_h total number of units in the frame population for stratum *h*;

n_h is the number of units in the sample in stratum *h*;

y_{hi} is the value of variable *y* of enterprise *i* in stratum *h*.

In the case of the quantitative variables in absolute terms (not in percentage terms), y_{hi} is the value of that variables. In the case of the qualitative variables, y_{hi} assumes the value 1 if the answer is "YES" and 0 if the answer is "NO". This way the total of this variable is the number of enterprises having answered "YES".

To compute the percentages, these totals are divided by either the total number of enterprises, in the case of the qualitative variables, or by the total of the quantitative variable for all enterprises, in the other case.

For the total population:

$$Y^p = \frac{\sum Y_h}{N} \qquad (3)$$

For each stratum:

$$Y_h^p = \frac{\frac{N_h}{n_h}\sum_{i \in h} y_{hi}}{N_h} \qquad (4)$$

In the grossing up, each enterprise *i* in strata *h* has the following weight

$$w_{hi} = \frac{N_h}{n_h}, \qquad (5)$$

which gives how many enterprises in the population this sampled unit represents.

Because the weighting factor w_{hi} is based on the number of enterprises, this is called weighting by number of enterprises.

Additional weighting by an auxiliary variable (number of persons employed, turnover, etc.)

As mentioned before, the additional weighting by an auxiliary variable z has the purpose, in the case of the qualitative variables, of computing indicators of the type "percentage of variable z", instead of "percentage of enterprises". For example "percentage of persons employed". In the case of the quantitative variables collected in percentage terms, this additional weighting by the base variable of the percentages is necessary to gross up the results.

In terms of absolute values, and not percentages, this means that the indicators computed using this additional weighting are of the type "number of persons employed", "turnover" or "purchases", instead of "number of enterprises" and the percentages are computed afterwards using the totals of these variables.

The formula for the computation of these indicators of the type "percentage of variable z" is actually the same in both types of variables, but the reasoning is slightly different. For this reason, they will be explained separately.

Qualitative / Binary variables

In this case we compute indicators of the type "percentage of variable z, for enterprises with y equal to some value" (where y is the qualitative / binary variable). For example "percentage of persons employed, for enterprises with internet access", in which case the qualitative / binary variable "Do you have internet access" equals "Yes". Or, in absolute values, "number of persons employed, for enterprises with internet access".

One way of doing it for the absolute value is simply to select the enterprises for which the qualitative variable equals the value of interest (in the example above, "Yes") and gross up the number of persons employed in those enterprises:

$$Y = \sum Y_h,$$

$$Y_h = w_{hi} \cdot \sum_{i \in h} z_{hi} y_{hi} \qquad (6)$$

where Y_h is the estimated absolute value of persons employed in enterprises, for which the qualitative variable equals the value of interest, for the total population in stratum *h*;

N_h total number of units in the frame population for stratum *h*;

n_h is the number of units in the sample in stratum *h*;

z_{hi} is the number of persons employed in enterprise *i* of stratum *h*.

y_{hi} is a binary variable assuming 1 for the enterprises for which the qualitative variable equals the value of interest, and 0 otherwise.

w_{hi} is the weight of enterprise *i* in stratum *h*., as defined in formula (5).

For the percentage, we divide this value by the grossed up total of persons employed in all enterprises:

$$Y_h^P = \frac{w_{hi} \cdot \sum_{i \in h} z_{hi} y_{hi}}{w_{hi} \cdot \sum_{i \in h} z_{hi}} \quad (7)$$

Which simplifies to:

$$Y_h^P = \frac{\sum_{i \in h} z_{hi} y_{hi}}{\sum_{i \in h} z_{hi}} \quad (8)$$

Quantitative variables

In this case the indicators computed are of the type "percentage of variable z, in relation to another variable", where the first one is a parcel of the second one. For example "percentage of persons employed with access to internet, in relation to the total number of persons employed". When defined in absolute values, it refers to the numerator only: "number of persons employed with access to internet".

One way of doing it for the absolute value is to use the individual percentages of each enterprise (as it is collected), multiply it by the base of that percentage (in the example above, "total number of persons employed") and gross it up. The total value for each stratum *h* would then be:

$$Y_h = w_{hi} \cdot \sum_{i \in h} z_{hi} y_{hi} \quad (9)$$

where Y_h is the estimated absolute value of persons employed with access to the internet, for the total population in stratum *h*;

N_h total number of units in the frame population for stratum *h*;

n_h is the number of units in the sample in stratum *h*;

z_{hi} is the number of persons employed in enterprise *i* of stratum *h*.

y_{hi} is a binary variable assuming 1 for the enterprises for which the qualitative variable equals the value of interest, and 0 otherwise.

w_{hi} is the weight of enterprise *i* in stratum *h*., as defined in formula (5).

For the percentage, we divide this value by the grossed up total of persons employed in all enterprises of stratum *h*:

$$Y_h^p = \frac{w_{hi} \cdot \sum_{i \in h} z_{hi} y_{hi}}{w_{hi} \cdot \sum_{i \in h} z_{hi}} \qquad (10)$$

Which simplifies to:

$$Y_h^P = \frac{\sum_{i \in h} z_{hi} y_{hi}}{\sum_{i \in h} z_{hi}} \qquad (11)$$

Formulas (6) and (9) are identical, and formulas (8) and (11) as well. Therefore, the grossing up method when an additional weighting by an auxiliary variable is used is the same for qualitative variables and quantitative variables, when collected in percentage values.

Further information: (Eurostat 2006). Handbook on methodological aspects related to sampling designs and weights estimation.

I.2.4. Survey type

Data collection method

Face-to-face interviews, telephone interviews, postal surveys and web-based or electronic surveys are all possible techniques of collecting data. Next the main strengths and weaknesses of each method are presented.

Postal survey

Strengths:

They are relatively inexpensive to administer. You can send the exact same instrument to a wide number of companies. They allow the respondent to fill it out at their own convenience. Mail survey doesn't allow personal contact between the researcher and the respondent; consequently the researcher impressions can't influence the response of the respondent.

Weakness:

But there are some disadvantages as well. Response rates from mail surveys are often very low and mail questionnaires are not the best vehicles for asking for detailed written responses. They also require a strong monitoring system (of sending reminders, etc.)

Personal interviews (face-to-face)

Strengths:

Interviews are a far more personal form of research than a postal survey. In the personal interview the interviewer works directly with the respondent (controlled situation). Unlike with mail surveys the interviewer has the opportunity to probe or ask follow-up questions. Interviews are generally easier for the respondent, especially if what is sought is opinions or impressions (individual behaviour can be observed and exchange of material/information between interviewer and respondent is possible). They have a very good response rate.

Weakness:

Interviews can be very time consuming and they are resource intensive. The interviewer is considered a part of the measurement instrument and interviewers have to be well trained in how to respond to any contingency. They are very expensive and consequently inapplicable for global and big surveys.

Telephone interviews:

Strengths:

Telephone interviews enable a researcher to gather information rapidly. Like personal interviews they allow for some personal contact between the interviewer and the respondent (questions can be repeated and interpreted). They allow the interviewer to ask follow-up questions.

Weakness:

Some small companies don't have publicly-listed telephone numbers. Some don't have telephone. Telephone interviews have to be relatively short or people will feel imposed upon. The inability to use visual aids is also a weakness. Difficulties occur in case information needs to be collected by the respondent (e.g. from the accounts).

Electronic survey:

Strengths:

With the growth of the internet (and particular the World Wide Web) and the expanded use of electronic mail for business communication, the electronic survey is becoming a more widely used survey method. Electronic surveys can take many forms. They can be distributed as electronic mail messages sent to potential respondents. They can be posted as World Wide Web forms on the internet. They can be distributed via publicly available computers. In some cases electronic surveys are placed on laptops and respondents fill out a survey on a laptop computer rather than on paper. It is less expensive to send questionnaires online than to pay for postage or for interviewers. It is easier to make changes to questionnaire, to copy and sort data. Questionnaires can be delivered to recipients in seconds rather than in days as with traditional mail. Research shows that respondents may answer more honestly with electronic surveys than with paper surveys or interviews. Due to the speed of online networks participants can answer in minutes or hours and coverage can be global. Since many enterprises prefer to report information electronically to more traditional data delivery, offering electronic surveys generally increases the total response rate.

Weakness:

Population and sample is limited to those with access to computer and online network. Thus, in reality electronic surveys can only be a supplement to other survey vehicles. The use of several survey vehicles simultaneously may prove not cost efficient due to increased logistical work. The open natures of most online networks make it necessary to invest considerable time and expertise in guaranteeing acceptable anonymity and confidentiality. Constructing the format of a computer questionnaire can be more difficult the first few times due to a researcher's lack of experience. More instruction and orientation to the computer online systems may be necessary for respondents to complete the questionnaire. Electronic surveys can have a high technical problem with hardware and software.

Nevertheless, in practice almost all EU Member-States collect the data through self-administered postal survey, more and more combined with the possibility to fill-out and post the questionnaire

electronically. The table below presents the information on the data collection methods used in the latest national surveys.

Table: Overview of the data collection methods

Source: Metadata/Quality reports for the survey on ICT usage and e-commerce in enterprises (2013)

Country	Stand alone survey?	Survey vehicle	Mandatory participation?
Belgium	Yes	First letter: invitation letter for web survey (Self-Administered Web Questionnaire) Reminder: paper questionnaire included (Self-Administered Web Questionnaire + Self-Administered Paper Questionnaire)	No
Bulgaria	Yes	Combined web- based and face to face interviews.	Yes
Czech Republic	Yes	E-mail with questionnaire in pdf format or self-administered mail out and mail or e-mail back. Electronic questionnaire (computer assisted program available on the web site) for electronic completion was developed.	Yes
Denmark	Yes	Self-administered web survey. Only possible to use ordinary mail by exception.	Yes
Germany	Yes	It is a combination of a self-administered mail and a web survey – respondents can either choose. The web-questionnaire is provided barrier-free and not barrier-free.	No
Estonia	Yes	1) Online channel for compiling and transmitting reports (eSTAT), available starting December 2012 2) Spreadsheet reports transmitted through the web (questionnaires downloaded from the web page) 3) Paper reports sent by mail or fax (questionnaires printed from the web page)	Yes
Ireland	Yes	Web survey	No
Greece	Yes	A combination of face-to-face interviews, self-administered mail and telephone interviews	Yes
Spain	Yes	This year, as it was done last year, a letter is sent with a password to fill in the survey via internet and the questionnaires are sent via mail only under request of the enterprises once they have received the first letter. Moreover, enterprises can call to a free-charge telephone line to get clarification of the questions and to be helped with the filling out of the questionnaire. In all cases there is a system of reminders, including letters and phone calls. The survey is mandatory and enterprises can be penalized in case of no response.	Yes
France	Yes	Web survey (using phone for reminders and further information). If the enterprise doesn't want to answer using the web site, self-administered mail survey.	Yes
Italy	Yes	Stand alone survey is based on self-compilation of electronic questionnaires (html questionnaire). Enterprises have possibility to fill on-line the questionnaire using a dedicated web site; an ID code and a password are provided to enterprises with mailed letter.	Yes
Cyprus	Yes	Face-to–face interview. A letter notification was sent by mail or e-mail to the IT manager of the enterprise just before the beginning of the data collection period. The IT manager was informed that an enumerator would contact him in order to arrange a meeting for the completion of the questionnaire. As an alternative the questionnaire could	Yes

		be completed on the website of the Statistical Service through an electronic version of the questionnaire. Computer Assisted Personal Interviewing (CAPI). The electronic questionnaire was developed in BLAISE.	
Latvia	Yes	The enterprises were able to submit the filled web-questionnaires electronically using our on-line submission system via Internet in CSB web-site (90% of net sample) or by fax/mail (10% of net sample).	Yes
Lithuania	Yes	An e-questionnaire with a possibility to fill online or print and fill in a paper version (combination of a postal survey and e-questionnaire). Enterprises without internet access are approached by traditional mail.	Yes
Luxembourg	Yes	Self-administered mail or web survey.	Yes
Hungary	Yes	Electronic survey via ELEKTRA system. ELEKTRA is a new project carried out by the Statistical Office which enables data suppliers to submit data completely electronically through a web form.	Yes
Malta	Yes	Surveys to all enterprises employing 10 or more employees were sent by post. Interviewers were utilised to solicit response and carry out face to face interviews with respondents.	Yes
Netherlands	Yes	Web survey (over 90% of respondents). Respondents only receive a paper form at their own request.	No
Austria	Yes	A web questionnaire is offered for transmitting the data to Statistics Austria via the Internet. So, a letter including a supplementary sheet for using the web questionnaire was sent to the respondents. As an alternative to the web questionnaire, the respondents are offered a PDF to be filled out which can be downloaded from Statistics Austria's website. A telephone hot line for the respondents has been established. Additionally, two written reminders (paper questionnaire including the explanatory notes, an accompanying letter, a pre-paid addressed envelope for returning the questionnaire in addition of a supplementary sheet for using the web questionnaire) have been sent out. For the first time, enterprises were only sent the login data for the web questionnaire in the first stage of the survey. After having a lower response rate at the first deadline, it was decided to enclose also a paper questionnaire to the two reminders.	No
Poland	Yes	Electronic questionnaire via statistical on-line portal with possibility of paper questionnaire.	Yes
Portugal	Yes	The national survey combines two modes of data collection: postal and electronic. The enterprises can send either the information by post or download the web questionnaire from the NSI site and return it electronically.	Yes
Romania	Yes	Self administrated mail survey. In case of non-responses the telephone interview will be used.	Yes
Slovenia	Yes	The survey type was a combination of web and self-administered mail survey. Enterprises were asked to fill out electronic version of the questionnaire in the eSTAT application. Paper questionnaire was sent out to enterprises which did not have the possibility to fill out the electronic questionnaire, due to lack of resources or knowledge required to register in the eSTAT application. Majority of questionnaires was filled out in electronic form. A brochure explaining the meaning of the survey with results from 2012 survey was also prepared.	Yes
Slovak Republic	Yes	The survey was included in the Program of state statistical surveys in 2012 – 2014. The questionnaires as well as the reminders were sent to the enterprises by post. The after the respondents were reminded also by phone.	Yes

Finland	Yes	Responses were requested to be given via web as principal method. No paper questionnaires were sent in the first place. However the possibility was given to respond in paper and getting the paper forms by printing the Questionnaire in web or ask for the paper questionnaire.	Yes
Sweden	Yes	Self-administered mail survey. The enterprises are invited to fill in a web questionnaire but have the possibility to ask for a paper version of the questionnaire. In the second reminder, a paper questionnaire will be sent out to the enterprises. Enterprises in the financial sector only have the possibility to fill in a web questionnaire.	Yes
United Kingdom	Yes	Self-administered mail survey.	Yes
Island	Yes	Combination: web-based, postal survey, and telephone interviews.	No
Norway	Yes	Self-administered web questionnaire.	Yes
Montenegro	Yes	Telephone interview	No
Croatia	Yes	Web survey – introduction letters were send by mail to all participants (enterprise managers) with guidelines for completion of web survey. In extreme situations (e.g. no internet connection available) respondents were allowed to complete survey on printed questionnaire or via CATI interview.	Yes
Macedonia	Yes	In general it is self-administered mail survey, combined with gathering data by e-mail or by telephone calls.	Yes
Serbia			
Turkey	Yes	Combination of techniques (face-to-face interview and self-administered mail survey) and also we have implemented electronic data collection via the Internet for this surveys in order to reduce the response burden of enterprises.	No

Independent versus embedded survey

The ICT usage survey in enterprises is not a short survey. If this survey is attached to another survey it is possible that the complete questionnaire is too long. The respondent needs too much time to answer all the questions. Consequently the quality of the answers decrease and the rate of non-response increases. Arguments in favour of embedding the ICT usage survey into an existing business survey are reduction of its total cost and the possibility to cross ICT usage data with other business information.

Mandatory survey versus voluntary survey

Voluntary surveys are usually cheaper, quicker and easier to manage. In mandatory surveys it is implied to send the questionnaire several times to the respondent, you have to write a reminder letter and sometimes you have to send an administrative offence. You have to wait a long time for all responses and your costs increase. The advantage of a mandatory survey is that your response quote is potentially 100%. In practice this appears to be only rarely reached. But a voluntary survey can settle this argument by increasing the sample size and sending also reminder letters to the respondents.

For a majority of the national surveys, response for enterprises in the sample is compulsory (see table above).

Contact person of the survey

In most cases the IT manager is the appropriate recipient of the survey on ICT usage in enterprises. However, not all small enterprises have an IT manager; in these cases either the owner or the general administrator should answer. In bigger companies it is helpful to send the questionnaire to the general management. They have an overview of all functions within the company. Sometimes you find a contact person of the company in your business register. In this case it is useful to send the questionnaire to this person.

I.2.5. Questionnaire – Data collection tool

To enhance the comparability across countries, Eurostat provides a Model Questionnaire to the NSI's. It is recommended to adopt the Model Questionnaire for the national survey. However, for practical reasons (for instance linked to the survey vehicle) this may not be appropriate. In this case, the statistical institute should nevertheless make sure that the routing and the filtering is followed.

Eurostat designs the Model Questionnaire in close cooperation with the national statistical institutes, the main users and in coordination with other organisations such as the OECD. The results of previous surveys are taken into account, for instance to assess the relevance of certain questions or items, or to assess whether a variable needs to be collected on an annual basis. For reasons linked to the annually renewed implementing measures for Regulation 808/2004, the Model Questionnaire needs to be finalised 9 months before the survey takes place. In practice, this means that the questionnaire reaches its final status in March-April of the year (**T-1**) preceding the reference year (**T**). The obvious disadvantage is that "last-minute" policy needs can't be included or that the results of the **T-2** survey can't be fully taken into account. The advantage is that this gives the NSI's the necessary time to translate the questionnaire and implement the survey.

The Model Questionnaire for the survey on ICT Usage and e-Commerce in Enterprises can be found in Annex I.3.1.

I.2.6. Quality control systems

Quality control systems are of course country-specific as most statistical institutes have standard procedures and guidelines for plausibility checks or logic tests of datasets.

Such controls can be executed on-line, at the moment of the data capture by the interviewer or the data entry in the statistical institute, or after the data entry process (a program checks the data and highlights the errors to be checked or corrected). On-line tests have the advantage that the errors can be corrected immediately (a lot of errors will be simple data-entry errors, for instance typing 17773 instead of 1773), a disadvantage is that one needs data-entry staff that is familiar with the concepts in the survey (in case they have to decide if a certain value will be accepted or not). The latter case of course doesn't apply with CATI or CAPI where the interviewer immediately enters the data.

Below, some of the most common errors or problems are briefly discussed.

Measurement error

There are a number of sources of measurement error: survey instruments (questionnaire), the respondent, the information system, the mode of data collection, the interviewer. This paragraph focuses on the latter, other types can back below. The first and probably most important step in the statistical process is the data collection or field work. One can design excellent questionnaires, excellent capture tools, excellent imputation methods and excellent data analyses, but as with any process the adagium "Garbage In Garbage Out"

also applies to statistics. If the interviewer is not well-trained, one can't expect input data of a good quality. A more important problem is usually the reliability of the interviewers. The interviewers may have an interest in finishing the interviewers in the shortest possible time. This can lead to interviews where certain questions are skipped (e.g. the enterprise looks old-fashioned, let's fill in none of the more sophisticated e-business functions are used) or interviews that never took place (e.g. the enterprise is so small, let's put down it does not use computers). Therefore the sensibilisation of interviewers (which can include threats of non-payment or exclusion for future interview rounds) and the follow-up is extremely important. The follow-up can be a check for coherence and consistency of the answers or follow-up phone calls to verify with the respondents whether the interview has actually taken place or whether the interview was of an acceptable quality (e.g. the respondent confirms that the interview only took 4 minutes, but on the questionnaire answers to 60 questions were recorded).

Invalid response

This aspect is relatively unimportant in the ICT usage survey as most answers are limited to Yes or No. However, it is possible that several items were ticked in questions where only one answer is expected (e.g. Yes **and** No).

Relationship error

Comparing the answers across the survey can reveal inconsistencies between the answers. The routing and filtering should normally guarantee that respondents are not presented with questions they can't answer. It is nevertheless possible that e.g. an enterprise not using electronic transmission of data in a format suitable for automated processing (question C1) indicates that it is sending e-invoices (in question C2) or that an enterprise with 20 persons employed (X2) indicates that the number of persons employed using computers with access to the www is 200 (question B2). In the first case, this is most probably an error related to the understanding of the questions by the respondent. In the latter case it is more likely that a "typing" error occurred.

Compulsory question left unanswered

Again, the routing and filtering should avoid this type of errors. CATI or CAPI programs usually improve the quality of the data capture in this respect. With traditional interviews or self-administered mail surveys, this error is more important.

Suspicious values

This can occur when the individual responses are compared with the average scores within a stratum and unexpected answers are found. In practice it will however be difficult to determine whether it concerns a mistake or an outlier.

In terms of quality of the survey *as such*, the methodology and outcomes of the survey can be benchmarked against other surveys:

Representativeness

It can be useful to do an *ex-post* check of the representativeness of the sample, e.g. does the sample have a representative size class distribution, is there some variability in the economic activities?

Year-to-year comparison at aggregate level

Comparing the results for the current year with the previous survey can also reveal quality problems where the growth is outside the range of the expected growth (e.g. the proportion of enterprises using the internet decreases …). In such cases, it is of course possible that the problem stems from the previous survey exercise …

For this purpose, it can be interesting to produce some simple tabulation of the survey results.

Coherence or consistency with other surveys

The results can be compared with results from related survey or studies. However, in case inconsistent results are observed, it is not always easy to identify which survey gave the "wrong" results.

I.2.7. Data processing

This chapter mainly discusses the treatment of non-response. Although the grossing-up methods can be considered as a part of the *data processing*, this topic is discussed above.

I.2.7.1 MISCLASSIFICATION TREATMENT

Misclassification occurs when an enterprise is included in the survey (because it belongs to a size class and sector of activity covered by the survey, at least according to register data used for stratification) but later information indicates that it should not have been included. In other cases misclassification of enterprises means that enterprises that (according to the registers) are classified in a size class or NACE category should have been classified in another size class or NACE category. The misclassification will then possibly lead to the fact that the enterprises should belong to another strata than the one used for stratification.

Such a situation can arise due to frame population imperfections. Frame imperfections can occur when there is a time lag between the actual situation for an enterprise and the information available in the registers. It often takes a certain period of time to update register information after a change in the number of employed persons or a change of sector of activity has occurred.

This time lag in updating register information implies that there is a difference between the target population (i.e. the population that the survey intends to cover) and the frame population (i.e. the population that the survey actually covers based on information available in registers).

Misclassification errors are part of the frame deficiencies *"… Coverage errors (or frame errors) are due to divergences between the target population and the frame population. Three types of coverage error are distinguished: (i) undercoverage, (ii) overcoverage and (iii) multiple listings. Another sort of frame deficiency is misclassification, incorrect information about frame units. Such a deficiency causes errors other than coverage errors)…." Reference: http://epp.eurostat.ec.europa.eu/portal/page/portal/ver-1/quality/documents/ESQR_FINAL.pdf (page 11).*

Recommendation in case of misclassification of enterprises

There are different possible options available in case of misclassification of enterprises. The options depend on the type of misclassification.

If it is obvious that an enterprise should not have been included in the survey for example because the number of employed persons have decreased and falls beyond the size cut-off limit during the reference period, the enterprise could be excluded from the sample, as it should have never been included.

However, such an approach could be hazardous as correction then only is made for enterprises that fall beyond the cut-off limit (e.g. 10 persons employed) and not for enterprises that had less than 10 employed persons according to register information and that during the reference period of the survey exceeded the cut-off limit. A more appropriate approach is in those cases to assume

that enterprises where the number of persons employed has decreased below the cut-off limit offset the enterprises that have increased in number of employed persons and that exceed the cut-off limit. Enterprises falling below the cut-off limit are then treated as respondents and not as over-coverage.

If the misclassification means that the enterprise should belong to another stratum than the one used for stratification, new strata should be built and the weights used in computations should be changed accordingly.

I.2.7.2 NON-RESPONSE TREATMENT[19]

Introduction

An important source of non-sampling error in surveys is the effect of non-response on the survey results. Non-response can be defined as the failure to obtain complete measurements on the (eligible) survey sample. The extent of non-response varies from partial non-response (failure to answer just one or some questions) to total non-response.

The latter case occurs when the respondent refused to participate in the survey or when the questionnaire was sent back to the statistical office as returned mail (e.g. when the address to the enterprise is incorrect). This type of non-response is called unit non-response (see I.2.7.3): the sample unit does not provide any of the data required by the survey. Unit non-response is generally handled by adjusting the weight of the enterprises that responded to the survey to compensate for those that did not respond.

Partial non-response or item non-response (see I.2.7.4) occurs when the respondents did not answer all questions because they did not understand or misinterpreted a question, alternatively refused to answer a question.

Effect of non-response on the quality of the data

Non-response (unit as well as item non-response) can seriously affect the quality of the data collected in a survey. Firstly, the characteristics (or answering pattern) of the non-respondents can be different from those collected among the sample units who did provide eligible answers. If such difference is systematic, serious bias can be introduced in the survey results. Secondly, the reduction of the sample size (overall or for certain questions) will increase the variance of the

[19] References for this chapter:

- Government Statistical Service (1997), *Report of the Task Force on Imputation – June 1996*. Government Statistical Service Methodology Series No. 3. London: Office for National Statistics.
- Statistics Canada (2003), *Household internet Use Survey, 2002 – Microdata User Guide*.
- European Commission (2003), *Household Budget Surveys in the EU: Methodology and recommendations for harmonisation 2003*. Luxembourg: Office for Official Publications of the European Communities.
- National Centre for Social Research (1999), *Item non-response* (different contributions). Survey Methods Newsletter Volume 10 No.2 1999.
- Adjusting for Missing Data, Mohadjer, L., Choudhry, H., Studies of welfare populations: Data Collection and research Issues, 2001.
- Introduction to Survey Quality, Biemer, P., Lyberg, L., Wiley series in survey methodology, 2003.
- Weighting for non-response, Lynn, P., Survey and Statistical Computing, 1996.
- Non-response in the Norwegian Business Tendency Survey, Wang, J., Statistics Norway, 2004.

estimates. Thirdly, non-response can have an impact on the total cost of a survey exercise. Not only because a larger initial sample may be necessary, but also because of higher unit costs of the last few percentages of respondents (due to multiple visits, sending of reminders, repeated telephone calls). Finally, non-response can be an indicator of poor overall quality of the survey and thus create an image or confidence problem.

Minimising non-response

As prevention is always better than cure, attention should be given to avoiding non-response rather than treating non-response. The number (and timing) of reminder letters or call backs, the length of the fieldwork period, the survey technique(s) used, the length of the survey (i.e. the response burden), the use and structure of advance letters, the dissemination of previous results or the mandatory nature of the survey can all have an impact on the number of non-contacts or refusals. The length, design and complexity of the questionnaire, the interviewer's training or the interviewer's style can have an impact on the item non-response.

As this issue is common to all surveys, it will not be discussed in detail in this manual. This chapter is mainly focussed on the treatment of non-response in the specific context of the ICT usage survey.

I.2.7.3 UNIT NON-RESPONSE

Introduction

Unit non-response is defined as enterprises that are included in the sample but that have not participated in the survey and for which information consequently is missing for all the questionnaire variables.

Unit non-response can introduce bias in the survey results especially in situations in which the non-responding units are not representative of those that responded. Non-response increases both the sampling error level, by decreasing the sample size, and non-sampling error level.

Weighting adjustment for unit non-response

The principal method for unit non-response adjustment is weighting. Most strategies for weighting for non-response involve dividing the respondents into a set of comprehensive and mutually exclusive groups, referred to as weighting classes. A weight is then applied to each class.

Weighting classes

In order to implement non-response adjustments, it is required to create weighting classes. It is desirable to divide the sample in "response homogeneity groups/classes". Within these classes the response rates should be as homogeneous as possible, and the response rate should be different among the classes. Data used to form these classes must be available to both non-respondents and respondents. Usually it is possible to get information on size, economic activity, legal status, location, and other variables in the business registers.

More advanced methods for creating weighting classes are methods like classification based on a categorical search algorithm or a logistic regression model using auxiliary variables to estimate the probability of response.

Sample-Based Weighting Adjustment

In sample-based weighting adjustment the weight adjustment applied in each class, is equal to the reciprocal of the ratio of selected sample size to respondents within each of these classes (the inverse of the response rate within each class). The grossing-up factor should then be multiplied by the non-response adjustment factor.

A simple example:

Size Class	Population (I)	Sample size (II)	Respondents (III)	Respondent with characteristic (IV)	Non-response adjustment Factor (V = II / III)	Initial Grossing-up factor (VI = I / II)	Adjusted Grossing-up factor (VII = V * VI)
Small	35 141	878	764	595	1.15	40.0	46.0
Medium	5 362	882	821	795	1.07	6.1	6.5
Big	761	761	624	543	1.22	1.0	1.2
Total	41 264	2 521	2 209	1 933			

Alternative forms of sample-based weighting are that the weights are not inverse response rates, but estimated coefficients of a regression model (where survey response is the left-side variable). In this case, the weights are reciprocals of estimated response rate by the regression model.

Population-Based Weighting Adjustment

Population-based weighting adjustment requires population estimates and class membership of respondents. If there is no data available about the non-respondents, population-based adjustment still is possible since this uses external control counts for the population and not data from the sample. The method is used to correct simultaneously for both non-coverage and non-respondents. The method is used similar to the sample-based method.

In population-based adjustment (post-stratification adjustment) the classes are created based on variables, which are known both for respondents and for the population. Weights are then applied in proportion to the ratio of population to achieved sample, so that the sums of the adjusted weights are equal to population totals for certain classes of the population.

A two-step procedure of first adjusting for non-response (sample-based adjusting) and then adjusting to known population counts is a common method that is used. However, this procedure is the same as a population-based weighting adjustment if the weighting classes in the sample-based and the population-based weighting adjustment are equal.

If the strata used in the stratification are used as classes in the weighting adjustment, there is no need for the weighting adjustment. The adjusted weighting procedure is then equal to the final grossing up/weighting procedure.

I.2.7.4 ITEM NON-RESPONSE

As already mentioned above, there are several reasons for the data being unavailable. These include the refusal to provide an answer, the inability to provide an answer, inadequate quality of the provided answer (e.g. implausible, incomplete, inconsistent with answers to other questions, etc.). It can be caused by either the respondent (e.g. refusal) or the interviewer (e.g. failure to record the answer adequately) but also by the survey design itself (e.g. ambiguous routing or filtering).

In case a particular questionnaire shows too many errors, or if too many data are missing, it can be assumed that the enterprise in question has not co-operated satisfactorily in the survey. Here, the best solution is probably to remove the enterprise from the database for respondents (but not from the frame) and adjust the weighting coefficients for the other enterprises accordingly. In other words, sampling units with a very high item non-response can better be classified as total non-response or unit non-response. However, if the enterprise has provided answers on some key questions of the questionnaire, it should probably be preferable to keep the answers provided by the enterprise (even though many questions are left blank).

The general assumption that should be used is that item non-response means that the answer to the question is:

- "NO" if the question is a qualitative question (e.g. yes/no-questions)
- Zero if the question is a quantitative question (e.g. e-commerce value questions)

The general recommendation is that imputations should be avoided, except logical imputations.

In principle two kinds of imputations could be considered: 1) to use data from answers provided by the enterprise in the previous year or 2) to use data from the current survey and impute the answer on the basis of answers provided by similar enterprises (e.g. enterprises belonging to the same strata).

To use data from the previous year is not recommended as the use of ICT is expected to change over time. It is e.g. not obvious that an enterprise that did not sell goods or services via computer networks during year t-1 still did not do so during year t. The data provided the previous year can however be used in order to find enterprises that should be contacted to verify the change from previous year.

To use the data of the current survey to generate an imputed value for a variable is not recommended as there is a risk of overestimation. In this survey possible underestimation due to item non-response is preferred to possible overestimation due to imputation.

As **no imputation is recommended**, it is very important that item non-response is kept as low as possible in order to minimise possible underestimation.

There are different methods to attain lower item non-response rates:

- Clarity of questions. Unclear questions might be more frequently "avoided" by the respondents. The content and form of the questions must be considered carefully. Pre-tests can be done before sending out the questionnaires in order to check that the questions are clear to the respondents.
- After conducting the survey, item non-response can be further reduced by contacting the respondents in order to get answers to the missing questions. Special efforts should be put in order to gather answers for questions where item non-response is especially high.

For e-commerce value questions it is recommended to check for enterprises with non-response in e-commerce value questions but having large e-commerce values in previous years. These enterprises should be contacted to verify the current situation.

For e-commerce value questions it is important to compare the answer to that of previous year. Especially enterprises stating that they have e-sales but that have not provided an estimate of the level of e-sales should be contacted to have an estimate. If the enterprise has provided an estimate of e-sales value previous year it could be checked if the value that was provided then can still be of use.

Item non-response for the background variables that are used in the grossing-up process should also be avoided. Here also, contact with the enterprises is the preferred method.

I.2.7.5 COMPUTING ACCURACY MEASURES

The use of a sample survey unavoidably leads to sampling errors. To get an idea of the sampling error, it is advised to compute this error for certain indicators. Such information is not only interesting to decide whether the indicator is reliable enough to be published, but can also help to determine whether the sample size should be increased in a next edition of the survey for certain strata where a large sampling error was diagnosed.

The sampling error reflects the fact that only a particular sample was surveyed rather than the entire population. As accuracy measure the standard error (the square root of the variance) of the estimator for proportions is recommended (see also Chapter 9 of the *Reporting template*, Annex I.3.3). The estimation of the sampling variance should take into account the sampling design (e.g. the stratification).

Additional guidance for the standard errors including thresholds regarding the levels that are considered acceptable can be found in section I.2.3 where the sampling design is discussed.

For the 2008 quality reports for the first time information on standard error calculation was requested; this was done in line with the agreement between the NSI's and Eurostat to study a revision of the sampling error calculation. Member States were invited to describe the method used during the survey of that year. On the basis of the quality reports it appeared that several ways of calculation were used of which the Taylor expansion method and a model-assisted method.

This first collection of more detailed information on sampling errors shows that further work at EU level is needed. The coming Working Groups and Task Forces will address this topic.

The computed accuracy measures will also determine the reliability of the estimates. If the standard error for a certain proportion or absolute value is too high, the cell needs to be suppressed.

In case the computation of the standard error for all cells is not feasible, the number of respondents underlying the estimate can be used to decide whether a proportion can be published (e.g. if based on less than 10 respondents, the estimate is not published, unless the stratum is take-all stratum and there are no issues related to confidentiality).

On the basis of information collected in the 2008 quality reports on the way NSI's are dealing with accuracy of the estimates, several elements are worth mentioning:

- National rules used for determining which data can be released and which not due to low reliability are not always the same as the rules used when transmitting data to Eurostat; this

might lead to situations where low quality data are released by the NSI whereas they are not by Eurostat or vice versa; this practice should be eliminated.

- Some countries seem to release data before checking the reliability in terms of sampling errors of the results; although speed of publication for this domain is very important, it should not be done at the cost of quality.
- The samples are not always designed in such a way that the quality criteria in terms of standard errors are systematically respected: some countries seem to give preference to groupings of economic activities that are relevant on national level instead of to those included in the Regulation on EU level.
- Exceeding the thresholds in terms of standard errors for some data does not always make the NSI qualify these data as unreliable; this could lead to confusion for the end-user as to the quality of the data.

See also Annex I.3.2, the transmission format; this document describes how cells that do not comply with the quality standards mentioned in section I.2.3 in terms of standard errors or are for other reasons not of a sufficient quality (high item non-response, question not in line with the Model Questionnaire, etc.) should be "flagged" as unreliable; in the latter case notes accompanying the data should explain why the data is considered unreliable.

It is noted that in the Quality report (2010 and subsequent) the header of the table under point 9.1, should in the last column "Standard error" refer to percentage points i.e. "pp", and not to %. The requested information concerns variables which are proportions expressed in percentages (example: the proportion of enterprises with access to internet is 0.85) or conceptually equivalent percentages (example: the percentage of enterprises with access to internet is 85%). Although the percentages use the sign '%', the unit of measurement of these variables are percentage points (the percentage of 85% is 85 times one percentage point). The absolute precision measure of standard error keeps the measurement unit of the variable. Therefore, the precision measure of the estimated proportions/percentages should be expressed in percentage points (absolute precision measure) and not as a percentage (relative precision measure which is derived by dividing the absolute precision measure by the estimated proportion/percentage). The absolute precision measure of standard error i.e. percentage points, can be easily interpreted and used for the calculation of confidence intervals when the estimated proportions/percentages are expressed in percentage points.

Example: 73% of all the enterprises used the internet for interaction with public authorities, during 2009. The standard error could be 1.0 percentage point and hence the 95% confidence interval would be: [73 - 1.0 * 1.96, 73 + 1.0 * 1.96] = [71.04, 74.96]

Hence, the footnote number 2 of the Quality report, should read as following: (2) The variable (estimated proportion) is to be expressed as a percentage and its standard error is to be expressed in percentage points; please use decimal points only and as many decimals as needed.

I.2.8. Survey execution reports

Drawing up reports after the execution of the survey is not only a tool for a self-assessment of the work carried out by the statistical office and a detection of possible points where there is space of improvement. Survey execution reports are also a tool for assessing the comparability from one year to another or from one country to another.

The methodological reports for the ICT usage surveys coordinated by Eurostat cover the following subjects:

1. Identification of the survey and the NSI.
2. General methodological information: reference period(s); survey period; survey vehicle; survey type; survey participation; methodological differences compared to previous data collection exercises.
3. Statistical unit, scope and target population: statistical unit; economic activities covered; size coverage in terms of number of persons employed; territorial coverage.
4. Information on the national questionnaire: deviations with the Model Questionnaire; additional questions; inclusion of optional questions.
5. Sampling frame: name and description of the sampling frame or register used; indication of use of different (versions of) frames at different stages of the statistical process; known shortcomings of the sampling frame, if any; frame distribution by size class and economic activity.
6. Sampling design: sampling method; indication on coordination with sampling for other surveys; sample distribution (by size class and economic activity).
7. Unit non-response, by giving information on: gross sample size; number of out-of-scope cases, number of other ineligible cases; number of elements usable for tabulation and grossing-up; non-response; net sample size; unit response rate; additionally, methods used for minimizing unit non-response; methods used for treating unit non-response; variables or items with response rates below 90%; methods used for treating item non-response.
8. Grossing-up procedures.
9. Sampling error: standard errors for a selected group of indicators or sub-indicators; the calculation method for the standard error; comments on reliability, representativeness and completeness of the data; reliability of aggregates for the calculation of European aggregates.
10. Reference to the dissemination of national results of the survey.
11. Problems encountered and lessons to be learnt.
12. The report includes the questionnaire in national language and, if available, in English.

In cases where the recommendations provided in this Methodological Manual are not followed, the report should clearly describe deviations, the impact of those on comparability, a justification for the deviation and – where possible – plans to comply with the recommendations.

The *reporting template* used for the European enterprise ICT usage survey, including some instructions, can be found in Annex I.3.3.

This template covers both the *Metadata report* (formerly known as *Interim Report*) and the *Quality report* (formerly known as *Final Report*) referred to in Article 7(4) of Regulation 808/2004 (see Annex I.3.3) and in Annex I, section 6 of Regulation 1023/2009.

I.3. ANNEXES

I.3.1. Model Questionnaire, version 1.0

See document ***MM2014_Annex I.3.1_Model_questionnaire ENT 2014 v 1.0.doc***

I.3.2. Transmission format

See document ***MM2014_Annex I.3.2_Transmission format ENT 2014 v 2.doc*** *(for more recent version see* https://circabc.europa.eu/w/browse/02d57877-79a5-45d1-8940-2e6648176b18 *restricted access)*

I.3.3. Reporting template

For the metadata reporting, the **ESS MH** is used; the metadata flow is: ***INFOSOC_ESQRSET_A***

Part II

Household survey

II.1. STATISTICAL PRODUCT

The statistical product is the clear and precise definition of the statistical information to be produced. It distinguishes itself from the production methodology. The production methodology is the way or method of doing, while the statistical product is its direct result. Different statistical methodologies can produce the same statistical product, being only different ways of doing the same thing. This means that as long as we guarantee that two figures concern the same statistical product, for instance for two different countries, they are comparable. This way, this distinction between the statistical product and the statistical methodology helps us to focus on those elements more important to assure comparability between the several national statistics and produce new ones at the EU level, i.e. the statistical product, while leaving to the discretion of the national statistical institutes the choice of the better statistical methodology to be applied in its own country taking into consideration its own specificities.

The elements that make up the statistical product, at an input level, are the statistical unit, the target population and the observation variables, and at the output level, the periodicity and the summary measures, aggregate variables and tabulation. Covering all the elements of the statistical product, the statistical concepts and the nomenclatures are the additional needed element to assure harmonization and comparability of statistics.

II.1.1. Statistical unit

The statistical unit is the base type of the elements of a group (also called population) that we want to observe or analyse. The basic statistical operations of classification, aggregation and ordering are done on the statistical unit.

The choice of the statistical unit is a matter of both the data collection process (namely the operational restrictions associated to collecting data from each type of statistical unit) and the conceptual framework chosen to observe and analyse the phenomenon. The statistical unit is the bearer of statistical characteristics or attributes, which we ultimately want to measure.

There are several types of statistical units, according to its usage. An *observation unit* represents an identifiable entity, about which data can be obtained. During the collection of data, this is the unit for which data is recorded. It should be noted that this may, or may not be, the same as the reporting unit. The *reporting unit* is the unit that reports to the survey authority. It reports information for the

observation unit(s). In certain cases it may be different from the observation unit. A reporting unit is a unit that supplies the data for a given survey instance.

In the ICT usage survey, the following statistical units are used (depending on the variable):

- households;
- individuals.

Ideally, the data collected on the household, should be reported by 'the household'. In most cases, it is of course not feasible to put the household around the table in order to collect their common answer, this is especially the case in a telephone interview. In general, one individual in the household will answer the household related questions having the household perspective in mind. This one individual can for instance be the head of the household or the individual which has been selected for the individual questions.

As the survey relates to ones usage of computers and Internet, it is necessary that the selected individual answers the questions personally. Proxy interviews may lead to errors in the data collection.

Different survey units, i.e. households and individuals, are used in different sections of the model questionnaire. The household approach is used when information on access to different electronic devices, type of Internet connection, and barriers to use of the Internet is collected. The individual approach is used when information on use of computers, use of the Internet, e-commerce and e-skills is collected.

The reasoning behind using a household approach when trying to describe access to ICT is that households are the platform for providing access for a number of individuals, i.e. the members of households. In order to make comparisons between European countries on household level it is important to take into account the differences in household demographics. One of the reasons for this is that some countries might have bigger household sizes, and composition than other countries, which also probably is correlated with the rate of access. To be able to make fruitful comparisons it also important that countries deploy the same household definition.

The individual approach is used in the context where the collected information to a larger extent refers to the individual's use of ICT. Also in this case it is import to take into account structural differences between countries, e.g. age demographics, in order to make fruitful comparisons. The problem is however not as large as when it comes to households since many people have a better knowledge of countries' differences in age structure than the differences in household structure.

II.1.2. **Target population**

A population is a collection of objects of the same class, which in statistical terms means a group of elements of the same statistical unit. There are two types of populations to be considered when producing statistics: the target population and the frame population.

The *target population* is the population of interest. It is defined by delimiting clearly the group of statistical elements for which some information concerning the all group we want to know. That delimitation is based on one or more attributes of the statistical unit. In the ICT usage survey, the target population for the different statistical units is:

- individuals: target population consists of all individuals aged 16 to 74;
- households: target population consists of all (private) households with at least one member aged 16 to 74.

The *frame population* is an operationalisation of the target population, consisting ideally in a complete list of the elements of the target population. Although a target population can be easily defined, in practice a list of all its elements is needed for its complete or partial (in case a sample is used) observation, and that can be very difficult to obtain. That list should be complete and include every element of the target population only once. However, most of the time it will suffer from both under-coverage and over-coverage. Usually files of statistical elements (registers) are maintained and updated, containing lists of statistical elements and also information on some attributes, usually used for delimiting target populations. Frame populations are usually extracted from those registers. The frame population will be further explained in chapter II.2.2, including an overview of the countries' practices for the 2011 survey.

II.1.3. **Periodicity**

The periodicity is annual, meaning the data is collected and compiled once per year. However, some variables can be observed with a lower frequency, e.g. variables which tend to be stable over time.

An annual survey should be a compromise or trade-off between response and collection burden and the need for relevant and recent information on a "fasting moving" study domain like the information society.

This periodicity is laid down in paragraph c) of Annex II of Regulation 808/2004 and in section 5 of Regulation 859/2013 (see also *Part III – EU Legislation*).

II.1.4. **Observation variables**

In the ICT usage survey, most of the observation variables are *qualitative*, i.e. the aim is not to collect information on quantities ("how many …"), frequencies ("how often …") or amounts ("how much …") but rather to obtain non-numerical or categorical information. In most cases, the observations variables in the ICT usage survey are binary (dichotomous), meaning the respondent answers with *yes* or *no*, e.g. *Have you used the Internet for selling goods or services.* The final statistics will then mainly be proportions (e.g. the number of *yes* answers divided by the number of respondents who answered the question). Other observation variables are qualitative too, but with more than two answering categories (e.g. highest educational level).

The attributes or characteristics for the Community survey on ICT usage in households and by individuals are listed in Part III of this document (Annex II of the Commission Regulation).

The operational versions of the observation variables are the questions in the survey questionnaire. A model for such questionnaire can be found in Annex II.3.1.

The definition of the observation variables or the model questions are discussed in more detail in the explanatory notes below, see *II.1.6*. ff.

II.1.5. Summary measures, aggregated variables, indicators and tabulation

Summary measures are numerical values defined by a statistical measure that is used to summarize the values for a specific variable for all statistical units in a specific group. Such measures can take the form of aggregates (e.g. total number of *yes*-answers on a specific question) or indicators (e.g. percentage of *yes*-answers).

Aggregates can be compiled for the total population or for the different subpopulations defined by the background variables (e.g. gender or educational level) or for subpopulations defined conditional on the answers of other study variables (e.g. Internet users versus non users).

E.g.:

- number of individuals having used a computer in the last 3 months at home
- number of individuals aged 16-24 having used a computer in the last 3 months at home
- number of individuals with having used a computer in the last 3 months at home but not at work
- number of Internet users having used it for downloading official forms (= number of individuals having used the Internet and having downloaded official forms)

To obtain the indicators (proportions, percentages, etc.), the aggregates need to be divided by the total population or subpopulation.

The denominator of such ratio can for instance be:

- total population of individuals in selected strata (e.g. "proportion of persons aged 16-24 having used a computer at home" = the *number of individuals aged 16-24 having used a computer in the last 3 months at home* divided by *the total number of individuals aged 16-24*)
- total number of computer users (individuals that used a computer at least once in last 3 months)
- total number of regular computer users (individuals that used a computer at least once a week in the last 3 months)

The detailed transmission format for sending the data to Eurostat, can be found in Annex II.3.2 *Transmission format*.

II.1.6. Explanatory notes

The explanatory notes in this chapter refer to the questions in the 2014 model questionnaire (see *II.3.1 Model questionnaire*). The structure of this chapter follows the model questionnaire, i.e. the explanatory notes are grouped per *module* and per *question*. It is recommended to have the model questionnaire at hand while reading this section.

For reasons of continuity, the explanatory notes for questions that are not part of the 2014 model questionnaire (but were part of the 2013, 2012, 2011, 2010, 2009, 2008, 2007, 2006 and 2005 model questionnaires) have only been retained if necessary. More information on former questions can be found in the 2013, 2012, 2011, 2010, 2009, 2008, 2007 and 2006 manuals.

The statistical unit for Module A is the household while the individual is the statistical unit for the other modules.

II.1.6.1 MODULE A: ACCESS TO INFORMATION AND COMMUNICATION TECHNOLOGIES

➔ Statistical unit: households

A1: Do you or anyone in your household have access to the Internet at home?, (by any device)?

[Scope: all households]

[Type: one single answer needed, i.e. *Tick only one* ; binary (*Yes/No*) + *Don't know* ; filter question]

'Access' does not refer to the "connectability" (i.e. can connections be provided in the households' area or street), but to whether anyone in the household could use the Internet at home if desired, even if just to send an e-mail ("connectivity").

Former versions of the question had the clause 'regardless of whether it is used' which is now dropped. First, it will in most cases be irrelevant as households will normally make use of a service they have subscribed to (and are paying for). It is however possible that the connection was installed by one of the household members' employer or is a default facility in the building where the family is living (without actually being used). But discussions in recent ISS meetings confirmed that no significant differences in access and use could be found. Thus the clause was removed. The *Don't know* answering category should be avoided by asking the respondent additional questions. However, in some cases the household member selected for the interview may not be aware what other members are doing with e.g. the household's computer.

In certain countries, it is possible to access the Internet without subscription contracts, but for instance by ad hoc payments whenever the Internet is actually used. This can be in form of sending a text message to receive a short duration access code. In this situation, a lot of households have the *possibility* to access the Internet from home (of course under the condition they have the necessary equipment, f.i. computer+modem or Internet enabled mobile phone) but will *not be using* it.

The question should be consistent with past surveys, this means not pointing especially to mobile phone devices since it would be difficult for some respondents to reflect on whether the mobile device enables internet or not or by which connection and at which location. People using internet via a mobile phone only away from home have access at home if desired in case they subscribed to internet flat rate or could use Wi-Fi access connected to a DSL or other router in a household or public Wi-Fi if this is possible at the location of the household. In 2011, the words "by any device" were added and should serve as sufficient clarification.

A2: On which of the following devices is the Internet accessed at home?

[Scope: all households with Internet access at home, i.e. A1 = Yes]

[Type: multiple answers allowed, i.e. *Tick all that apply*]

The question will measure the diversity of devices for internet access used at home, including Smart TVs.

a) Desktop computer

b) Portable computer (laptop, netbook, tablet)

c) Other mobile devices (mobile phone or smart phone, media or games player, e-book reader)

d) TV connected to the internet (Smart TV)

The TV is used as a device to access the web and other internet services. The use of IPTV (TV broadcast received from the internet subscription) and watching web TV not be included in this item on Smart TV or connected TV. Smart TVs require integrated internet device to use web applications or over-the-top (OTT) set top boxes functioning as internet device for accessing apps, checking e-mails, browsing, chatting etc.. Households with Smart TV are also likely to connect other devices to the TV for using it as second screen.

e) Games console

f) Don't know

A3: What types of Internet connection are used at home?

[Scope: all households with Internet access at home, i.e. A1 = Yes]

[Type: multiple answers allowed, i.e. *Tick all that apply*]

The respondents should indicate which types of connections are actually used. On the basis of the listed types of connections, the grouping into broadband can be done. The question does not investigate the speed for data transport in bit/s, nor the type of device used for internet access. 'No' answers to a) or b) can be used to group into non-broadband internet connections.

Below is given an explanation of the most commonly used technologies, simplified for the 2014 by listing two broadband connections (fixed and mobile) as well as two optional items on other, narrowband connections (fixed and mobile) in order to facilitate responses. If the NSI finds it more appropriate to use the brand names or company names than using the references to the technologies (ADSL, cable, etc.), this is also a possibility (of course under the condition the brand names can be recoded to the different technologies based on a one-to-one relationship).

a) Fixed broadband connections, e.g. DSL, ADSL, VDSL, cable, optical fibre, satellite, public WiFi connections

The examples added in the wording of option a) can be explained as follows:

DSL (e.g. ADSL, SHDSL, VDSL)

DSL (Digital Subscriber Line) stands for technology that transports data at high speeds (i.e. theoretically higher than 144 kbit/s when compared to narrowband type of connections but

today normally at least around 0.5 Mbit/s) over the existing copper network. DSL technology is a type of Internet connection which is significantly faster than dial-up access and therefore designated as broadband.

Important DSL technologies include:

ADSL: This term is used for DSL where more bandwidth is allocated to download than upload (Asymmetric DSL).

SHDSL: Single pair high-speed DSL covers symmetric high speed DSL.

VDSL: Very high bit-rate DSL, with speeds up to 53 Mbps, compared to about 8 to 10 Mbps for ADSL.

Wired fixed (cable, fibre, Ethernet, PLC, etc.)

The wired fixed connections may or may not have a router for Wi-Fi access in the household, i.e. for wirelessly routing the signal to the users' device over a short distance. They are classified as wired fixed.

These technologies include:

Cable modem: This technology allows high-speed Internet access to be delivered over a cable TV network or communal aerial systems. Like DSL technologies the data is transported over the existing copper network.

LAN (Local Area Network): High-speed Internet connections via housing networks (connections common to several households, e.g. residents in a college).

Fibre optic cable: A technology which is not based on copper wire. The technology uses lasers or light emitting diodes and can provide unlimited bandwidth potential.

Ethernet: Can apply to home networks or housing networks, e.g. residents in a college, and relate to wired LAN technology, but is in general more typical for business type of internet connections to broadband. When Ethernet internet access is possible, it could be based on fibre-optic or copper twisted pair; standard Ethernet data rates could be up to 10 Gbit/s.

PLC: means power line communication. Broadband over ordinary power lines in general can range from 256 Kbit/s to 2.7 Mbit/s.

Fixed wireless (satellite, public WiFi)

Satellite technologies: Provides wireless access and are predominantly used in remote areas not reached by cable or DSL networks.

Satellite internet access usually requires installation of satellite dish/modem and monthly subscription fees. The issue applies especially for home internet access.

Bi-directional high speed satellite internet services for Internet access are possible and are marketed. Bi-directional means that receive and transmit functions are performed via the satellite. There is no need for a telephone line.

Fixed wireless technologies (FWA): Technologies in this group are all wireless technologies that provide access between fixed points (provider to the household, not within the household).

Examples are WiFi or WiMAX. Public WiFi or WiMAX may be accessible to anyone within a certain area. Inside a building different households could connect to WiFi. WiMAX connectivity provide fixed wireless access similar to Wi-Fi functionality but at much higher range (up to about 50 km).

Wi-Fi access connected to a DSL router in a household is not concerned. This should be included in item a) on DSL. Cable with a router for Wi-Fi access concerns item b).

To highlight this, in 2010 the word "public" was added before WiFi or WiMAX. However, WiMAX is not very widespread (if not existing at national level the item could be omitted in national questionnaires). For the 2012 survey, only the words public WiFi remained.

Public WiFi access can be free of charges, but sometimes not. Provision by municipalities may be within the reach of where households are located and is a main example for item c) (rather than hotspots at hotels, cafés etc.).

b) Mobile broadband connections (via mobile phone network, at least 3G, e.g. UMTS, using (SIM) card or USB key, mobile phone or smart phone as modem)

The examples added in the wording of option a) can be further explained as follows:

Only connections to 3rd generation (3G) mobile networks or higher (3+), such as for example UMTS radio technology network, or higher will classify for broadband. UMTS (Universal Mobile Telephone System) give access to the Internet at much higher speed compared to 2nd or 2,5 generation networks such as GPRS. Any access to 3G+ networks such as HSDPA (High Speed Download Packet Access) or 4G should be included.

Usual devices used for 3G, 3G+ mobile connection are mobile phones or smart phones, laptops or similar portable devices.

UMTS: Relates to 3G (third generation mobile technologies) and mobile phone or **smart phone** over broadband. UMTS broadband connectivity can also be obtained through portable computers with 3G modem.

The term GPRS (General Packet Radio Systems) stand for 2nd generation mobile technologies (though GPRS is sometimes referred to as 2,5 generation).

Mobile technology is a wireless technology. However, this should not be confused or mixed with fixed wireless technology or wired fixed as described above. As mentioned above, the possibilities that DSL or cables have WiFi routers will not classify for wireless access. For more clarity the word 'wireless' visible in earlier surveys has been skipped.

LTE (Long Term Evolution): Related to 4G (fourth generation mobile technologies). This standard was introduced in 2009 with around 100 Mbit/s download and 50 Mbit/s upload capacity, enabling much higher speed for internet use than UMTS. LTE has the ability to manage fast-moving mobiles and supports multi-cast and broadcast streams.

Mobile phone network (at least 3G, e.g. UMTS) using mobile phone or smart phone

3G mobile phones or smart phones allow use of a web-browser, e-mailing or similar internet application via 3G networks, e.g. UMTS.

Mobile phone network (at least 3G, e.g. UMTS) via a (SIM) card or USB key

Laptops or other portable computers (e.g. netbook) can normally be used with an USB key ('surf stick', dongle) or a built-in SIM card for access to 3G networks.

Other, narrowband connections - optional -

c) Dial-up access over normal telephone line or ISDN

Connects the household via dial-up either through an analogue modem or ISDN (Integrated Services Digital Network). Both types can be categorised as narrowband.

d) Mobile narrowband connection (less than 3G, e.g. 2G+/GPRS, used by mobile phone or modem in laptop)

Broadband over the mobile phone may still be limited despite recent increases or not available in many rural areas or certain regions. Second generation (2G) mobile technologies such as GPRS (General Packet Radio Service) are designated as narrowband connection. GPRS is a low capacity speed mobile system which enables Internet access. GSM, also 2G where transmission is done via circuit switched mode, is intentionally not mentioned as example. 2G cellular systems combined with GPRS are often described as 2.5G, that is, a technology between the second (2G) and third (3G) generations of mobile telephony. Enhanced Data rates for Global Evolution (EDGE) is an intermediate technology that brings second-generation GSM closer to third-generation capacity for handling data speeds up to 384 Kbits/s.

This option includes cases where a mobile phone or a laptop card only allows narrowband connection at home because broadband services are not available in certain areas.

Options on mobile connections should not be ticked if not used at home. It could be expected that people with mobile phones or smart phones having subscribed to internet flat rate are occasionally using mobile connection at home to check news etc although the main intention is normally usage away from home. People using such phones only away from home and if they find a free Wi-Fi should answer "no" to b or d. If public WiFi is possible at the location of the household, the option on fixed wireless should be ticked if actually used.

If the phone is used for Wi-Fi connection to the DSL router at home, option a) will apply. Mobile connections may also be used via a 3G modem (USB key or card) connected to a desktop, but use is expected to be marginal; if used item f) should apply.

A4: What are the reasons for not having access to the Internet at home?

[Scope: all households without Internet access at home, i.e. A2 = No]

[Type: multiple answers allowed, i.e. *Tick all that apply*]

The question replaces the former two on barriers to internet access and barriers to broadband access of the i2010 benchmarking framework. Some few modifications were introduced i.e. the former item 'don't want internet' was deleted, 'broadband not available in the area' was added and the item on security concerns moved up in the listed answer options.

The items give an idea of the household reasons or barriers for not having access to the internet at home. Respondents should indicate all reasons that are relevant to them, not only the main reason(s).

a) Have access to the Internet elsewhere

All other places, e.g. workplace, but not at the home address.

b) Don’t need Internet (because not useful, not interesting etc.)

May apply to people that do not find it useful or interesting.

c) Equipment costs too high

People in this category find equipment costs too high, that is, the previously defined devices in question A1 (including the set-up costs, e.g. the cost of a modem for internet connection).

d) Access costs too high (telephone, DSL subscription etc.)

Access costs can be telephone charges, subscription fees for broadband use etc.

e) Lack of skills

The respondent does not have sufficient knowledge to use or to get access to the Internet. The respondent may indicate that he/she perceives Internet access as too complicated. Then this item will apply since the barriers lies in a lower level of skills compared to others.

f) Privacy or security concerns

The respondent is worried about giving personal details over the Internet due to private or security reasons. Respondents might also be worried about virus attacks and similar security problems even without giving personal details over the Internet. In this respect, the spreading of spyware is a major concern (spyware is a program which has installed itself on the computer to monitor and report on the respondent’s computer activity; the development of spyware has also led to numerous anti-spyware tools to prevent spyware activity on the computer).

g) Broadband internet is not available in our area

May apply to people living in households in areas or regions where it is not possible to subscribe to broadband because of missing telecommunication and mobile network infrastructures.

h) Other

Instead of 'none of the above, but other' in the previous questionnaires, the option was changed to 'other'. This allows checking on the importance of other reasons than those mentioned above. It may also include aspects of the former item on physical disability (see e.g. 2012 survey).

II.1.6.2 MODULE B: USE OF COMPUTERS, LOCATION AND FREQUENCY OF USE

➔ Statistical unit: individuals

This module asks about the individuals' own computer use at any location (home, work, other places).

A computer is a desktop, laptop netbook or tablet computer. The following devices should not be considered as computers: Smart phones, PDA with mobile phone facilities, games consoles, media players or e-book readers.

B1: When did you last use a computer (at home, at work or any other place)?

[Scope: all individuals]

[Type: one single answer needed, i.e. *Tick only one* ; filter question]

The question addresses use of a computer by individuals of any type, i.e. desktop, laptop, netbook, tablet or handheld. A computer should be defined as multi purpose machine, a personal computer, powered by one of the major operating systems, i.e. Macintosh (Apple), Linux or Microsoft (Windows, XP or NT, Vista). PDA's (handheld computers, e.g. palmtops) without phone facilities are also be included but are hardly used anymore. [Other equipment with embedded computing abilities, e.g. cell phones or smart phones including PDA phones, VCR's, TV sets, multi-media players, etc should be excluded.]

In summary: Computer means "any type: desktop, laptop, netbook, tablet, excluding smart phone" (see also above text which was added as introduction in the model questionnaire).

Use of a computer can be at any place, not only at home or work. This has been added to the question in brackets.

B2: How often on average have you used a computer in the last 3 months?

[Scope: all individuals having used a computer in the last 3 months, i.e. 1st option in B1]

[Type: one single answer needed, i.e. *Tick only one*]

For definition of a computer see above, B1.

The three answer choices from 2013 onwards refer to daily, weekly and less than weekly usage.

Persons who use a computer at work on a daily basis during the week but who don't use a computer at home during the weekend should tick option *Every day or almost every day*.

II.1.6.3 MODULE C: USE OF THE INTERNET

➔ Statistical unit: individuals

This module asks about the individuals' own internet use at any location (home, work or other places). Any internet enabling device should be considered. Examples are desktop, laptop, netbook or tablet computers as well as smart phones, games consoles or e-book readers.

C1: When did you last use the Internet?

[Scope: all individuals (even those who reported never having used a computer in B1 as they may have accessed the Internet via other means, e.g. smart phones) **]**

[Type: one single answer needed, i.e. *Tick only one* ; filter question]

This question covers any use of the internet - whether at home, at work or from anywhere else and whether for private or professional purposes.

The time breakdowns are self-explanatory: *Within the last 3 months*; *Between 3 months and a year ago*; *More than one year ago*; *Never used it*. The question conforms to time series.

For better clarity, the explanation "via any device desktop, portable or handheld, including mobile or smart phones" has been added in brackets in this and other internet related questions in the past; in recent questionnaires, the module(s) starts with a text box explaining these issues.

C2: On average how often did you use the Internet in the last 3 months?

[Scope: individuals who used the Internet in the last 3 months, i.e. 1st option *(Within the last 3 months)* in C1]

[Type: one single answer needed, i.e. *Tick only one*]

The response categories to this question could be slightly ambiguous. Respondents should therefore be presented with all three options and should select which one best describes their behaviour.

a) Every day or almost every day

If asked this should refer to more than 4 days each week. As in question B2, persons who use the Internet at work on a daily basis during the week but who don't use it at home during the weekend should tick option *a) Every day or almost every day*.

b) At least once a week (but not every day)

If asked this should refer to between 1 and 4 days each week

c) Less than once a week

C3: Do you use any of the following mobile devices to access the Internet away from home or work?

[Scope: individuals who used the Internet in the last 3 months (i.e. 1st option in C1)]

[Type: multiple answers allowed, i.e. *Tick all that apply* or d]

This question refers to the actual *use* as a means of access not to the 'theoretical' possibility of *access*.

The words added for the 2010 survey "via wireless connection" were removed to ensure that home wireless access or at work are really excluded. Originally they aimed to signal that both types of technologies are included, the use of mobile technologies such as GPRS and UMTS which are wireless as well as fixed wireless technologies such as public WiFi or WiMAX. Replacement by the words "away from home or work" were found much more appropriate and easier to understand.

To distinguish real mobile use (in the sense of 'on the move' or nomadic use) from situations where the portable computer serves as a desktop or connects to WiFi at home or at work, the general wording "away from home or work" is supposed to clearly exclude such cases.

Examples included in mobile use away from home are use of WiFi with own devices at relatives, friends and neighbours home (considered as a form of public WiFi since the access code has been made public) or mobile internet use by employees at the place of clients.

The question applies the simplifications done for the questions on the mobile use of the internet in the specific module D of the 2012 survey (see also 2012 manual for further details).

a) Mobile phone (or smart phone)

if yes to a) a1) via mobile phone network

a2) via wireless network (e.g. WiFi)

The words "or smart phone" in brackets have been added since the 2010 questionnaire. Smart phones are associated with advanced functionalities, e.g. having a full, soft or slide-out keyboard, or advanced touch screen and internet access (e.g. iPhone). A PDA with mobile phone facility should be regarded as smart phone. The type of mobile network supported does not seem to be the primary feature; there are apparently such on the market which are not always able to link with UMTS or 3G, 3G+ networks but 2G for internet access. There is generally the possibility to use another less than 3G network if this is not available. Since 2012, a distinction between 3G and 2G network usage has been skipped in the sub-items.

Item **a1** will measure the extent to which mobile phone networks were used by individuals, no matter whether they are 2G e.g. GPRS mobile networks, 3G or above 3G.

Item **a2:** Mobile devices can be used to access the internet at hotspots with public WiFI or WIMAX at hotels, airports, public places, cafés etc..

At hotels or airports it may be necessary to buy this access for a certain time. At a cafe, the owner may provide it for free to attract customers. WiFi or WiMax may also be provided by municipalities. If there are cases where an individual uses WiFi at another person's home, this should count since the WiFi access code has become public to the individual. On the other hand, plugging out the device of another person and using a fixed connection (e.g. DSL with wireless router) at another person's home should be rare and should not be counted.

b) Portable computer (e.g. laptop, tablet) away from home or work

Item **b)** portable computers (e.g. laptop, netbook or tablet computer) via mobile or wireless connection away from home or work as important device to access the Internet remotely.

If b) was answered 'yes', the following two sub-items should be addressed:

b1) via mobile phone network, using USB key or (SIM) card or mobile phone as modem

Examples for mobile phone networks are GPRS (General Packet Radio Systems, stands for 2nd generation mobile technologies - 2G), UMTS (Universal Mobile Telephone System, 3G). HSDAP, High Speed Download Packet Access, is linked to 3.5 G networks; 4G networks are available in more and more countries/regions.

b2) via wireless network (e.g. WiFi)

For further explanations, please see above a) sub-item a2).

c) Other devices

Other devices can be used for mobile internet access:

First, any handheld computer (palmtop, PDA) with mobile telephone facility should be considered (in 2010 and before this had still a separate answer category but usage is declining). A handheld computer is a computer that can be used while you are holding it and stored in a small bag or for instance in a pocket. Personal Digital Assistants (PDAs) are another term for handheld computers. A PDA uses a pen rather than a keyboard.

As some PDAs can also be used for mobile telephony (e.g. hi-tech mobile phones with some kind of integrated PDA), then they should be regarded as smart phones, item a).

Second, any other mobile devices could be used for internet access, e.g. e-book readers, handheld games console, or MP3 players should be mentioned if used in addition to a) and/or b) or if none of the other options a) to b) apply. Tablet computers should be regarded as portable computers, item b). The iPad for example is a tablet computer.

d) I don't access the internet via any mobile device away from home or work

This option applies if none of the options a), b) or c) was answered.

C4: For which of the following activities did you use the Internet in the last 3 months for private purposes?

[Scope: individuals who used the Internet in the last 3 months (i.e. 1st option in C1)]

[Type: multiple answers allowed, i.e. *Tick all that apply*]

This set of responses refers to personal or private use only (ie. NOT for work related purposes) but access could have been from anywhere – including place of work.

The question has been modified for 2014 following the Digital Europe Benchmarking Framework 2011-2015.[20] The framework, providing the major rationale for the EU ICT usage surveys, includes a set of biennial usage indicators. Some biennial indicators mentioned in the framework are measured annually because of the high interest (e.g. internet banking). For 2014, the following have been selected for question C4.

Communication

a) Sending / receiving e-mails

It includes the use of e-mail for sending messages to friends or for getting information on goods/services.

b) Telephoning over the Internet / video calls (via webcam) over the Internet

General issues: Telephoning over the Internet is a relatively inexpensive method to communicate. Voice-over-Internet Protocol (IP) or VoIP are offered by specialised enterprises but also by internet service providers (if asked, the respondents will often not know if VoIP is used). Peer-to-peer telephony has become important. The user needs to install a little program (such as Skype) for making free calls over the internet to anyone else who also has this software. Usually, one can also make calls to normal fixed or mobile lines via a pre-paid credit.

Item p) is an activity where installment of a program or activating it for making live audio or combined with visual communication between persons (e.g. Skype) is involved.

Video calls (via webcam) over the Internet

The item was changed in 2008. The former C5b item in 2007 was not found relevant since videoconferencing is normally understood as being used by companies and for business related purposes and instead of face-to-face meetings. Often it is not based on IP network.

The term video calls reflect more current trends in live audio and visual communication between persons. It is simply referring to web camera as additional mean and via Internet (IP based networks). Web cameras, usual small digital cameras are used. The features can be made accessible in programs for Internet calls/telephoning (such as Skype).

An explanation "using applications, e.g. Skype or Facetime" could be added directly in the answer option in national questionnaires (item C4c in the 2014 model questionnaire).

c) Participating in social networks (creating user profile, posting messages or other contributions to facebook, twitter, etc.)

[instead of 'posting messages to social media' originally foreseen in the benchmarking framework for 2014; a change was requested and agreed]

Social networking can be distinguished from other communication and content activities by the aspect of creating a profile on certain websites. Besides global platforms such as facebook or MySpace there are many others focussing on certain subjects or existing offline communities.

[20] https://circabc.europa.eu/sd/d/ee295abf-562a-4d16-99c0-2540438dbdde/i2010_HLG_benchmarking_framework%20fin.pdf

The terms facebook or twitter in the wording of the answer option are meant as examples as they are widely known. National questionnaires could add any other network frequently used at the national level.

Being a member of a network with selected other members who share interests and activities is an essential characteristic of a social network. Online social networks are websites where members can store information about themselves, typically in the form of profiles, and make contact with other members, thus creating a network of personal connections. In addition, user-created content can be uploaded by anyone, such as texts, photos, music files and video clips, which often act as the centre for interaction within the network (e.g. YouTube, MySpace). Social networking may take place anonymous or with real personal data. Information provided often includes photographs and CVs. Members' networks of connections are usually displayed as an integral piece of their self-presentation.

Social networking sites thus include tools for posting personal data into a profile, uploading user-created content, allowing personalised interaction and communication with others by posting messages, and defining social relationships by determining who has access to data, who can communicate with whom and how.

Many different social networks exist that cater to different needs. Examples are (source: http://ec.europa.eu/information_society/activities/social_networking/facts/index_en.htm):

Generalist social networks: MySpace, Facebook, Skyrock, Bebo, Netlog, Hyves, StudiVZ.de, Piczo, Zap.lu, MSN, Giovani.it, Arto.dk, Yahoo, One.lt, Grono, Tuenti, Aha.bg

Content-based platforms, where users can watch or upload content such as videos or pictures: Youtube, Dailymotion, Flickr

Micro-blogging networks: Twitter

Virtual environments: Second Life, Habbo Hotel

Access to information

d) Reading online news / newspapers / news magazines

This should include all types of online newspapers and magazines either free of charge or under payment.

The wording "Reading or downloading…" in former questionnaires was modified and "downloading" removed since information on downloading (or direct screening) was found irrelevant.

e) Finding information about goods or services

Using the Internet to seek for information about any household good, for example, films, music, video-games, books, e-learning material, clothes, electronic equipment computer software or services for example banking, financial or health services. It should not include transactions, e.g. purchases of any goods or services, although one will usually look up information on a good or services before actually purchasing it.

(In the former 2009 survey, two optional sub-items d1 and d2 have been added to investigate whether looking up information on products on websites have actually influenced shopping or

purchases, and e-shopping and purchases over the Internet – these were not continued in 2010, 2011).

Use of entertainment

f) Listening to web radios

This covers both live streaming (real-time) and radio 'on demand' (batch, i.e. the user can listen programs later on). This does not result in any saving of information to, for example, the respondents computer. [revised: in 2012, C3f read listening to web radios or watching web TV]

g) Playing or downloading games, images, films or music

This refers to actually playing or downloading games, images for use, films, movies or videos for direct watching (but not live streaming) or downloading, or music for listening (but not live streaming) or downloading.

h) This applies if item g) has been ticked:

Playing networked games with other persons

Creativity

i) Uploading self-created content (text, images, photos, videos, music etc) to any website to be shared

This can involve uploading of own produced content to own website or to any other website with the purpose of sharing it with others. Websites enabling sharing files with others are e.g. social networking sites, file sharing services (e.g. Picasa and other cloud services, e.g. Dropbox).

j) Creating websites or blogs

This refers to practical experience in doing all necessary steps for setting up a website or blog by the respondent him-/herself (not done by others for the respondent), including getting an URL (registering a domain name), signing up for website hosting, and getting the finished site into search engines. It can involve nowadays the instalment and use of site setting up software requiring fewer skills on technical design than a few years ago.

e-health

k) Making an appointment with a practitioner via a website (e.g. of a hospital or a health care centre)

Sending an e-mail is excluded. The e-health related item will check to which extent the population makes use of websites of hospitals and health care centres providing opportunities for online appointments via web forms. These services may not be applicable in a few countries (if validated by other reliable sources, it can be skipped in the national questionnaire).

Other on-line services

l) Using services related to travel and accommodation

Includes using the Internet for ascertaining information or to purchase goods and services in relation to travel and travel related/holiday accommodation, for example travel tickets, hotels or any other type of accommodation or web sites containing information for tourists.

m) Selling goods or services, e.g. via auctions (e.g. eBay)

Selling goods or services on-line, e.g. via eBay website, does not require an *electronic* payment transaction, i.e. the transaction or 'deal' is done on-line but the payment and/or delivery can take place off-line.

Putting an advertisement on a website to, for example, sell a second-hand bicycle or a spare ticket for an event, should not be included here as the transaction is in general not concluded on-line in an automatic manner (but via a phone call or informal e-mail).

n) Internet Banking

This includes electronic transactions with a bank for payment, transfers, etc. or for looking up account information.

Electronic transactions for other types of financial services are not covered by this category (purchase over the Internet of shares, financial services, insurances is covered by item h*)* of question F2, see Module F. A simple information search on e.g. shares or financial services is of course included under item e*) Finding information on goods and services*.

II.1.6.4 MODULE D: USE OF CLOUD SERVICES

➔ Statistical unit: individuals

The following questions mainly concern the use of Internet services for storing files for private purposes in the last 3 months prior to the interview.

These services offer the possibility to store on a server accessible over the internet. Many of them are used for sharing files with others. In particular media files are usually large and storing them on the internet will provide opportunities to overcome limitations in connection with own devices used for storing.

Cloud computing simply means that individuals move from using their own resources (hardware, software, human resources) to using more resources of a cloud computing provider as a service via the Internet. According to the US National Institute of Standards and Technology cloud computing technology is defined as follows:

Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g. networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model promotes availability and is composed of five essential characteristics (on-demand self-service, ubiquitous network access, metered use, elasticity, resource pooling), three service models

(software, platform and infrastructure as services) and four deployment models (hybrid clouds, private cloud, community cloud and public cloud).

The cloud characteristics most popular and relevant for individuals and deserving special attention in the survey refer to the possibility to store data on a server of cloud operator offering this infrastructure related service to the general public (public cloud). Such data and media files (including office documents, photos, videos, music) can also be shared or communicated with others. The individual may store his/her own files or those for which he has usage/sharing rights. Moreover, software applications can be run in the cloud to edit files (e.g. word processing) or to play or edit media files.

The measurement focus of 'Cloud computing' should be the storage of files and software on remote servers offered by dedicated services. How to differentiate such new services from already popular services, e.g. Facebook or web-based email accounts that use cloud computing technologies, posed a major challenge for the design of the module in the questionnaire.

D1: Did you use storage space on the Internet to save documents, pictures, music, video or other files, e.g. Google Drive, Dropbox, Windows OneDrive (formerly Skydrive), iCloud, Amazon Cloud Drive?

[Scope: individuals who used the Internet in the last 3 months (i.e. 1st option in C1)]

[Type: one single answer needed, i.e. *Tick only one* ; binary (*Yes/No*)]

The question does not ask about the use of personal websites or social media (e.g. blogs, social networking sites such as Facebook) or the use of e-mail accounts, but asks about the use of dedicated services. It asks about active usage of online/cloud storage applications and not about having access.

For a better understanding regarding cloud services providing storage space on the Internet, most popular services have been added directly in the question. Further popular national examples could be mentioned as necessary.

For further information:

Google Drive: http://learn.googleapps.com/tour-drive

Dropbox: https://www.dropbox.com/tour/1

Windows Onedrive: https://onedrive.live.com/about/en-gb/

(formerly Windows Skydrive: http://windows.microsoft.com/en-us/skydrive/download)

iCloud: http://www.apple.com/icloud/

Amazon Cloud Drive:
https://www.amazon.com/gp/feature.html/ref=amb_link_396402442_3?ie=UTF8&pf_rd_m=ATVPDKIKX0DER&pf_rd_s=center-8&pf_rd_r=0NDVGHS4CK78JY0HSJDN&pf_rd_p=1681451922&pf_rd_t=6301&docId=1000796931&pf_rd_i=clouddrive_authportal_us

D2: When sharing documents, pictures or other files electronically, which of the following did you use?

[Scope: individuals who used the Internet in the last 3 months (i.e. 1st option in C1)]

[Type: multiple answers allowed, i.e. *Tick all that apply – or option e) in case no file sharing services were used at all ; filter question*]

The question lists tools and means for sharing files with others, online and offline, including e-mail services and personal and social networking sites, e.g. Facebook. Because e-mail accounts and personal and social networking sites should not be considered as cloud services, the question serves as a filter question. Option c) respondents and those who use storage space for file saving ('yes' in D1) are cloud service users and routed to the next question D3.

a) E-mails with attached files

b) Personal websites (e.g. blogs) or social networking sites (e.g. Facebook)

c) Storage space on the Internet, e.g. Dropbox, Windows ***OneDrive (formerly Skydrive),*** Google Drive or file sharing services, e.g. Picasa, Flickr

Internet storage space often provides opportunities for sharing files as mentioned in the examples of question D1. Item c) checks the use of file sharing provided by cloud services and popular examples are again mentioned directly. There are also dedicated file sharing services (Picasa, Flickr) which use internet storage space.

Applications like Picasa and Flickr allow the uploading and sharing of photos and videos. Content can be made public or visible to groups in both applications. Picasa (Google) is mainly used to organise, edit and share photos. Flickr is more a borderline case. It also provides photo file storing and sharing services and in addition allows networking, often used among photographers' groups, with facilities to comment and discuss. Since they are typical file sharing services using internet storage space in the first place, the examples should remain in c) (and not in b including social networking).

d) Other means not using internet, e.g. USB stick, DVD, Bluetooth

e) None, did not share files

This item is important to ensure that non-response can be measured properly. Deriving a variable for non-users from "no" answers to options a)-d) was not found appropriate for all survey types.

Since cloud services for file sharing are services for file saving using internet storage space at the same time, there should be no cases where someone answers "no" to D1 and "yes" to D2c. The routing for "no" in D1 could theoretically directly lead to question D7 on awareness. However, respondents may not always think of saving when using web storage applications for sharing. Eurostat will derive an aggregated variable with the scope of D1 “yes” or D2c “yes” for review and publication.

D3: Did you use Internet storage space to save or share the following?

[Scope: individuals who used the Internet in the last 3 months (i.e. 1st option in C1) and who used cloud services (i.e. 'yes' in D1 or 'yes' to option c in D2)]

[Type: multiple answers allowed, i.e. *Tick all that apply*]

The question aims to obtain more information on user behaviour and the type of files used on internet storage space for file saving, sharing, reading or playing. It provides further insights about the type of files used in web storage applications. Or music or video files were purchased in webiTunes is a web shop for music, films etc.. With the tool iTunes match, the content can be saved in iCloud for access and listening or watching anywhere and iCloud is included as example (question D1). Content purchased from web stores may be made available without additional cost or with payment (depending on size of use) in web storage services provided by the same or other businesses.

The question was designed because it could provide in a simple way further insights about the type of files used in web storage applications

a) Texts, spreadsheets or electronic presentations

This concerns files and documents created with word processing and other software (e.g. excel, powerpoint).

The files were usually created with software installed on the own computer and then uploaded in Internet storage space or with software available on internet storage space, saved there and eventually shared with others.

b) Photos

Photo files were usually transferred from other devices (camera, smart phone), then uploaded in Internet storage space, saved there and eventually shared with others.

c) E-books or e-magazines

E-books or e-magazines purchased from web stores may be made available in internet storage (cloud) services. For example, electronic version of books purchased from the Amazon/Kindle web shop can be delivered and automatically downloaded to the Kindle Cloud Reader.

d) Music

Music files in mp3 or other format can be uploaded in Internet storage space. It also involves internet storage space that is specifically designed for storing digital audio files, including so-called music locker services, such as e.g. iTunes Match, Amazon Cloud Player, Google Play. Web services such as iTunes providing files for streaming or downloading are not dedicated cloud services.

e) Videos (including films, TV programmes)

Video files can be uploaded in Internet storage space. Video content purchased from web stores may be made available without additional cost or with payment (depending on size of use) in web storage services provided by the same or other businesses. For example, iTunes is a web shop for music, films etc.. With the tool iTunes match, the content can be saved in iCloud for access and listening or watching anywhere and iCloud is included as example in question D1. TV programme files can be downloaded from provider's websites, e.g. as MPEG-4 Movie file type and then saved in a dedicated cloud service.

f) Other

Includes any other types of files not mentioned above. For example: private e-mail files may also be saved in dedicated internet storage space, although this practice is expected to be less frequent.

D4: Have you paid for Internet storage space or file sharing services?

[Scope: individuals who used the Internet in the last 3 months (i.e. 1st option in C1) and who used cloud services (i.e. 'yes' in D1 or 'yes' to option c in D2)]

[Type: one single answer needed, i.e. *Tick only one* ; binary (*Yes/No*)]

Some services providing internet storage space are for free, some not. It usually depends on the size of memory space used for file storing. Internet storage space is evident in all cloud services enabling saving, sharing, creating, editing and playing of files. However, the words "…or file sharing services" have been added for reasons of possible non-awareness by respondents.

Payment can be done over the internet (e.g. providing credit or debit card or pre-paid account details, transfer via internet banking) or not via the internet (e.g. ordinary bank transfer).

D5: What are the reasons for using Internet storage space to save or share files?

[Scope: individuals who used the Internet in the last 3 months (i.e. 1st option in C1) and who used cloud services (i.e. 'yes' in D1 or 'yes' to option c in D2)]

[Type: multiple answers allowed, i.e. *Tick all that apply*]

The respondents should state their own actual reason(s) and own motivation that has led to using internet storage space. The following answer options appear rather clear. On better information security and protection against data loss, item c), respondents may think of advantages of using the cloud service as file back-up service.

a) Possibility to use files from several devices or locations

b) Possibility to use larger memory space

c) Protection against data loss

d) Possibility to share files with other persons easily

e) Access to large libraries of music, TV programmes or films

D6: When using storage space on the Internet or file sharing services, have you ever experienced any of the following problems?

[Scope: individuals who used the Internet in the last 3 months (i.e. 1st option in C1) and who used cloud services (i.e. 'yes' in D1 or 'yes' to option c in D2)]

[Type: multiple answers allowed, i.e. *Tick all that apply – or option h) in case of no problems*]

The following answer options are expected to provide a rather complete list of possible problems that may be faced when using cloud services. These problems for the relatively new cloud phenomena may have appeared at any time in the past; a reference period of the last three months has been omitted by intention.

a) Slow speed of access or use

Depending on the type of internet connection and the size of files, access or usage may involve a large amount of time.

b) Incompatibility between different devices or file formats

For example, documents in internet storage space that cannot be read in MS Word format.

c) Technical server problems, e.g. unavailability of the service

This resulted in no access to the service and uploaded data on the internet.

d) Disclosure of data to third parties due to security problems or breach

Cloud users often may not be aware if there was any disclosure of data and security breach. But if experienced, the answer option applies.

e) Unauthorised use of personal information by service provider

Respondents often may not be aware of any unauthorised use of personal information. But if experienced, the answer option applies.

f) Terms and conditions of service provider ambiguous or difficult to understand

For example, a certain limitation in storage space that was not visible in the service conditions.

g) Difficulties moving files from one service provider to another

There could be restrictions for storage of any files, in particular concerning media files purchased in web shops and stored in internet space provided by the same web shop/seller.

h) No problems were faced

Item h), will allow a variable on the percentage of individuals who experienced no problems when using internet storage space. Non-response can be avoided in paper questionnaires through this item.

Item h) should be answered in case no problems were faced at all. However, since there remains some uncertainty regarding other problems than a)-g) after having adopted the questionnaire, an item "other" is advised to be added in the national questionnaires for verification.

D7: Were you aware of the existence of services providing storage space on the Internet?

[Scope: individuals who used the Internet in the last 3 months (i.e. 1st option in C1) and who did not use cloud services (i.e. 'no' in D1 or 'no' to option c in D2)]

[Type: one single answer needed, i.e. *Tick only one* ; binary (*Yes/No*), filter question for D8]

The question is for non-users of cloud services, i.e. those who did not use internet storage space for saving or sharing of files.

The reason for non-use of cloud services is expected to be in most cases a lack of awareness, thus a filter question to collect evidence. Those individuals who know about such services but did not use them should then be directed to question D8, checking different reasons for non-usage.

D8: What are the reasons for not using services providing storage space on the Internet?

[Scope: individuals who used the Internet in the last 3 months (i.e. 1st option in C1) and who did not use cloud services (i.e. 'no' in D1 or 'no' to option c in D2)]

[Type: multiple answers allowed, i.e. *Tick all that apply*]

Besides using other solutions for file saving and sharing, respondents may avoid usage because of lack of knowledge, security or trust concerns. The respondents should state their own reason(s) and own motivation that has led to the non-use of internet storage space.

a) Save files on own devices, e-mail account or rarely/never save files

b) Share files using other ways (e-mail, social media, USB stick) or do not share files online with others at all

c) Lack of knowledge about how to use such storage space

This applies if the user did not know how to use the service and did not try or tried but was not able to perform the activity. The potential results could reveal needs for improving skills.

d) Concerns about security or privacy

This refers to e.g. worries about misuse of private files or files with sensitive data saved on internet storage space.

e) Concerns about reliability of service providers

Reliability refers to delivery of cloud platforms and services that are accessible all the time, secure and private. The item will check worries about reliability expectations.

D9: Have you used software run over the Internet for editing text documents, spreadsheets or presentations, e.g. Google Docs, Office 365?

[Scope: individuals who used the Internet in the last 3 months (i.e. 1st option in C1)]

[Type: one single answer needed, i.e. *Tick only one* ; binary (*Yes/No*)]

Google Docs, Office 365, Google Sheets and Google Slides are examples for online word processing, spreadsheet and presentation software. These web based applications and services for (collaborative) editing are linked or accessible via the provider's cloud storage services. All individuals who used the internet, not only cloud users who answered yes in D1 and yes to D2c should be routed to this question because respondents may not be aware of the file storage/saving features of the online service.

D10: Have you used software run over the Internet for editing pictures or videos, e.g. Picasa, WeVideo, JW Player?

[Scope: individuals who used the Internet in the last 3 months (i.e. 1st option in C1) – *optional question*]

[Type: one single answer needed, i.e. *Tick only one* ; binary (*Yes/No*)]

As for editing of documents etc (question D9) storage space can be bundled with software run over the internet for editing of pictures or videos. The examples Picasa, WeVideo and JW Player are popular and have been inserted directly in the question for a better understanding by respondents.

For further information:

http://picasa.google.com/

http://www.wevideo.com/

http://www.jwplayer.com/cloud-video-player/

D11: Have you used services over the Internet for playing music or video files uploaded or saved in internet storage space, e.g. Amazon Cloud Player, iTunes in the Cloud (iCloud), Google Play Music?

[Scope: individuals who used the Internet in the last 3 months (i.e. 1st option in C1) – *optional question*]

[Type: one single answer needed, i.e. *Tick only one* ; binary (*Yes/No*)]

The question will measure cloud services going beyond storage and providing applications for listening to music or watching films (and eventual sharing).

The examples Amazon Cloud Player, iTunes in the Cloud (via iCloud; iTunes is not to be understood as a cloud service) and Google Play Music are popular and have been inserted directly in the question for a better understanding by respondents.

For further information:

http://www.amazon.com/gp/dmusic/marketing/CloudPlayerLaunchPage/ref=sv_dmusic_0

http://www.apple.com/support/itunes/in-the-cloud/

https://play.google.com/about/music/unlock/

II.1.6.5 MODULE E: USE OF E-GOVERNMENT

➔ Statistical unit: individuals

This module asks about electronic contacts vie the Internet with public authorities and some public services. Contacts through manually typed e-mails should be excluded.

Contact and interaction with public authorities and public services include websites concerning citizen obligations (e.g. tax declaration, notification of moving), rights (e.g. social benefits), official documents (e.g. ID card, birth certificate), public educational services (e.g. public libraries, information on the enrolment in schools or universities), public health services (e.g. services of public hospitals).

E1: 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?

[Scope: individuals who used the Internet in the last 12 months (i.e. 1st and 2nd options in C1)]

[Type: multiple answers allowed, i.e. *Tick all that apply*]

Definition of public authorities or public services used in the survey

Contact or interaction with public authorities or public services over the internet is done normally via websites provided by these entities. The wording in former questionnaires changed to read "contact or interact" and "public authorities and public services". Limiting the questions only to public authorities was found too narrow. On the other hand, not all public services should receive attention. This also complies with former descriptions in the manual and is now made more explicit in the wording.

Regarding adding the word "contact", the reason was mainly because of translation problems for the word "interact" in some countries.

Contact or interaction with public authorities or public services within the scope of the e-Government module has been explained directly in the first question of the module and is defined as follows:

Contact or interaction with "websites concerning citizen obligations (e.g. tax declaration, notification of moving), rights (e.g. social benefits), official documents (e.g. ID card, birth certificate), public educational services (public libraries, information on the enrolment in public schools or university), public health services (includes services of public hospitals), "excludes manually typed e-mails.

Excluded are public postal and public transport services, public utilities (e.g. electricity), public news and weather forecast services. The reference period will be "in the last 12 months" to consider the more irregular nature of interactions with public authorities/services.

Enrolment in higher education could also be done by parents for their children. Therefore, not only students should be concerned.

As for all internet related questions, internet activities including e-government may be performed on any such activities enabling device, i.e. "via any device, desktop, portable or handheld device including mobile or smart phones".

Contact or interaction with public services or administrations

The three items a), b) and c) make use of the conceptual framework for the levels of interaction between enterprises or citizens and public authorities, used by DG CONNECT. The framework of four levels of interaction was developed to measure the level of sophistication offered by websites supplying basic public services. That way, one can distinguish websites that only provide information on a public service (1st level) from websites that allow to download forms (2nd level) or to submit information in an online form (3rd level) up to websites that allow to completely treat the public service or administrative obligation without the need for any paperwork whatsoever (4th level of sophistication). For the time being, only the first three levels are considered in the household survey.

It is expected or logical that a particular website that reaches a certain level of sophistication also include all lower levels, for instance a website offering the possibility to download a form should certainly provide information on the particular public service that form would refer to. The concept of the four levels of sophistication was adopted for the measurement of the usage of e-government by enterprises (and the 3 first levels for citizens). However, the interpretation of the results will have to differ slightly from the interpretation done for the sophistication of the public websites. At usage level, it is natural to assume that a user (enterprise or citizen), before knowing which form to download (2nd level), will obtain information on the public service he is interested in (1st level), but he/she will only do it once or twice and not every time it downloads the form. Therefore, when measuring the usage every year, it is perfectly possible that there are users who have downloaded forms with obtaining information in those cases in which the user already had all the information he needed (although not very probable, as users may well constantly search for information on several different issues). In this sense, it is possible that the respondents answers *Yes* to item *b* without having answered *Yes* to *a* (or *Yes* to *c* without having answered *Yes* to *a* and/or c).

a) Obtaining information from web sites

Includes searching to obtain any type of information from public authority web sites. Public authorities' web sites include local or central government offering information or services. It would be helpful to provide a list of local/central government web sites in the format of a showcard to facilitate respondents to recall public authority web sites that they have used.

Searching any type of information from public authority web sites may constitute some overlap with former items in the questionnaire (looking for online job offers on labour office maintained website, searching for public health organisation's information, etc.), see also manual of 2006 survey. As the competence of the government or the 'public authorities' differs strongly from one country to another, the concept has been kept as wide as possible: public services and administration. While 'administration' refers to obligation and rights one has as a citizen (the so-called red tape), public services can refer to non-administrative tasks or competences of government bodies, e.g. offering the public library's catalogue on-line.

The service providers can be at national or regional level, but also at the level of cities or municipalities. The service providers can even be 'semi-governmental', e.g. public libraries, hospitals, universities, etc.

b) Downloading official forms

Includes downloading official forms from public authorities' websites for any purpose of use (e.g. for information or for requesting a service). The downloaded files could be in formats such as PDF or Microsoft Word.

c) Submitting completed forms

Instead of "sending filled in forms" in former questions, the wording changed to "submitting completed forms". This includes completed forms sent via internet (to public authorities' websites) only. Forms downloaded, printed, filled in and sent by post should not be included in this category.

The downloading or sending or transmitting completed forms may involve the need to "sign in" at the website and provide personal details.

There could be software offered by third parties which is authorised or acknowledged by public authorities and which could be used for the transmission of forms via the internet to public authorities.

Income tax declaration required to be submitted by an individual is an example, but: If the individual gave order to a company or service and the service sent the tax declaration form, this should not count. If the declaration was required to be sent for a couple and the other person (husband, wife or partner) sent this declaration, this should also not count.

Other examples are: completing web forms for notifying new address, requesting school certificate, reserving book in public library, providing requested information (e.g. from national statistical offices).

E2: What were the reasons for not submitting completed forms to public authorities' websites for private purposes in the last 12 months?

[Scope: individuals who used the Internet in the last 12 months (i.e. 1st and 2nd options in C1) and individuals who answered "no" to option c in E1)]

[Type: multiple answers allowed, i.e. *Tick all that apply*]

The question enables to obtain a denominator of people having had a need or no need to submit forms. *Example: Results suggest that 50% of those that have the need to fill a form did it at least once electronically when the other 50% did it by paper by various reasons - e.g. because they preferred to do it or because use of such e-service was not available or it was available but the use was too difficult etc.*

Compared to 2013, question D6 was embedded in a larger specific module on e-government. For 2014 the question was simplified, listing only the most important answer options. It should be noted that this may have an impact on comparability over time when using the option "other reason".

a) Did not have to submit official forms at all

If "yes" to a), go to question F1, if "NO", go to b)-h) of question E2

There could be a note added in the national questionnaire if certain situations are known. E.g. like: adult persons concerned by a statistical census and having the possibility to submit forms online, item a) should not be ticked.

b) There was no such website service available

Either no website was available or a website existed but it did not offer the possibility to submit official forms.

c) Lack of skills or knowledge (e.g. did not know how to use website or use was too complicated)

This category applies if the user did not know how to use the service and did not try or tried but was not able to perform the activity. The potential results could reveal needs for improving skills of the population.

d) Concerns about protection and security of personal data

This refers to e.g. worries about giving personal details over the Internet. Names, addresses, but also private financial information are examples

e) Lack of or problems with electronic signature or electronic ID/certificate (required for authentication /using the service)***- optional -***

In some countries there are obligations of having an electronic signature or electronic certificate in order to be able to submit online forms, but there may be a lack of these opportunities or problems with such a service.

f) Another person did it on my behalf (e.g. consultant, tax advisor, relative or family member)

The reason for those having had a need but for having not submitted tax forms for example could be that other persons did it on behalf of the respondent. One should also consider that for example subscription to higher than the current level of education, a parent may be required to do this for the child younger than 18 years or relatives submit social benefit claims on behalf of a disabled parent.

g) Other reason

II.1.6.6 MODULE F: USE OF E-COMMERCE

The following questions concern the purchase of goods and services over the Internet (e-commerce, e-shopping) for private use via any device (desktop, portable or handheld, including mobile or smart phone).

Purchases refer to ordering goods or services for which payment is required, but the payment does not have to be done online. Orders via manually typed e-mails, SMS or MMS should be excluded.

F1: When did you last buy or order goods or services for private use over the Internet (excluding manually typed e-mails, SMS, MMS)?

[Scope: individuals who already used Internet, i.e. 1st, 2nd or 3rd option in C1]

[Type: one single answer needed, i.e. *Tick only one* ; filter question]

As for all internet related questions, internet activities including e-commerce may be performed on any such activities enabling device: desktop, portable or handheld device including mobile or smart phones.

Purchases via SMS or MMS should be excluded since they do not make use of the internet via web browsers or web apps but mobile telephone networks. E-mail is internet related, but should be excluded.

Recurrent payment activities (e.g. for transport tickes) for which no ordering process is involved should not be counted.

Purchases of financial investments, such as shares, should be included in this question. The ordering of goods and services should also include confirmed reservation for accommodation, participation in lotteries and betting and obtaining information services from the Internet that are directly paid for.

This question also applies to purchases that are made via Internet auctions, such as "eBay".

Only individuals that actually placed the order over the Internet should answer this question, even if the order was carried out on somebody else's behalf. Therefore, individuals that had other people ordering for them should not be included. For the time of the transaction, the date when the goods or services were ordered is relevant, not the date of delivery or payment.

Orders via manually written e-mails should – in line with past definitions, including the updated definition of OECD - not be included. The placing of orders is done on websites of sellers or webshops.

Delivery via electronic means or payment via electronic means is not a requirement for an order via Internet (as reconfirmed by the recent OECD definition on e-commerce).

Goods and services that were obtained via the Internet for free should be excluded. Such goods are e.g. free software ("freeware"), reservations in restaurants or any kind of information obtained via Internet for free (such as downloading pdf-files). Software that can be downloaded from the Internet and used for free for a certain period but loses functionality after this time should only be counted as a purchase when the product is finally paid for.

Orders that are placed for professional use should be excluded.

F2: What types of goods or services did you buy or order over the Internet for private use in the last 12 months?

[Scope: individuals who already used Internet (1, 2 or 3 in C1) *and* who bought/ordered over the Internet in the last 12 months, i.e. 1st (Within the last 3 months) **or 2nd option** (Between 3 months and a year ago) **in F1]**

[Type: multiple answers allowed, i.e. *Tick all that apply*]

a) Foods or Groceries

Groceries include necessities of "daily use" such as sanitary products, tobacco products, cosmetics or flowers. This category also includes drinks. Newspapers should be excluded from this category.

Pharmaceuticals were mentioned as another example in previous manuals, but there is a separate item c) on medicine. Pharmaceuticals, in particular those for which no prescription is required could be part of the product basket in some grocery web shops. Because of item c) mentioning them is no longer advised.

b) Household goods (e.g. furniture, toys, etc)

Includes e.g. furniture, toys, washing machines, micro-wave ovens, dish washers, vehicles, gardening equipment, plants, antiques, art, tools, collectibles, etc.

Household goods exclude consumer electronics such as the products addressed in items j) computer hardware and k) electronic equipment, incl. cameras.

c) Medicine

Other used terms for medicine are medicaments or medication, drugs. These can include both prescribed and non-prescribed medicaments, including those containing antibiotics and other substances for which prescription is usually necessary, as well as vitamins.

d) Clothes, sports goods

Includes shoes, textiles, accessories.

e) Computer hardware

Includes any accessories to computer hardware, such as modem and printer.

f) Electronic equipment (incl. cameras)

Includes e.g. mobile phones, cameras, radios, TVs, stereos, DVD players, video recorders.

g) Telecommunication services (e.g. TV, broadband subscriptions, fixed line or mobile phone subscriptions, uploading money on prepaid phone cards, etc)

Note that this item is about ordering/subscribing. Just paying telecommunication bills is not an e-commerce activity.

h) Share purchases, insurance policies and other financial services

Includes the purchase of securities of all kinds. Transfers from account to account and similar services via Internet where no purchase is carried out are not included (account management is addressed in C5m on Internet banking). Since financial services e-commerce include purchasing of shares, ordering of a loan and insurances, the item has been reformulated (before: share purchases, financial services, insurance).

i) Holiday accommodation

Includes e.g. hotel reservation. Only bookings of accommodation should be included. The mere search for those services and booking inquiries without legal binding should be excluded.

j) Other travel arrangements (transport tickets, car hire, etc)

Includes e.g. travel tickets and vehicle hire.

Only bookings of travel tickets and other travel services should be included. The mere search for those services and booking inquiries without legal binding should be excluded.

k) Tickets for events

Includes tickets for concerts, cinema, sport events, theatre, and alike.

l) Films, music

Includes DVDs, CDs, Video and music cassettes and similar devices that can store films, music or spoken word. Delivery of films and music on-line via files is also included. Only goods that are paid for should be mentioned here.

m) Books, magazines, newspapers (including e-books)

Subscription to newsletters and alike – if not free of charge – also belong to this category.

n) e-learning material

E-Learning materials are documents which are used for multi-media learning in digital form, but also comprise CD-Roms that are used for educational purposes.

o) Video games software, other computer software and software upgrades

Includes paid video and computer games and paid computer software on any media or devices.

p) Other

Goods and services not mentioned above. Could include e.g. jewellery, fan articles, and a number of services ordered via Internet, e.g. information services (from databases), subscription to paid cloud service.

F3: 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.?

[Scope: individuals who already used Internet (1, 2 or 3 in C1) *and* who bought/ordered over the Internet in the last 12 months and answered 'yes' to l, m, n, o in F2]

[Type: multiple answers allowed, i.e. *Tick all that apply*]

This question aims to get information on how the respondent got the products that can be downloaded or accessed directly from the websites rather than delivered at home by post.

Digitally delivered products include those products which can be stored on physical media (such as CD or tape) or printed but are actually delivered over the Internet in digitised form as a computer file (and thus independently of those physical media). Such content products can be screened or downloaded and stored on own devices. Depending on sales conditions, they may be shared with others or not allowed to share.

a) Films, music

Delivery of films /music on-line via files (e.g. a video file or an MP3 file).

b) (Electronic) books, magazines, newspapers, e-learning materials

c) Computer software (incl. computer and video games and software upgrades)

F4: From whom did you buy or order goods or services for private purposes over the Internet in the last 12 months?

[Scope: individuals who already used Internet (1, 2 or 3 in C1) *and* who bought/ordered over the Internet in the last 12 months, i.e. 1st (Within the last 3 months) **or 2nd option** (Between 3 months and a year ago) **in F1]**

[Type: multiple answers allowed, i.e. *Tick all that apply*]

Buying goods and services from other Member States contributes to the EU policy objectives of achieving a single European market and a single information space. The question is asking about the country of origin/residence of the seller.

a) National sellers

b) Sellers from other EU countries

c) Sellers from the rest of the world

d) Country of origin of sellers is not known

Multi-national companies should be treated as national sellers when it is known from the website that they are registered as a company with an address in the surveyed country. The term "national sellers" includes the trade business or sales offices established in the country by foreign owners (development, production, other distribution may be located in the home country and/or globally). This approach is more feasible and relevant at the European level than asking about websites in national or foreign language.

If the seller and his sales office address are not known, item d) should be ticked.

F5: How did you pay for goods or services ordered over the Internet for private use in the last 12 months?

[Scope: individuals who already used Internet (1, 2 or 3 in C1) *and* who bought/ordered over the Internet in the last 12 months, i.e. 1st (Within the last 3 months) **or 2nd option** (Between 3 months and a year ago) **in F1]**

[Type: multiple answers allowed, i.e. *Tick all that apply*]

The question should reveal the *payment methods* used for e-commerce in the last 12 months and whether preference is more for on- or offline payment (is was previously included in the 2009 survey).

a) Provided credit or debit card details over the Internet

b) Provided pre-paid card or prepaid account details over the Internet

Several types of services are possible. The provider of the payment service can include a seller itself or an intermediary, such as Paypal or click&buy etc. Details of prepaid account to be given during the purchase can just involve the e-mail address which is usually required to enter in the case of Paypal on the payment webpage.

The method of pre-payment for the card or account (can be e.g. by credit card) does not matter for the applicability of the item b). If e.g. credit card details are directly provided during the order process, than item a) should be ticked.

c) Electronic bank transfer via Internet banking

d) Payment not via Internet (in cash, via ordinary bank transfer etc)

II.1.6.7 MODULE G: E-SKILLS

G1: Which of the following computer related activities have you already carried out?

[Scope: individuals who already used a computer, i.e. 1st, 2nd or 3rd option in B1]

[Type: multiple answers allowed, i.e. *Tick all that apply*]

The purpose of this question is to assess the skills level relating to general computer use of the respondents (questions on e-skills levels were included before in 2005, 2006, 2007, 2009, 2011, 2012, 2013 surveys).

Although the items a) to f) are more or less ordered from less complex to more complex, a respondent doesn't necessarily need to have ticked e.g. item a, b and c before item d can be ticked. In other words, there is no strict order in the items presented.

a) Copying or moving a file or folder

b) Using copy and paste tools to duplicate or move information within a document

E.g. in a Word document.

c) Using basic arithmetic formulas in a spreadsheet

E.g. in an Excel sheet.

d) Compressing files

This activity applies when the respondent has compressed a file by using Winzip or related programs to limit the size of large files.

e) Connecting and installing new devices, e.g. a printer or a modem

This activity covers both the hardware and the software side of the operation, even if both can be relatively straightforward nowadays because of plug-and-play USB connections.

f) Writing a computer program using a specialised programming language

Specialised programming language refers to high-level computer languages such as BASIC, Pascal or for instance SAS programming, which might only be familiar to programmers, system analysts and so forth.

The above items existed in previous questionnaires and items a) to f) should be considered for classifying and comparing from one year to another the categories low level computer skills (1-2 items ticked), medium (3-4 items ticked) and high level skills (5-6 items ticked).

*The following options g) to j) are **new** since the 2011 survey*:

g) Transferring files between computer and other devices (from digital camera or from/to mobile phone, mp3/mp4 player)

h) Modifying or verifying the configuration parameters of software applications (except internet browsers) - *optional*

Most computers are pre-configured with some computer software but users may change the settings to execute the software in the desired manner. This can include the change of the set of basic options or preferences in common software applications (e.g. word processing, spreadsheet, database, presentation, e-mail, photo editing or computer games). Examples are setting the default folder to open and save files, change security settings or changing security settings. More complex computer applications can require users to perform post-installation configuration to make the software truly operational.

i) Creating electronic presentations with presentation software (e.g. slides), including e.g. images, sound, video or charts

Electronic presentations could be powerpoint slides but also photo presentations in form of books etc. for which use of specific software installed, downloaded or directly used via the internet would be required.

j) Installing a new or replacing an old operating system

E.g. replacing Windows Vista by Windows 7 operating system for the computer.

k) None of the above

This is a closing item, enabling checking if the before-mentioned items were reviewed by respondents. Persons with "no computer skills" are those who have no experiences in any of the above items a) to j) *plus* those who have never used a computer.

The order of items g) to j) reflect some degree of sophistication, with g) the transfer of files between devices being at the more bottom end as it can be done by using the USB port of the computer for plug and play connections. Item i) relates to the use of presentation software. Items h) and j) reflect more the aspects of IT expert proficiency, but users in other jobs than IT specialist jobs may well have certain capabilities to perform such activities.

II.1.6.8 MODULE H: SOCIO-DEMOGRAPHIC BACKGROUND CHARACTERISTICS

[Scope: all households / all individuals]

[Type: H2, H5, H6, H7, H8, H12, H13: one single answer needed, i.e. *Tick only one* ; H1, H14, H15, H16: quantitative ; H3, H4, H9, H10, H11: open question, i.e. description needed (or coded by the interviewer)]

This module does not focus on ICT-related study variables, but on background variables to put the results on the study variables in perspective. These background variables should enable to relate the outcome of the surveys to the sociologic, demographic and economic background of the observed statistical units.

From the 2012 survey onwards, the socio-demographic background characteristics all reflect the agreed Core Social Variables to be included in all Eurostat social surveys.

For an overview about the guidelines concerning all characteristics below, please consult the latest **CSV Guidelines of Eurostat** on public Circa:

https://circabc.europa.eu/w/browse/c9df3412-21d8-4ef6-9312-fd68a9ff0c9e

or check the document in the following Circabc folder:

Europa > European Commission > CIRCABC > Eurostat > EMIS annexes >
Library > data_-_database > theme_3_-_popul > isoc > methodological_informati

In the model questionnaire, this module is inserted at the end of the survey. This was an arbitrary choice, the module can also be the starting point for the survey (in most cases, the information will partly be available from other sources, e.g. registers or the survey the ICT usage questions are embedded in).

The questions on age (H1), gender (H2), country of birth (H3), country of citizenship (H4), legal marital status (G5), de facto marital status (H6), educational attainment (H7), employment situation (H8) and occupation (H9) are to be answered by the individual(s) who is (are) selected within the household.

The questions on regions of residence (H10, H11), geographical location (H12) and type of locality (H13) refer to the household, but the answer at household level should coincide with the individual level anyway as the individuals are part of the household. Usually these two variables do not need to be directed to the respondents as the NSI's can derive the information on the basis of the household's address.

The questions on the household composition (H14, H15) and household income (H16) of course refer to the household level.

H1: Age

The data transmitted to Eurostat is broken down by age group. Whether the year of birth or date of birth or the age is collected, is relatively unimportant. In a postal survey, the NSI could simply ask the respondent to tick the appropriate age group. However, in most cases, this information will be available from the sampling frame (e.g. the population register).

Whether the reference point is the age during the interview or rather the age during the reference period, is not expected to have a significant impact on the results and comparability.

H2: Gender

Self-explanatory.

H3: Country of birth

The individual's country of birth should be coded according the following categories:

Native-born or foreign-born.

For foreign-born: born in another EU Member State or born in non-EU country.

H4: Country of citizenship

The individual's country of citizenship should be coded according the following two categories:

1. National, or 2. non-national.
That means the individual has or does not have the citizenship of the reporting country.

For non-national: national of other EU Member State or national of non-EU country.

More information on the background variables H3 and H4 (mandatory since 2011) can be found in the Report on Core Social Variables of Eurostat.[21]

In the case of dual or multiple citizenship, the following approach should be used:

1. If person has citizenship of the country of usual residence, will normally be recorded in the survey with the first priority.

2. If no citizenship is that of the country of residence, another EU MS citizenship has priority.

3. In other cases person may choose which country of citizenship will be recorded in survey.

H5: Legal marital status (optional)

21

https://circabc.europa.eu/w/browse/c9df3412-21d8-4ef6-9312-fd68a9ff0c9e

or check the document in the following Circabc folder:

Europa > European Commission > CIRCABC > Eurostat > EMIS annexes >
Library > data_-_database > theme_3_-_popul > isoc > methodological_informati

Legal marital status is defined as the (legal) conjugal status of each individual in relation to the marriage laws (or customs) of the country (i.e. de jure status).

categories Labels

1Unmarried (i.e. never married)

2Married (including registered partnership)

3 Widowed and not remarried (including widowed from registered partnership)

4 Divorced and not remarried (including legally separated and dissolved registered partnership)

All individuals in survey will be classified as unmarried (never married), married, widowed and not remarried or divorced and not remarried.

Information on the legal marital status of each person should be collected for all persons aged 16 to 74 years.

H6: De facto marital status (optional)

De facto marital status is defined as the marital status of each individual in terms of his or her actual living arrangements within the household.

categories Labels

1Person living in a consensual union

2Person not living in a consensual union

De facto marital status is used for identifying persons living in consensual union. Two persons are taken to be partners in a consensual union when they have usual residence in the same household, are not married to each other, and have a marriage-like relationship to each other.

Distinction will be made between persons living in consensual union and persons not living in consensual union.

Information on the de facto marital status should be collected for the same age groups as for the legal marital status. I.e. it should be collected even if respondents answered they are married.

Person who lives with another partner than his/her legal husband/wife is considered living in consensual union.

Person who lives with partner with whom he/she is legally married is not living in consensual union.

The two definitions on marital status are in accordance with the Conference of European Statisticians Recommendations for the 2010 Censuses of Population and Housing (United Nations Economic Commission for Europe and the Statistical Office of European Communities, 2006).

H7: Educational attainment level

The educational attainment level of an individual is the highest ISCED level successfully completed, the successful completion of educational programme being validated by a recognised qualification (or credential), i.e. a qualification officially recognized by the relevant national education authorities.

The expression 'level successfully completed' must be associated with obtaining a certificate (qualification). Eurostat recommends for all household surveys to use "diploma approach" in the collection of data on educational attainment, e.g. to ask about the diploma obtained. In cases where there is no certification (e.g. in primary education), successful completion must be associated with full attendance or the acquired competences to access the upper level. Persons who have not completed their studies (dropped-out) should be coded according to the highest level they have successfully completed.

In order to obtain comparable information for the different countries, the levels of education are linked to the UNESCO's International Standard Classification of Education – **from 2014 ISCED 2011 will be implemented in all EU surveys**.

A detailed description of the classification is available from the UNESCO Institute for Statistics' website: http://www.uis.unesco.org/Education/Documents/isced-2011-en.pdf

For the purpose of compliance with the core variables for Eurostat social surveys, information on the educational attainment level should be transmitted according to ISCED 2011 codes: 0, 1, 2, 3, 4, 5, 6, 7 and 8:

- **0 - Less than primary education**
- **1 - Primary education**
- **2 - Lower secondary education**
- **3 - Upper secondary education**
- **4 - Post-secondary non-tertiary education**
- **5 – Short-cycle tertiary education**
- **6 – Bachelor's or equivalent level**
- **7 – Master's or equivalent level**
- **8 – Doctoral or equivalent level**

The main tool to be used for the coding of educational attainment is the national ISCED integrated mapping (prepared in each country by education statisticians). This mapping classifies in ISCED all national educational programmes and formal qualifications (educational attainment).

For all questions related to the measurement of educational attainment according to ISCED 2011, please consult the methodological guidelines available on:

https://circabc.europa.eu/sd/a/07bba876-187a-41cf-8f3f-90c324e3bcea/ISCED%202011%20-%20Guidelines%20on%20educ_attainment%20(1st%20version%20-%202013-05).pdf

H8: Employment situation

The survey in the past always did not contain a number of detailed questions to assess an individual's employment situation or labour status. This variable was and is more related to the LFS question on the *self-declared labour status* as recommended in the Core Social Variables

(CSV) Guidelines (than on the detailed ILO labour status questions) where the respondent gives a self-perception regarding his/her activity status. In addition, employment situation includes as well the core variable *Status in employment (optional variables self-employed, employee, employee with a permanent job or contract of unlimited duration or with a temporary job or contract of limited duration)*.

The variables on employment situation have been extended since 2012, covering as well all other variables of Eurostat's Core Social Variables, agreed by the Directors of Social Statistics, at least with optional status as visible in the model questionnaire.

For further reference and explanations on all variables, please refer to the updated **CSV document:**

https://circabc.europa.eu/w/browse/c9df3412-21d8-4ef6-9312-fd68a9ff0c9e

or check the document in the following Circabc folder:

Europa > European Commission > CIRCABC > Eurostat > EMIS annexes >
Library > data_-_database > theme_3_-_popul > isoc > methodological_informati

The explanatory notes below have been included in the previous manuals and are kept in case of need. The concepts and definitions used are based on those used in Eurostat's labour market statistics and recommended by the International Labour Organisation (ILO).

Employee

Employees are defined as persons who work, by agreement, for a public or private employer and who receive compensation in the form of wages, salaries, fees, gratuities, payment by results or payment in kind; irrespective of the number or hours worked (full time or part time) and the duration of the contract (fixed or indefinite).

This category includes non-conscripted members of the armed forces, apprentices or trainees receiving remuneration, seasonal workers and persons on strike.

This category does not include voluntary workers and family workers who don't receive a regular monetary wage.

Persons who are simultaneously working in their own professional practice and for a public or private employer (e.g. doctors with their own cabinet and working in a hospital) should be classified according to the status where they work a more important number of hours.

Self-employed

Self-employed persons, with or without employees, are defined as persons who work in their own or jointly-owned business, professional practice or farm for the purpose of earning a profit.

This category includes own-account workers, family workers who don't receive a regular monetary wage, freelancers (although a person who has been regularly retained by a single employer for some time may also be regarded as an employee) and persons who look after one or more children that are not their own on a private basis and receiving a payment for this service.

Persons who are simultaneously working in their own professional practice and for a public or private employer (e.g. doctors with their own cabinet and working in a hospital) should be classified according to the status where they work a more important number of hours.

Unemployed

In accordance with the ILO standards used in the Labour Force Survey, unemployed persons comprise persons aged 15 to 74 who are (the three conditions must be fulfilled):

(a) without work during the reference week (the week of the interview), i.e. neither have a job nor are at work (for one hour or more) in paid employment or self-employment;

(b) currently available for work, i.e. are available for paid employment or self-employment before the end of the two weeks following the reference week;
(c) actively seeking work, i.e. have taken specific steps in the four weeks period ending with the reference week to seek paid employment or self-employment or who found a job to start later, i.e. within a period of at most three months.
The following are considered as specific steps in "actively seeking work":
- having been in contact with a public employment office to find work, whoever took the initiative (renewing registration for administrative reasons only is not an active step),
- having been in contact with a private agency (temporary work agency, firm specialising in recruitment, etc.) to find work,
- applying to employers directly,
- asking among friends, relatives, unions, etc., to find work,
- placing or answering job advertisements,
- studying job advertisements,
- taking a recruitment test or examination or being interviewed,
- looking for land, premises or equipment,
- applying for permits, licences or financial resources.

Student

This category refers to individuals in formal education (in educational institutions such as school, university, etc.) and not belonging to any of the other categories of employment situation.

When strictly applying the LFS definitions, students with small jobs have to be classified as employees. However, due to the absence of the long list of questions used to determine the employment status in the LFS, students who work only a few hours per week may be classified as students.

Other not in the labour force

This category refers to persons who are not employed or unemployed. Students are considered as a separate category (see above).

"Other not in the labour force includes persons in retirement or early retirement, persons in compulsory military service, persons fulfilling domestic tasks (without being an employee or self-employed, e.g. housewives) or other inactive persons.

As some countries embed the ICT usage survey in the Labour Force Survey or other socio-economic surveys using LFS definitions, it is useful to briefly discuss the correspondence between the five above categories and the 9 categories used in the LFS variable *Main labour status* (see column 101 in the LFS handbook or in the Annex to Regulation 1575/2000).

ICT usage survey *Employment situation*		**Labour Force Survey** *Main Labour Status*
Employee	→	LFS col.101, code 1 *(carries out a job or profession, including unpaid work for a family business or holding, including an apprenticeship or paid traineeship, etc.)*
Self-employed	→	LFS col.101, code 1 *(carries out a job or profession, including unpaid work for a family business or holding, including an apprenticeship or paid traineeship, etc.)*
Unemployed	→	LFS col.101, code 2 *(unemployed)*
Student (not in the labour force)	→	LFS col.101, code 3 *(pupil, student, further training, unpaid work experience)*
Other not in the labour force	→	LFS col.101, codes 4, 5, 6, 7 and 8 *(in retirement or early retirement or has given up business (=4); permanently disabled (=5); in compulsory military service (=6); fulfilling domestic tasks (=7); other inactive person (=8).*

The above correspondence table shows a former 1-1 relation between the breakdown in the ICT usage survey and in the Labour Force Survey (with codes 4, 5, 6, 7 and 8 combined into one single category as the subcategories are less relevant for the ICT usage survey or could be problematic in terms of low cell frequencies).

The breakdown "employee or self-employed" is obligatory since the survey year 2007 in the ICT usage survey. Up to 2006, the additional breakdown in employee and self-employed could not be distinguished directly from column 101 in the LFS as they are grouped in code 1. A pragmatic approach - with an acceptable risk of coding error – was suggested to additionally use the *Professional status* from the LFS (column 26) and file all persons (in code 1 of column 101) who appear to be *self-employed* or *family worker* (codes 1, 2 or 4 in LFS column 26) under code *self-employed* for the ICT usage survey and all others under *employees*. In other words, the auxiliary information from LFS column 26 was then only used to redirect some persons into the group of *self-employed* while the rest remain in the group of *employees*. This recoding suggestion supposes a certain correlation between the professional status according to the ILO definitions (= col.26) and ones self-assessment on the labour status (= col.101).

H9: Occupation

[Scope: individuals in employment, i.e. 1st *(employee)* or 2nd option *(self-employed)* in H8]

The new classification ISCO-08 on occupations (adopted in 2008) should be considered from 2011 onwards. It will be used in most of Eurostat's social statistics and it is based on the ILO's International Standard Classification of Occupations.

For information on the new ISCO-08, please see
http://www.ilo.org/public/english/bureau/stat/isco/index.htm

1. *ICT/non-ICT and manual/non-manual workers*

ICT professionals

It is now possible to code and store the occupations at the more aggregated level of submajor groups (2 digit). The major groups (1 digit) don't allow to distinguish ICT professionals from other workers.

At least the following sub-major groups should be considered for ICT professionals:

25 Information and communications technology professionals

35 Information and communications technicians

If possible, some ICT occupations classified in other groups should also be taken into account:
133 ICT service managers

2356 ICT teachers

7422 ICT installers and servicers

Non ICT professionals

The category *Non ICT professionals* consists of the individuals in all other ISCO Unit Groups (in the case of 4-digit coding) or all other ISCO Minor Groups (in the case of 3-digit coding).

For the breakdown manual / non-manual workers, the following applies:

Manual workers

This category corresponds to major groups 6 to 9 of ISCO-08:
- Major group 6: Skilled agricultural, forestry and fishery workers;
- Major group 7: Craft and related trades workers;
- Major group 8: Plant and machine operators and assemblers;
- Major group 9: Elementary occupations.

Non-manual workers

This category corresponds to major groups 0 to 5 of ISCO-08.
- Major group 1: Managers;
- Major group 2: Professionals;
- Major group 3: Technicians and associate professionals;
- Major group 4: Clerical support workers;
- Major group 5: Service and sales workers;
- Major group 0: Armed forces.

The classification is available in the RAMON-database (>classifications):
http://ec.europa.eu/eurostat/ramon
http://ec.europa.eu/eurostat/ramon/other_documents/index.cfm?TargetUrl=DSP_ISCO_08

The basis for the classification is the job and the skills. A job is defined as the set of tasks and duties to be performed. Skills are the abilities to carry out the tasks and duties of a job. Skills consist of two dimensions: skill level – which is usually related to the level of educational attainment - and domain specialisation.

The purpose of the variable "occupation" as a background characteristic is not to collect data on ICT usage broken down by individual occupations (this would necessitate very large samples), but rather by groups of occupations. E.g. manual versus non-manual or ICT-jobs versus non ICT-jobs.

2. *Occupations according to all 2-digit ISCO-08 categories*

The transmission of all occupations recoded into at least 2-digit ISCO-08 categories is requested in the CSV guidelines, but optional currently in the ICT survey.

A table with the 2-digit ISCO classification is provided in the Transmission Format.

For further information, please refer to the above link on the CSV Guidelines.

H10: Region of Residence – NUTS 1

The NUTS 1 regions have been introduced as being mandatory in 2008. Most, but not all countries are concerned. More information about NUTS 1 regions can be found at:

http://epp.eurostat.ec.europa.eu/portal/page/portal/nuts_nomenclature/introduction

The new NUTS classification, valid from 1 January 2012 is called 'NUTS 2010/EU-27' and is also described in the following publication:

http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-RA-11-011/EN/KS-RA-11-011-EN.PDF

Further information can be found on the web link:
http://ec.europa.eu/eurostat/ramon/index.cfm?TargetUrl=DSP_PUB_WELC

To see which codes have been changed between version NUTS 2006 and NUTS 2010:

http://epp.eurostat.ec.europa.eu/portal/page/portal/nuts_nomenclature/history_nuts

H11: Region of Residence – NUTS 2 (optional)

The NUTS 2 regions have been introduced as being optional in 2008. Most, but not all countries are concerned. More information about NUTS 2 regions can be found at:

http://epp.eurostat.ec.europa.eu/portal/page/portal/nuts_nomenclature/introduction

The new NUTS classification, valid from 1 January 2012 is called 'NUTS 2010/EU-27' and is also described in the following publication:

http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-RA-11-011/EN/KS-RA-11-011-EN.PDF

Further information can be found on the web link:
http://ec.europa.eu/eurostat/ramon/index.cfm?TargetUrl=DSP_PUB_WELC

To see which codes have been changed between version NUTS 2006 and NUTS 2010:

http://epp.eurostat.ec.europa.eu/portal/page/portal/nuts_nomenclature/history_nuts

H12: Geographical location

From 2014 onwards, a new classification of regions according to the Cohesion Policy 2014-2020 should be applied.[22] The following types of regions are to be distinguished according to their levels of development:

Less developed region: GDP < 75% of EU-27 average
Transition region: GDP 75% to 90% of EU-27 average
More developed region: GDP > 90% of EU-27 average

The map of European regions is available via the following link:

http://ec.europa.eu/regional_policy/what/future/img/eligibility20142020.pdf

A draft list of regions is available and shown on the next pages (endorsement by Commission decision is pending). The final list is planned to be inserted in the next update of the manual. DG REGIO informed that in conformity with art. 90 of the Common Provisions Regulation on cohesion policy (http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:347:0320:0469:EN:PDF), this list is based on the "2007" breakdown of NUTS regions; any more recent change in the NUTS classification will not change the eligibility of the areas in this list.

[22] http://ec.europa.eu/regional_policy/what/future/index_en.cfm

Regions by category, 2014-2020

MFF Council conclusions, 08/02/2013

MS	NUTS2 region *		Population 1000 inh.	Eligibility status 2014-2020	GDP/head, index EU27=100, average 2007-08-09
BE	BE10	Région de Bruxelles-Capitale / Brussels Hoofdstedelijk Gewest	1058.7	more developed	220.2
BE	BE21	Prov. Antwerpen	1722.5	more developed	135.3
BE	BE22	Prov. Limburg (B)	829.4	more developed	96.1
BE	BE23	Prov. Oost-Vlaanderen	1414.2	more developed	105.0
BE	BE24	Prov. Vlaams-Brabant	1064.2	more developed	123.4
BE	BE25	Prov. West-Vlaanderen	1152.4	more developed	110.2
BE	BE31	Prov. Brabant Wallon	374.6	more developed	116.0
BE	BE32	Prov. Hainaut	1301.8	transition	76.4
BE	BE33	Prov. Liège	1056.7	transition	86.7
BE	BE34	Prov. Luxembourg (B)	265.3	transition	78.5
BE	BE35	Prov. Namur	466.9	transition	81.6
BG	BG31	Severozapaden	923.0	less developed	27.5
BG	BG32	Severen tsentralen	928.1	less developed	29.2
BG	BG33	Severoiztochen	991.6	less developed	35.9
BG	BG34	Yugoiztochen	1123.6	less developed	35.0
BG	BG41	Yugozapaden	2114.8	less developed	71.4
BG	BG42	Yuzhen tsentralen	1541.7	less developed	29.9
CZ	CZ01	Praha	1221.6	more developed	173.8
CZ	CZ02	Střední Čechy	1214.5	less developed	74.3
CZ	CZ03	Jihozápad	1199.2	less developed	70.2
CZ	CZ04	Severozápad	1139.4	less developed	64.2
CZ	CZ05	Severovýchod	1501.1	less developed	66.0
CZ	CZ06	Jihovýchod	1657.1	less developed	72.8
CZ	CZ07	Střední Morava	1232.2	less developed	64.1
CZ	CZ08	Moravskoslezsko	1249.6	less developed	67.9
DK	DK01	Hovedstaden	1654.8	more developed	147.5
DK	DK02	Sjælland	819.8	transition	88.6
DK	DK03	Syddanmark	1196.4	more developed	110.5
DK	DK04	Midtjylland	1241.3	more developed	113.7
DK	DK05	Nordjylland	579.1	more developed	108.9
DE	DE11	Stuttgart	4006.5	more developed	137.2
DE	DE12	Karlsruhe	2738.3	more developed	131.9
DE	DE13	Freiburg	2195.8	more developed	113.1
DE	DE14	Tübingen	1807.7	more developed	122.3
DE	DE21	Oberbayern	4316.8	more developed	161.9
DE	DE22	Niederbayern	1192.5	more developed	115.0
DE	DE23	Oberpfalz	1085.0	more developed	122.9
DE	DE24	Oberfranken	1085.3	more developed	112.5
DE	DE25	Mittelfranken	1712.4	more developed	132.1
DE	DE26	Unterfranken	1330.6	more developed	115.7
DE	DE27	Schwaben	1787.1	more developed	119.8
DE	DE30	Berlin	3422.3	more developed	101.7
DE	DE41	Brandenburg - Nordost	1144.1	transition	76.4
DE	DE42	Brandenburg - Südwest	1385.0	transition	88.9
DE	DE50	Bremen	662.1	more developed	158.6
DE	DE60	Hamburg	1770.5	more developed	187.1
DE	DE71	Darmstadt	3782.8	more developed	156.6
DE	DE72	Gießen	1050.6	more developed	107.7
DE	DE73	Kassel	1234.9	more developed	113.8
DE	DE80	Mecklenburg-Vorpommern	1672.0	transition	82.5
DE	DE91	Braunschweig	1628.2	more developed	109.7
DE	DE92	Hannover	2153.1	more developed	112.8
DE	DE93	Lüneburg	1698.8	transition	83.9
DE	DE94	Weser-Ems	2479.7	more developed	101.1
DE	DEA1	Düsseldorf	5197.9	more developed	132.7
DE	DEA2	Köln	4386.0	more developed	117.0
DE	DEA3	Münster	2609.4	more developed	98.6
DE	DEA4	Detmold	2054.7	more developed	109.8
DE	DEA5	Arnsberg	3710.7	more developed	106.9
DE	DEB1	Koblenz	1503.1	more developed	96.1
DE	DEB2	Trier	515.2	more developed	94.5
DE	DEB3	Rheinhessen-Pfalz	2017.3	more developed	105.4
DE	DEC0	Saarland	1033.2	more developed	112.3
DE	DED1	Chemnitz	1495.3	transition	83.0
DE	DED2	Dresden	1642.3	transition	88.4
DE	DED3	Leipzig	1068.4	more developed	90.0
DE	DEE0	Sachsen-Anhalt	2397.7	transition	83.9
DE	DEF0	Schleswig-Holstein	2834.2	more developed	100.6
DE	DEG0	Thüringen	2278.6	transition	83.7
EE	EE00	Eesti	1341.2	less developed	67.7
IE	IE01	Border, Midland and Western	1188.4	more developed	95.6
IE	IE02	Southern and Eastern	3237.1	more developed	151.1

GR	GR11	Anatoliki Makedonia, Thraki	606.8	less developed	67.8
GR	GR12	Kentriki Makedonia	1940.6	less developed	74.8
GR	GR13	Dytiki Makedonia	293.4	transition	85.6
GR	GR14	Thessalia	736.2	less developed	71.0
GR	GR21	Ipeiros	353.6	less developed	64.4
GR	GR22	Ionia Nisia	230.1	transition	84.6
GR	GR23	Dytiki Ellada	740.7	less developed	66.2
GR	GR24	Sterea Ellada	555.0	transition	85.9
GR	GR25	Peloponnisos	592.9	transition	75.1
GR	GR30	Attiki	4073.6	more developed	120.1
GR	GR41	Voreio Aigaio	200.4	transition	75.5
GR	GR42	Notio Aigaio	306.7	more developed	112.4
GR	GR43	Kriti	607.7	transition	83.2
ES	ES11	Galicia	2734.7	more developed	91.0
ES	ES12	Principado de Asturias	1058.4	more developed	96.6
ES	ES13	Cantabria	572.9	more developed	101.8
ES	ES21	País Vasco	2134.4	more developed	135.2
ES	ES22	Comunidad Foral de Navarra	609.4	more developed	130.5
ES	ES23	La Rioja	313.0	more developed	113.3
ES	ES24	Aragón	1303.8	more developed	114.3
ES	ES30	Comunidad de Madrid	6219.4	more developed	135.4
ES	ES41	Castilla y León	2503.0	more developed	99.4
ES	ES42	Castilla-La Mancha	1996.8	transition	83.3
ES	ES43	Extremadura	1079.2	less developed	71.8
ES	ES51	Cataluña	7241.5	more developed	121.0
ES	ES52	Comunidad Valenciana	4931.6	more developed	93.4
ES	ES53	Illes Balears	1054.2	more developed	111.3
ES	ES61	Andalucía	8090.6	transition	79.9
ES	ES62	Región de Murcia	1425.0	transition	87.2
ES	ES63	Ciudad Autónoma de Ceuta	72.3	more developed	93.8
ES	ES64	Ciudad Autónoma de Melilla	69.3	transition	88.4
ES	ES70	Canarias	2055.5	transition	89.4
FR	FR10	Île de France	11687.7	more developed	175.3
FR	FR21	Champagne-Ardenne	1336.1	more developed	95.0
FR	FR22	Picardie	1906.6	transition	82.9
FR	FR23	Haute-Normandie	1826.0	more developed	93.7
FR	FR24	Centre	2533.5	more developed	91.1
FR	FR25	Basse-Normandie	1468.0	transition	84.2
FR	FR26	Bourgogne	1638.2	more developed	92.0
FR	FR30	Nord - Pas-de-Calais	4021.1	transition	87.7
FR	FR41	Lorraine	2344.7	transition	84.7
FR	FR42	Alsace	1840.7	more developed	101.1
FR	FR43	Franche-Comté	1165.2	transition	86.4
FR	FR51	Pays de la Loire	3521.5	more developed	96.6
FR	FR52	Bretagne	3157.8	more developed	90.4
FR	FR53	Poitou-Charentes	1756.4	transition	85.9
FR	FR61	Aquitaine	3187.1	more developed	94.6
FR	FR62	Midi-Pyrénées	2849.0	more developed	95.3
FR	FR63	Limousin	741.5	transition	83.0
FR	FR71	Rhône-Alpes	6136.3	more developed	108.9
FR	FR72	Auvergne	1341.4	transition	86.9
FR	FR81	Languedoc-Roussillon	2594.5	transition	84.9
FR	FR82	Provence-Alpes-Côte d'Azur	4899.9	more developed	99.6
FR	FR83	Corse	304.8	transition	86.8
FR	FR91	Guadeloupe	448.0	less developed	65.6
FR	FR92	Martinique	398.1	less developed	71.6
FR	FR93	Guyane	222.4	less developed	51.1
FR	FR94	Réunion	813.8	less developed	65.4
HR	HR03	Jadranska Hrvatska	1466.3	less developed	60.4
HR	HR04	Kontinentalna Hrvatska	2966.9	less developed	63.0

IT	ITC1	Piemonte	4411.1	more developed	113.4
IT	ITC2	Valle d'Aosta/Vallée d'Aoste	126.5	more developed	131.4
IT	ITC3	Liguria	1612.3	more developed	110.9
IT	ITC4	Lombardia	9690.3	more developed	133.8
IT	ITD1	Provincia Autonoma Bolzano/Bozen	496.1	more developed	145.0
IT	ITD2	Provincia Autonoma Trento	516.4	more developed	123.7
IT	ITD3	Veneto	4853.6	more developed	121.0
IT	ITD4	Friuli-Venezia Giulia	1225.4	more developed	118.5
IT	ITD5	Emilia-Romagna	4304.7	more developed	129.1
IT	ITE1	Toscana	3689.7	more developed	112.6
IT	ITE2	Umbria	888.5	more developed	97.8
IT	ITE3	Marche	1559.8	more developed	105.6
IT	ITE4	Lazio	5591.8	more developed	121.0
IT	ITF1	Abruzzo	1327.7	transition	86.6
IT	ITF2	Molise	320.6	transition	83.6
IT	ITF3	Campania	5810.6	less developed	66.5
IT	ITF4	Puglia	4077.7	less developed	68.4
IT	ITF5	Basilicata	590.6	less developed	74.5
IT	ITF6	Calabria	2006.7	less developed	66.9
IT	ITG1	Sicilia	5032.5	less developed	68.1
IT	ITG2	Sardegna	1667.5	transition	79.1
CY	CY00	Κύπρος / Kypros	792.4	more developed	97.0
LV	LV00	Latvija	2265.7	less developed	54.3
LT	LT00	Lietuva	3357.7	less developed	58.6
LU	LU00	Luxembourg (Grand-Duché)	488.4	more developed	273.4
HU	HU10	Közép-Magyarország	2911.6	more developed	105.1
HU	HU21	Közép-Dunántúl	1103.7	less developed	56.3
HU	HU22	Nyugat-Dunántúl	998.0	less developed	60.9
HU	HU23	Dél-Dunántúl	957.0	less developed	43.4
HU	HU31	Észak-Magyarország	1230.1	less developed	39.6
HU	HU32	Észak-Alföld	1508.4	less developed	40.3
HU	HU33	Dél-Alföld	1330.1	less developed	42.2
MT	MT00	Malta	411.7	transition	78.7
NL	NL11	Groningen	574.2	more developed	175.5
NL	NL12	Friesland (NL)	644.2	more developed	107.0
NL	NL13	Drenthe	488.5	more developed	102.6
NL	NL21	Overijssel	1122.7	more developed	114.4
NL	NL22	Gelderland	1987.1	more developed	112.3
NL	NL23	Flevoland	381.2	more developed	99.2
NL	NL31	Utrecht	1205.7	more developed	155.5
NL	NL32	Noord-Holland	2637.3	more developed	150.2
NL	NL33	Zuid-Holland	3473.3	more developed	134.3
NL	NL34	Zeeland	380.9	more developed	122.6
NL	NL41	Noord-Brabant	2429.6	more developed	132.2
NL	NL42	Limburg (NL)	1123.7	more developed	117.1
AT	AT11	Burgenland (A)	282.1	transition	82.8
AT	AT12	Niederösterreich	1600.3	more developed	102.3
AT	AT13	Wien	1680.0	more developed	160.9
AT	AT21	Kärnten	560.3	more developed	105.0
AT	AT22	Steiermark	1205.9	more developed	108.3
AT	AT31	Oberösterreich	1408.6	more developed	123.6
AT	AT32	Salzburg	528.1	more developed	142.6
AT	AT33	Tirol	701.7	more developed	129.8
AT	AT34	Vorarlberg	366.7	more developed	131.0
PL	PL11	Łódzkie	2552.7	less developed	52.7
PL	PL12	Mazowieckie	5195.8	more developed	91.1
PL	PL21	Małopolskie	3283.3	less developed	49.2
PL	PL22	Śląskie	4651.4	less developed	61.2
PL	PL31	Lubelskie	2164.1	less developed	38.9
PL	PL32	Podkarpackie	2098.3	less developed	39.1
PL	PL33	Świętokrzyskie	1274.1	less developed	44.7
PL	PL34	Podlaskie	1192.4	less developed	42.0
PL	PL41	Wielkopolskie	3392.2	less developed	60.0
PL	PL42	Zachodniopomorskie	1692.8	less developed	50.8
PL	PL43	Lubuskie	1008.9	less developed	49.3
PL	PL51	Dolnośląskie	2877.7	less developed	61.9
PL	PL52	Opolskie	1035.3	less developed	47.3
PL	PL61	Kujawsko-Pomorskie	2066.9	less developed	49.1
PL	PL62	Warmińsko-Mazurskie	1426.9	less developed	42.2
PL	PL63	Pomorskie	2215.5	less developed	55.3
PT	PT11	Norte	3745.2	less developed	62.8
PT	PT15	Algarve	428.1	transition	86.1
PT	PT16	Centro (P)	2384.2	less developed	65.3
PT	PT17	Lisboa	2813.5	more developed	110.7
PT	PT18	Alentejo	758.9	less developed	72.3
PT	PT20	Região Autónoma dos Açores	244.3	less developed	73.2
PT	PT30	Região Autónoma da Madeira	246.8	more developed	103.0

RO	RO11	Nord-Vest	2722.1	less developed	41.7
RO	RO12	Centru	2524.0	less developed	44.0
RO	RO21	Nord-Est	3719.9	less developed	28.3
RO	RO22	Sud-Est	2820.9	less developed	36.3
RO	RO31	Sud - Muntenia	3285.5	less developed	37.6
RO	RO32	Bucureşti - Ilfov	2243.5	more developed	106.6
RO	RO41	Sud-Vest Oltenia	2264.2	less developed	34.4
RO	RO42	Vest	1923.8	less developed	50.3
SI	SI01	Vzhodna Slovenija	1082.5	less developed	73.3
SI	SI02	Zahodna Slovenija	945.0	more developed	106.6
SK	SK01	Bratislavský kraj	614.1	more developed	168.5
SK	SK02	Západné Slovensko	1864.4	less developed	68.0
SK	SK03	Stredné Slovensko	1350.2	less developed	56.9
SK	SK04	Východné Slovensko	1578.1	less developed	48.8
FI	FI13 **	Itä-Suomi	656.0	more developed	*94.0*
FI	FI18	Etelä-Suomi	2643.0	more developed	134.2
FI	FI19	Länsi-Suomi	1347.0	more developed	105.0
FI	FI1A **	Pohjois-Suomi	640.3	more developed	*94.0*
FI	FI20	Åland	27.3	more developed	151.1
SE	SE11	Stockholm	1966.4	more developed	170.7
SE	SE12	Östra Mellansverige	1540.7	more developed	104.9
SE	SE21	Småland med öarna	806.7	more developed	108.9
SE	SE22	Sydsverige	1359.5	more developed	108.0
SE	SE23	Västsverige	1845.7	more developed	116.7
SE	SE31	Norra Mellansverige	825.2	more developed	104.8
SE	SE32	Mellersta Norrland	370.3	more developed	111.4
SE	SE33	Övre Norrland	508.2	more developed	113.3
UK	UKC1	Tees Valley and Durham	1163.0	transition	78.5
UK	UKC2	Northumberland and Tyne and Wear	1408.8	more developed	90.5
UK	UKD1	Cumbria	495.9	transition	89.5
UK	UKD2	Cheshire	1003.2	more developed	117.1
UK	UKD3	Greater Manchester	2582.1	more developed	98.4
UK	UKD4	Lancashire	1446.2	transition	84.9
UK	UKD5	Merseyside	1351.2	transition	80.2
UK	UKE1	East Yorkshire and Northern Lincolnshire	914.3	transition	85.8
UK	UKE2	North Yorkshire	788.8	more developed	96.3
UK	UKE3	South Yorkshire	1308.3	transition	84.5
UK	UKE4	West Yorkshire	2207.7	more developed	98.1
UK	UKF1	Derbyshire and Nottinghamshire	2070.2	more developed	95.2
UK	UKF2	Leicestershire, Rutland and Northamptonshire	1660.6	more developed	107.8
UK	UKF3	Lincolnshire	695.0	transition	79.8
UK	UKG1	Herefordshire, Worcestershire and Warwickshire	1266.4	more developed	95.1
UK	UKG2	Shropshire and Staffordshire	1517.6	transition	83.9
UK	UKG3	West Midlands	2622.0	more developed	97.4
UK	UKH1	East Anglia	2325.9	more developed	102.3
UK	UKH2	Bedfordshire and Hertfordshire	1682.0	more developed	119.9
UK	UKH3	Essex	1703.0	more developed	92.7
UK	UKI1	Inner London	3031.5	more developed	327.0
UK	UKI2	Outer London	4643.2	more developed	99.2
UK	UKJ1	Berkshire, Buckinghamshire and Oxfordshire	2202.0	more developed	146.7
UK	UKJ2	Surrey, East and West Sussex	2654.2	more developed	116.2
UK	UKJ3	Hampshire and Isle of Wight	1856.1	more developed	111.0
UK	UKJ4	Kent	1654.3	more developed	92.5
UK	UKK1	Gloucestershire, Wiltshire and Bristol/Bath area	2306.8	more developed	117.4
UK	UKK2	Dorset and Somerset	1233.3	more developed	94.0
UK	UKK3	Cornwall and Isles of Scilly	531.1	less developed	73.2
UK	UKK4	Devon	1135.5	transition	88.1
UK	UKL1	West Wales and The Valleys	1892.4	less developed	70.1
UK	UKL2	East Wales	1096.1	more developed	102.7
UK	UKM2	Eastern Scotland	1979.3	more developed	111.5
UK	UKM3	South Western Scotland	2291.4	more developed	101.3
UK	UKM5	North Eastern Scotland	452.5	more developed	153.9
UK	UKM6	Highlands and Islands	445.6	transition	84.1
UK	UKN0	Northern Ireland	1774.4	transition	86.4

** in addition, Mayotte will be treated as a less developed region*

*** regions merged by Commission Regulation (EU) 31/2011 of 17 January 2011; category is determined at the level of the modified region*

Changes compared to classifications until 2013:

Up to 2006, Former "Objective 1" promotes the development and structural adjustment of regions whose development is lagging behind, i.e. whose average per capita GDP is below 75% of the European Union average.

From 2007 to 2013, the list of objective 1 regions (valid until end 2006) related to the list of **convergence regions (including related phasing out regions**), see also: OJ L 243/44 (6.9.2006) "Commission Decision of 4 August 2006 drawing up the list of regions eligible for funding from the Structural Funds under the Convergence objective for the period 2007-2013. The list was extended: In EU-27, Bulgarian and Romanian NUTS regions were convergence regions. Non-convergence including related phasing out regions (i.e. all other regions) were classified as **regional competitiveness and employment region**. The list of NUTS regions eligible for support from the Structural Funds under the Convergence Objective (ex-Objective 1), including phasing out Convergence regions, was valid for the period 2007-2013.

Further information including a list of NUTS regions was provided in previous manuals.

H13: New Degree of Urbanisation

The new classification on the Degree of Urbanisation has become applicable since 2012.

For further information (including tables in excel with the classification by country):

http://ec.europa.eu/eurostat/ramon/miscellaneous/index.cfm?TargetUrl=DSP_DEGURBA

The following largely represents an extract of the methodology document:

Definitions

The concept of "urbanisation" has been introduced to indicate the character of the area where the respondent lives. Three types of area have been identified and defined using a criterion of geographical contiguity in combination with a minimum population threshold based on population grid square cells of 1 km2. These grid cells all have the same shape and surface, which avoids distortions caused by using units varying in size. The three types to be distinguished are:

- densely-populated (Code 1)
- intermediate (Code 2)
- thinly-populated (Code 3)

The classification with areas by countries and respective coding is available via the above website. Yearly updates by June may be provided if countries communicated changes..

The degree of urbanisation creates a classification of all LAU2s (Local Administrative Units – Level 2) as follows:

Thinly populated area (alternative name: rural area)

- More than 50% of the population lives in rural grid cells.

Intermediate density area (alternative name: towns and suburbs/small urban area)

- Less than 50% of the population lives in rural grid cells and
- Less than 50% lives in high-density clusters

Densely populated area: (alternative names: cities/large urban area)

-At least 50% lives in high-density clusters[23]

In the above, the following definitions are used:

- Rural grid cells: Grid cells outside urban clusters
- Urban clusters: clusters of contiguous[24] grid cells of 1 km2 with a density of at least 300 inhabitants per km2 and a minimum population of 5 000.
- High-density cluster: Contiguous[25] grid cells of 1 km2 with a density of at least 1 500 inhabitants per km2 and a minimum population of 50 000. (Alternative names: urban centre or city centre)

In order to classify properly LAU2s based on the grid cell approach described, a few additional correction rules must be provided:

- If the LAU2s do not have a raster equivalent, they are classified according to the share of territory in rural grid cells and high-density clusters.
- Thinly populated LAU2s may be classified as intermediate or densely populated due to border effects if rural grid cells cover most of the territory. For that reason, LAU2s with a population below 5000 inhabitants[26] and 90% of its area in rural grid cells are reclassified as rural area.
- Very small densely populated LAU2s may be classified as thinly populated due to the coarse[27] resolution of the population grid. For that reason, LAU2s with an area less than 5 km² but with a share of surface outside rural grid cells higher than 30 % are reclassified as intermediate density or densely populated according to the share of the correspondent cluster.

[23] Furthermore, each high-density cluster should have at least 75% of its population in densely populated LAU2s. This also ensures that all high-density clusters are part of at least one densely populated LAU2, even when this cluster represents less than 50% of the population of the LAU2.

[24] Contiguity for urban clusters does include the diagonal (i.e. cells with only the corners touching). Gaps in the urban cluster are not filled (i.e. cells surrounded by urban cells).

[25] Contiguity for high-density clusters does not include the diagonal (i.e. cells with only the corners touching) and gaps in the cluster are filled (i.e. cells surrounded by high-density cells).

[26] Please note that this threshold refers to the population in the LAU2, whereas the threshold used in the definition of an urban cluster refers to the set of contiguous grid cells – the cluster – which may cover cells belonging to several LAU2s.

[27] "Coarse" in relation to the small area of these particular LAUS2s.

As LAU2s vary considerably in area, this methodology will lead to a closer match between a high-density cluster and densely populated LAU2s in countries with small LAU2s than in those with large LAU2s. To take this difference into account, the classification can be adjusted as following:

- A densely populated LAU2 can be classified intermediate as long as 75% of its high-density cluster population remains in densely populated LAU2s.
- An thinly populated or intermediate density LAU2 can be classified as densely populated if it belongs to a group of LAU2s with a political function and if the majority of population of this group of LAU2s lives in a high-density cluster.

A LAU2 consists of municipalities or equivalent units in the 27 EU Member States and correspond to the following entities:

Belgium Gemeenten/Communes
Czech Republic Obce
Denmark Kommuner
Germany Gemeinden
Estonia Vald, Linn
Greece Demotiko diamerisma/Koinotiko diamerisma
Spain Municipios
France Communes
Ireland DEDs/Wards
Italy Comuni
Cyprus Dimoi, koinotites
Latvia Pilsētas, novadi, pagasti
Lithuania Seniūnijos
Luxembourg Communes
Hungary Települések
Malta Kunsilli
The NetherlandsGemeenten
Austria Gemeinden
Poland Gminy
Portugal Freguesias
Slovenia Občine
Slovakia Obce
Finland Kunnat /Kommuner
Sweden Kommuner
United KingdomWards (or parts thereof)
Bulgaria Naseleni Mesta
Croatia
Romania Communes+Municipiu+Orajse
Turkey Köy
Iceland Sveitarfélag
Norway Kommuner
Switzerland Gemeinden / Communes / Comuni

Note: This new methodology of classifying urban and rural areas has been agreed by DG for Regional Policy, DG for Agriculture and Rural Development and Eurostat. It replaces the methodology used in the LFS so far.

The older methodology can be consulted in previous versions of the explanatory notes/manual.

The definition of urban clusters is drawn from the new methodology to classify urban and rural regions developed and agreed by DG for Regional Policy, DG for Agriculture and Rural Development, Eurostat and JRC in 2010. The definition of high-density clusters is drawn from work by the OECD and the DG for Regional Policy on a new metropolitan area definition done in 2011.

For further information: see document with methodology notes via the above web link.

H14 and H15: Household type

"Household" refers either to one person living alone or a group of people living together in the same dwelling unit with at least one person of the age of 16 – 74 years (see above: "General outline of the survey", Scope - age limit).

Excluded are permanent members of institutions such as hospitals, prisons, etc.

The household composition (as used in the tabulation scheme) is determined using two mandatory indicators: the total number of persons in the household (including all age groups) and the number of children under 16. The presence (and the number) of children is collected separately as this is a driving force for adoption of ICTs or Internet in a household.

In addition, several optional variables have been added as visible in the model questionnaire, corresponding to the CSV (see link to document above). If countries have chosen to collect the equivalised net monthly income, at least the number of persons aged 14-15 should be additionally collected, otherwise the calculation will not be possible.

H16: Household income (optional)

The "income" concept used in this survey is the net monthly income of the household (not the individual!). The focus is on the household income as one's access to and use of ICTs will rather be related to the total income of the household he/she is part of than by his/her personal income.

For persons in paid employment, this refers to the monthly 'take-home' pay, i.e. the pay after deducting income taxes, employee's social security contributions and any voluntary contributions.

Bonuses regularly paid at the time of each payment – such as regular overtime hours, bonuses for team, night or weekend work, tips and commissions - are included.

Bonuses that are not paid at each pay period (f.i. annual profit shares) but that are by accident paid in the considered pay period are excluded as they will lead to a non-representative estimate.

Irregular or annual bonuses should be excluded. No "redistribution" of such bonuses – i.e. dividing by 12 and adding to the monthly estimate – is required as this would unnecessarily increase the recall burden on the respondent. This may however make international comparisons less relevant as the impact of such bonuses tends to differ heavily from one country to another.

Where applicable and feasible, an estimate of the monetary value for the *payments in kind* should be added.

Although the *income from self-employment* may be difficult to measure, the respondents can be asked for an estimate of their disposable income.

Income from allowances such as unemployment benefits, pensions and sickness, disability or invalidity allowances should refer to the last monthly allowance received before the interview (or the monthly average for a recent period if this is more easy to collect or estimate).

Income from investments (assets, savings, stocks, shares, etc) is to be excluded as this information is difficult to measure and/or collect.

Although undesirable, derogations from the above definition of the household income may be justified in cases where the data can be obtained from slightly different income-related questions in the survey vehicle the survey on ICT usage is linked to. Especially when using such already available data is the only way of providing information on this sensitive, optional variable.

Ideally, information should be collected by asking the respondent for an estimate of the actual amount. Where it is not desirable or feasible to collect such information, the respondent can indicate a size band. These size bands should be compatible with the breakdown used in the tabulation scheme. This tabulation scheme foresees the use of income quartiles rather than income classes. Such approach should make cross-country comparisons more meaningful. Not only can this overcome the dispersion of income levels across countries in Europe, the use of quartiles can also avoid threats to comparability caused by the use of different concepts by different countries (e.g. gross versus net earnings, whether or not including regular monthly bonuses, whether or not including $1/12^{th}$ of the annual bonuses, whether or not including an estimate of the value of payments-in-kind, etc.).

If countries have chosen to collect in addition the equivalised net monthly income, at least the number of persons aged 14-15 should be also collected, otherwise the calculation will not be possible.

For additional information, see transmission format in *Annex II.3.2 Transmission format* as well as the CSV Guidelines, link provided above.

II.2. PRODUCTION METHODOLOGY

II.2.1. Timetable – Survey period

According to Annex II of Commission Regulation (EC) No 859/2013 of 5 September 2013 implementing Regulation (EC) No 808/2004 of the European Parliament and of the Council concerning Community statistics on the information society, "the reference period for the statistics to be collected is the first quarter of 2013" (see Part III of this manual, Annex III.2). As the majority of the questions in the model questionnaire and the variables in the above mentioned legal act refer to the last three months – which means in practice: *the three months preceding the interview* – the data collection should take place immediately after the first quarter. Taking into account that the fieldwork can take several weeks or even months, the recommendation is to interview the respondents in April-May (or more general: in the second quarter). This way, a satisfying overlap with the theoretical reference period is guaranteed.

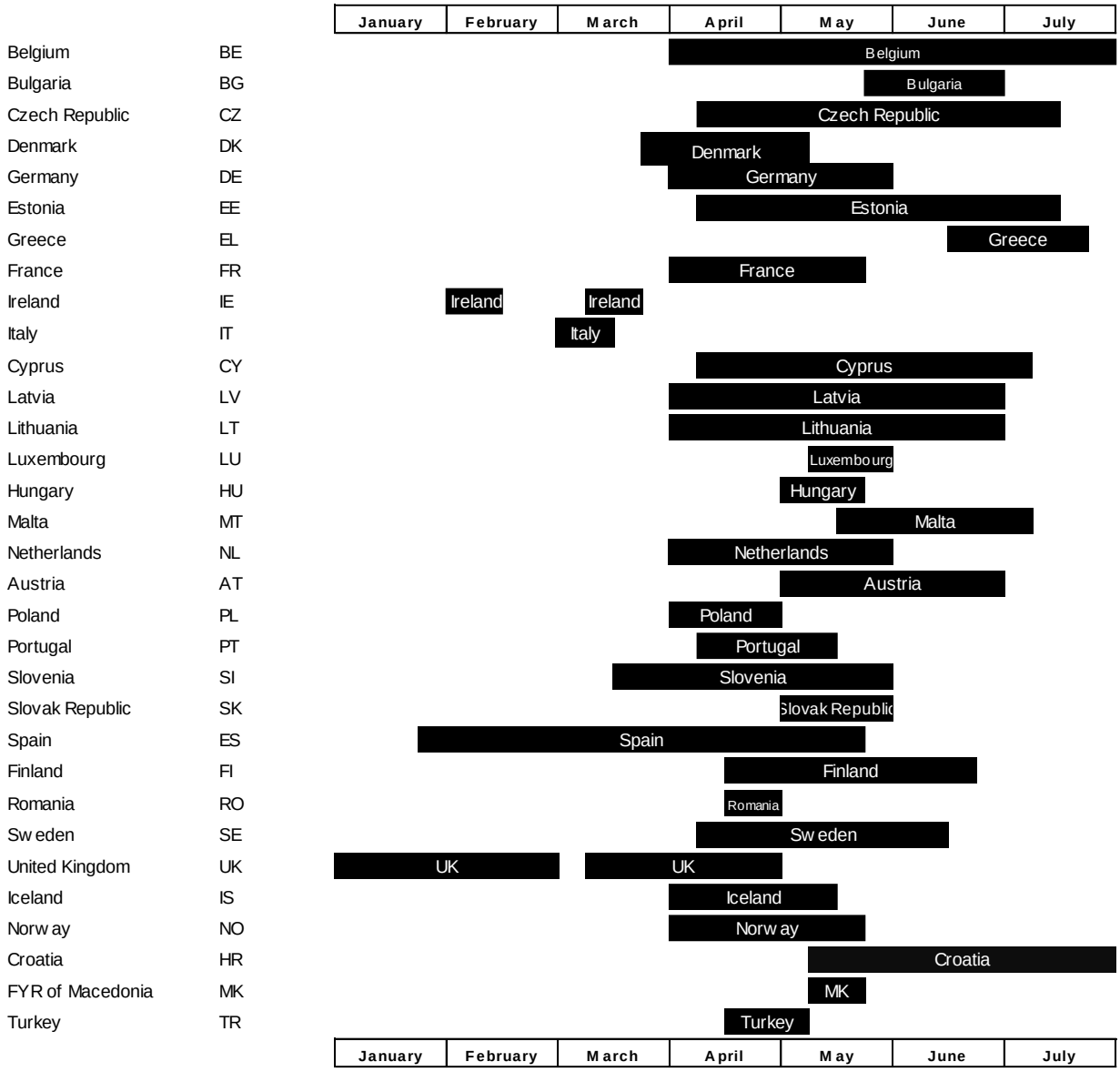
As mentioned in the previous paragraph, for most questions the reference period is 3 months. The choice for this 3 months' reference period is a trade-off between recall bias and seasonal bias. Using a reference period of 12 months filters out seasonality, but a recall period of 12 months may be too demanding for the respondents. In general, priority has been given to limiting the intellectual burden for the respondents, only for the questions related to e-commerce (where an important seasonal effect can be expected in consumption patterns throughout the year, e.g. more purchases Christmas, more booking of holidays during Summer) a longer reference period is used.

The figure below present an overview of the survey periods in the 2011 survey.

Looking at the results of previous surveys, the pattern for the last 3 months is not very difficult from the last 12 months' pattern. In terms of e.g. Internet use, this means that only few people will state they have used the Internet in the past 12 months, but not in the past 3 months. The harmonisation of the data collection periods across countries should further reduce the possible risk of seasonal bias when comparing across countries or from one year to another.

Figure: Overview of the survey periods (2011 survey)

Source: Quality reports for the Community survey on ICT usage in households and by individuals (2011)



II.2.2. **Frame population**

This issue was already discussed in chapter II.1.2 on the *target population*. The *frame population* (of *sampling population*) is the frame from which the sample will be drawn. Ideally, this list of units should be equivalent to the target population as both overcoverage and undercoverage can induce bias and affect the reliability of the survey results.

E.g. overcoverage due to inclusion of persons aged 75-79 in the sampling frame (making the simplistic assumption that the frame only has information on one's age using age bands, not the actual age or year of birth): as this age group to score relatively low on ICT usage indicators, the population total may show a downwards bias due to inclusion of this group.

E.g. undercoverage due to the fact that some remote areas are not covered in the register: as such areas usually have a lower penetration of ICTs, the overall indicators may be relatively overestimated for such country.

To allow for an assessment, it is important to include the coverage and the sampling frame when reporting metadata (see Chapters 3 and 5 in the reporting template, Annex II.3.3 –).

The table below presents the sampling frames used in the national statistical institutes for the 2011 survey.

The second table gives an overview of the target and non-target population. In most countries, the target population of individuals aged 16 to 74 represents approximately 75% of the total population. The target population for households covers the population relatively more (approximately 90%) as only households where *all* members are over 74 are excluded. This information has to be kept in mind when disseminating the results: the results are not representative for the total population but only for the subpopulation of persons in the age group 16 to 74. This difference can be important as the out-of-scope groups may have a different profile in terms of ICT use (heavy use among persons younger than 16 but low use among persons over 74).

Table: Overview of the sampling frames (2011 survey)

Source: Quoted from the quality reports for the Community survey on ICT usage in households and by individuals (2011)

Austria	The "Zentrales Melderegister (ZMR)" is used as sampling frame for the Continuous Labour Force Survey. A sub-sample of this survey was drawn for the ICT survey.
Belgium	Population Register (contains all residents of Belgium, regardless of nationality, together with information on household) Timeliness: Households drawn in Population Register according to the situation of a few months before the fieldwork between April and July. No geographical restriction
Bulgaria	The sample for ICT usage in households and by individuals was selected from the register prepared for the Population Census 2011 comprising Census regions, numerating units, dwelling and households. The register was based on the Population Census 2001 data base currently updated with data from the Information System "Demography" data (ISD) and the National Civil Registration System and supplemented with data from data bases on new built dwelling, maintained by the municipalities. Persons living in collective households and in institutions were excluded from the target population. The sample includes only the households with members between 16 and 74 years old. Sample was extracted from the databases last updated as of 31.12.2010.
Cyprus	The sampling frame used for the selection of the sample was the Population Census 2001 register. In addition, the newly constructed housing units since 2001 were also included in the frame. A list of domestic consumers of electricity that incorporates new customers was obtained from the Electricity Authority of Cyprus (EAC). It has already been tested that there is one to one correspondence between the definition of the housing units of EAC (domestic consumer of electricity) and the statistical definition of household. The problem of timeliness exists, since the sampling frame used was the Population Census 2001 register which has as reference date the 1st of October 2001. However, to minimise the problem, the newly constructed housing units, obtained from the Electricity Authority of Cyprus, are incorporated in the frame. The results will be adjusted to the 2009 population estimates.
Czech Republic	Census Enumeration Districts (CEUs) constitute the first-stage sampling units. CEUs are small geographical areas covering the whole territory of the country. They are used as enumeration districts during the census, but their use is more general. Continuously updated geographical register is maintained by the CSU, where these units form the basic geographical layer, on which subsequent aggregations are based. This register is the base for an integrated hierarchical geographical information system and is the base for databases of regional indicators and statistical data.

	For each CEU, a list of all buildings is maintained in the register. This list is updated from administrative data of the construction authorities (new buildings', flats' or commercial premises' acceptation protocols, demolitions' protocols). For each building, the number of dwelling units is recorded. Some overcoverage exists in the CEUs register in terms of recorded residential status (usually about 4% of the samples are coded as "administrative waste" – non residential (recreational or commercial space, temporarily empty flats – no persons with usual residence, demolished buildings not yet deleted from the register). Small CEUs (in terms of number of residential dwellings) are currently dropped from the sampling frame due to practical sampling reasons (approximately 1.26% of the dwellings).
Denmark	The Central Population Register (CPR) contains general personal data on individuals with residence in Denmark. Individuals without registered address in the CPR do not enter the survey. Individuals having their name and phone number protected from use in research purposes do not enter the survey.
Estonia	The Population Register is used as sampling frame. Population Register is an electronic database containing the general personal data about all Estonian citizens and foreigners residing in Estonia. Population register is administered by the Estonian Ministry of the Interior. Time-lag between sample selection and start of survey is 5 months. Frame includes persons who have moved abroad.
Finland	Population database maintained by Statistics Finland. It is based and updated from the official Central Population Register of the Population Register Centre. It contains the demographic information of resident population, address and residence information etc.
France	The database is the local residence tax database, for phone survey and internet/paper survey. Timeliness : last update is January 2009.
Germany	Quota sample from the German micro-census 2009 German micro-census is a 1%-sample of all households in Germany; continuous mandatory household sample survey
Greece	The sample of households for the ICT survey of the year 2011 was consisted of a sub-sample of the sample used in EU-SILC survey of the years 2009-2010. 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 2009-2010. The secondary sampling units are the sampling households of the EU-SILC survey containing individuals aged 16-74 years old and belonging to the selected primary sampling units. The final sampling unit is one person randomly selected among the household members aged sixteen to seventy four years. 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.
Hungary	In 2011 the sampling frame comes from the Register of Dwellings which is based on the 2001 census and is regularly updated by new and demolished dwellings.
	The Register of Dwellings is regularly updated. Information on the amount of under-coverage will be available after the census 2011 only. A known shortcoming is that there is no information on the occupancy of new dwellings which may cause a bit higher non-response due to empty dwellings.
Ireland	Census of population
Italy	Public register of households (name of the components of the household, address of the household) There is a lag between last update of the public register of households and the moment of the sampling.

Latvia	Two sampling frames are built for each sampling stage due to two-stage sampling is used. At the first stage the list of Population census 2000 counting areas, which has some changes in 2009, is used as sampling frame. All territory of Latvia was divided in small territories (smaller than NUTS4) during the Population census 2000. The list contains information about the number of households in each counting area. At the second stage sampling frame is built from The Population Register, statistical register of dwellings and statistical register of households. The number of households in counting areas is updated every month. The time lag between last update of the list and the moment of actual sampling was one month. Second stage sampling frame was built using the copy of Population register given at the end of year 2010. Both statistical register of dwellings and statistical register of households was also updated using the Population register. So the time lag between last update of the registers and the moment of actual sampling was 1 month.
Lithuania	The Residents' Register. The Residents' Register Service under the Ministry of the Interior is the manager of the Residents' Register. Data of the Residents' Register in on-line mode are submitted to Statistics Lithuania. The Residents' Register database includes data on the residents of the Republic of Lithuania: the citizens of Lithuania, the citizens of foreign countries or persons without citizenship, declaring the place of residence in Lithuania or registering any changes of the civil state in a registry office. Due to the absence of a sampling frame for selecting households, a sample of individuals aged 16 and older from the Residents' Register is drown. Households whose members are the selected persons are surveyed. The Residents' Register is updated regularly. All persons are obliged to declare their place of residence, i.e. to submit data on the address of the place of residence to an institution responsible for the declaration of the place of residence. However, not all movements of the population within the country are reflected, whereas not all persons report about changing the address to a respective institution or the declared place of residence isn't the main place of residence. Consequently, the households, living at the selected person's address, are surveyed.
Luxembourg	Worked with RLD (random last digit) by using computer generated national fixed phone numbers randomly chosen to be called. A last birthday rule was applied. In order to reach representative results, a quota system for the individuals to be interviewed was applied. Households without fixed phone line and households only using mobile phones were not called.
Malta	The sample was derived from a database of persons and households held by NSO. This database is based on a register derived from the 2005 Census of population census and is updated regularly using auxiliary information, including Electoral Register and other auxiliary information. The NSO makes regular updates on its household and persons database which in turn is used as sampling frame for this survey. Still there is a time lag between these updates and the time when actual sampling takes place – this may cause some shortcomings. Furthermore no amount of updating will ultimately eliminate all biases. In many cases these biases relate to under-coverage problems which are generally attributed to population sub-groups e.g. migrants, newly-wed couples etc.
Netherlands	Sample frame: Municipal Population Register (GBA). This register contains all the persons who are registered in Dutch municipalities Advantage of using the Municipal Population Register is the actuality of information from the registered individuals. Information about birth, death, move (house), new addresses et cetera is coming in daily. Monthly the information is incorporated in the mentioned register. After the register is brought up to date in month t, individuals are selected for interviewing in month t+2
Poland	The basis for selection of Primary Sampling Units (PSU) and dwellings was the Domestic Territorial Division Register. The Register includes (among others) information about:

	- enumeration statistical district (ESD) - dwelling addresses The first stage sampling frame was a record of 29 thousand Area Survey Points (ASPs). They are enumeration statistical districts or sets of them which consist of at least 250 dwellings for urban area and 150 dwellings for rural area. The second stage sampling frame was based on the register of inhabited dwellings in the area survey points from the first stage. Sampling frame does not cover collective households like students' residences, hostels, social welfare homes, prisons, convents etc.
Portugal	The ICT survey sample is a sub-sample of the Master Sample (MS) - the sampling frame used by Statistics Portugal for household surveys. The MS was designed and selected using the information of the last Census of Population and Housing (Census 2001). It is constituted by private dwellings and it excludes collective households and institutions since they represent 1% of the total population residing in Portugal. The MS has almost 750000 private dwellings (535000 of which are as usual residence, the remaining are vacant, seasonal or for secondary use). The MS is a stratified one-stage cluster sample. In each stratum the clusters were selected systematically with probability proportional to size (number of private dwellings of usual residence). The stratification was done at NUTS III level and the clusters are geographical areas constituted by one or more contiguous statistical sections (census enumeration areas). Since the end of 2006 the MS is being updated. Each quarter a set of approximately 100 areas are updated in the field. The time lag between the Census and the selection of the sample leads to some out-of-scope units, i.e., dwellings that are not occupied (vacant, secondary use, business, etc.). Despite all the efforts to update the Master Sample there will always be cases where the interviews cannot be made.
Romania	The household surveys carried out by NSI-Romania are based on the use of a master sample EMZOT. EMZOT is a Multifunctional Sample of Territorial Areas, made by the data registered from Population and Dwelling Census in 2002. It is a data base including approximate 1.500.000 dwellings, selected according to probabilistic criteria, serving as sample base for all household surveys, in 2004-2013. The sample of 780 centres of research are distributed in the whole counties and in Bucharest: 427 in urban area and 353 in rural area.
Slovak Republic	Population and housing census is organized by SO SR in 10-year periodicity. The last one used for this survey was realized in May 2001. On the base of census results the information on whole population was accessible - household type, age, sex, education level, employment situation, location. Meanwhile SO SR organizes surveys on population increase and decrease by age and by sex. Census register from 2001 is used as sampling frame for survey on information and communication technologies (ICT) in households and individuals. Sampling frame for individuals was created on the state of population as of 31 December 2010. (The data for census organized by SO SR in May 2011 were not included in survey.) Negative side of the sample frame was the time difference between Household Register (May 2001) and state of population as of 31 December 2010. Missing social and demographic data (level of education, employment) were included into the state of population as of 31 December 2010 by statistical methods used in SO SR.
Slovenia	Units of observation are persons aged between 10 and 74 years at the time of interviewing and their households. The basis for the sampling frame is the Central Population Register (CRP). Target persons (persons aged between 10 and 74 years) are randomly selected from the register. For the allocation of persons, in addition to the CRP also the frame of the enumeration units is used, which serves as a sampling frame for the selection of units at the first stage in two-stage sampling designs.
Spain	The sampling frame is the Population Register managed by INE and updated in a continuous way.
Sweden	The sampling frame is the Total Population Register (TPR) kept by Statistics Sweden. The TPR obtains notifications of births, deaths, migrations etc daily from Tax Authorities.

	The quality of TPR is good. Births and deaths cause very small under- and overcoverage problems. Immigration causes some undercoverage because of the time lag between entry to Sweden and population registration. This undercoverage refers to immigrants with a non- Nordic citizenship. Emigration causes overcoverage because the population register is not always informed about departures.
United Kingdom	The Omnibus Survey uses the Postcode Address File of 'small users' as its sampling frame. Private household addresses in Great Britain are included in this frame. The sampling frame does not include addresses in the Isles of Scilly and the Scottish Highlands (north of the Caledonian Canal) and Scottish Islands in Great Britain. Northern Ireland has not been included in the sample frame for 2011, but estimates will be provided.
Iceland	The National Register of Persons is used as a sampling frame for the ICT survey. All Icelandic as well as foreign citizens domiciled in Iceland are registered in the National Register of Persons. Basic information such as gender, age and sample municipality are always found in this register. To be able to monitor every changes Statistics Iceland´s survey center has unlimited access to the National Register of persons during the data collection period. The sample is drawn out of a national register of individuals, but an equivalent register for households is not available.
Norway	The Statistical Population Register (SPR) is the central demographic population database at Statistics Norway. The database is updated from the administrative Central Population Register several times a month.
Croatia	Updated public telephone books were used this year, as the most up to date sampling frame. At the moment this is the best available "register" of households. Landline telephone penetration in Croatia is around 98%. Additional 2% were drafted from mobile operator registers since they do not own landline. In the remaining households, without landline or mobile phone, roughly 60% have no members in the eligible age range. Sampling frame does not cover 100% of eligible population, since very small portion of households which do not have landline or mobile phone, were not included in sampling frame. Since no other sources were found eligible to be used as an alternate sampling frame, we decided to use Croatian telephone directories only.
Republic of Macedonia	The frame for household's selection was the data from the Census of Population, Households and Dwellings, 2002. Time lag between last update of the sampling frame of households (Population Census, October 2002) and the moment of the actual sampling (January 2010) is about 7 years. Usually because of the time lag some households or persons from the subpopulations are not covered.
Turkey	Updated National Address Database linked with Address Based Population Registration System

Table: Overview of the target and non-target population (2011 survey)

Source: Quality reports for the Community survey on ICT usage in households and by individuals (2011)

	Individuals			Households		
	Target population	**Non-target population**	Target population as % of total population	**Target population**	**Non-target population**	Target population as % of total population
Belgium	8.020.000	2.940.000	73%	4.230.000	630.000	87%
Bulgaria	5.468.207	1.567.884	78%	2.540.341	253.191	91%
Czech Republic	8.220.000	2.303.000	78%	3.820.000	490.000	89%
Denmark	4.060.000	1.450.000	74%	2.400.000	400.000	86%
Germany	62.500.000	18.300.000	77%	35.400.000	4.300.000	89%
Estonia	1.017.000	323.200	76%	555.000	:	
Greece	8.277.574	2.636.945	76%	3.746.946	377.119	91%
Spain	34.576.941	3.929.036	90%	15.494.265	1.796.442	90%
France	45.670.000	17.950.000	72%	25.230.000	3.060.000	89%
Ireland	3.225.123	1.257.402	72%	1.512.931	116.747	93%
Italy	45.146.882	14.791.723	75%	21.598.222	2.866.610	88%
Cyprus	591.006	212.051	74%	263.967	17.033	94%
Latvia	1.733.541	496.100	78%	:	:	
Lithuania	2.553.632	775.407	77%	1.264.000	125.000	91%
Luxembourg	381.781	130.059	75%	156.451	15.484	91%
Hungary	7.605.000	2.248.000	77%	3.430.000	392.000	90%
Malta	312.830	104.778	75%	135.795	7.882	95%
Netherlands	13.100.000	3.400.000	79%	6.600.000	900.000	88%
Austria	6.400.000	2.000.000	76%	3.300.000	300.000	92%
Poland	28.503.984	9.726.096	75%	12.592.289	744.751	94%
Portugal	8.012.132	2.628.807	75%	3.630.326	392.873	90%
Romania	16.700.000	4.800.000	78%	7.400.000	:	
Slovenia	1.541.185	467.803	77%	646.867	:	
Slovak Republic	4.247.316	1.181.001	78%	1.897.007	79.042	96%
Finland	3.900.000	1.343.000	74%	2.200.000	250.000	90%
Sweden	7.000.000	1.480.000	83%	:	:	
United Kingdom	45.502.718	16.289.282	74%	23.661.291	3.210.817	88%
Iceland	223.492	89.677	71%	114.403	10.398	92%
Norway	3.582.581	1.337.724	73%	1.986.032	215.755	90%
Croatia	3.426.813	1.010.947	77%	1.451.730	25.647	98%
FYR of Macedonia	1.601.879	450.505	78%	564.238		
Turkey	48948935			18937559		

II.2.3. **Sampling design**

The population of interest of this survey, i.e. the group of statistical units around which it was intended to investigate is made up of households (with the exclusion of permanent members of cohabitation) and individuals living in Member States.

As the country details in the table below show, there are great differences among national statistical institutes as regards sampling design. Some countries use a sample based on individuals as primary sampling units. For other countries the primary sampling units are represented by households registered in the public register and in general the second sampling units are all the actual components belonging to each household included in the sample.

The survey should be based on a probability sample from which results representative of the population could be derived, considering the agreed breakdowns defined in the questionnaire.

The final sample units should be the individuals but each participating country should design its sample selection according to what is most efficient to that country.

The sampling design and the resulting sample size (see below) should be appropriate for obtaining accurate, reliable and representative results on the survey characteristics and breakdowns specified in the Regulation and the model questionnaire. Taking into account the unbalanced distribution of e.g. Internet users over the different groups of society, this condition can be difficult to hold for all indicators and all breakdowns, but this requirement should at least be followed for the main indicators described in the Benchmarking Digital Europe Framework 2011-2015 and Digital Agenda for Europe scoreboard. For further information, see

https://ec.europa.eu/digital-agenda/node/30065.

(Former indicators: the *e*Europe Benchmarking Indicators, listed in the Council Resolution of 18 February 2003 on the implementation of the eEurope 2005 Action Plan, Official Journal of 28/02/2003, C48, p. 2-9; from 2006 to 2010: the i2010 benchmarking indicators, see http://ec.europa.eu/information_society/eeurope/i2010/docs/benchmarking/060220_i2010_Benchmarking_Framework_final_nov_2006.doc).

This objective should be achieved for the overall proportions as well as for the proportions relating to the different subgroups of the population. In line with section II.2.7.3 below on accuracy measures and the template for reporting quality in annex II.3.3, the estimated standard error (standard deviation), expressed by the square root of the estimate of the sampling variance, shall not exceed 2 percentage points of the overall proportions and shall not exceed 5 percentage 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.

The aim of such requirement is to ensure the collection of a complete dataset – without empty, confidential or unreliable cells - for these indicators. In case this aim cannot be reached, clarifications should be provided (see also II.2.7.3).

Table: Overview of the sampling design and sampling methods (2011 survey)

Source: Quoted from the quality reports for the Community survey on ICT usage in households and by individuals (2011)

Austria

For the Continuous Labour Force Survey (LFS) every quarter of the year appr. 22,500 households are surveyed. Every household remains in the sample for five quarters (this means that 20% of the surveyed households are replaced every quarter).
For the ICT survey a gross sub-sample of appr. 5,700 households has been drawn. The LFS is based on "reference weeks". This means that the answers of the individuals in the household concerning the LFS module should refer to the situation in the reference week. The selected household of a certain reference week can be contacted up to three (five) weeks after the reference week. This means that if a household cannot be reached immediately after the reference week, the interviewers have more time to get the interview.
The households that were selected for the ICT survey were from the reference weeks 17 to 24 in the year 2011.
All household members between 16 and 74 in the sampled households are surveyed.

Belgium

The LFS sampling designs defines 12 strata (corresponding to NUTS2 units, with one exception: the Liège province split into two). We use a 2 stage sampling design. Within each stratum, all (private) households have the same selection probability, while the selection probability of an individual is
-zero if the individual is not in the 16-74 age class;
-inversely proportional to the number of persons 16-74 in the household s/he belongs to if the individual does belong to the 16-74 age class.
Initial LFS sample size: 14.619 households .
Individual within household: person aged 16-74 with the most recent birthday at the end of the Labour Force Survey reference week.

Bulgaria

Two-stage cluster sampling design is used for ICT usage in households and by individuals in Bulgaria. Sampling units as follows:
-on the first stage - the census enumeration units (PSU) are selected;
-on the second stage - the households are identified.
The sample is stratified by administrative-territorial districts in the country (NUTS3) and the household's location (urban\rural). As a result 56 strata are formed (28 of urban and 28 of rural population). Municipalities and settlements are ranged according to the number of their population within each stratum.
The number of census enumeration units (PSU) is calculated for each strata included in the sample.
The clusters on the first stage are chosen with a probability proportional to the population size (number of households) in the PSUs. Systematic sampling of the secondary units (households) selected in each primary unit is applied. Each PSU contains the same size - 6 households.

Cyprus

The sampling frame is stratified into urban and rural strata by district.
The size of the sample is predetermined on the basis of the average size of households having at least one individual in age group 16-74. The size of the sample should satisfactorily serve the desirable and acceptable confidence limits of the survey results, taking always into consideration time and cost constraints.
The households are then allocated in each stratum according to the number of households in the population census register.
The selection of the sample in the urban areas is done by using simple systematic random sample. A random start is selected, and by using the sampling interval N/n, the households for each district (urban areas) were selected.
The selection in rural areas is done two stages. The villages of each district are the Primary Sampling Units and the households the Ultimate Sampling Units.

All the villages of each district are listed in ascending order of size. The sample is drawn with Probability proportional to size (P.P.S.), the latter being determined from the number of individuals. Once a PPS sampling is employed the number of households in each selected village is fixed at 20. In cases of very small villages with less than 20 households, these villages are attached to other neighboring ones and they are considered as one village.
For the selection of the households in each selected village, simple systematic random sampling is employed in each village separately.

Czech Republic

The initial gross sample size was 6780 dwellings. Dwellings were selected using stratified two-stage sampling design. 1130 CEUs - census enumeration units were first sampled as primary sampling units with probability proportional to their size. The first stage sample is stratified by region (LAU1) and size of municipality to which the CEU belongs. In the second stage, 6 dwellings were sampled in each sampled CEU (simple random sampling without replacement).

All individuals aged 16+ in households living in selected dwellings are interviewed.

Germany

The ICT survey is based on a representative quota sample. The population of households (microcensus 2009) was divided into groups by the following characteristics:
•Federal State (16 Bundesländer);
•Household Type (4 types): 1) one person household; 2) two persons or more (>16 years) without children under 16; 3) one person (>16 years) with at least one child under 16; 4) two or more persons (>16 years) with at least one child under 16;

•Social status of head of the household (5 types): 1) self employed; 2) employee; 3) retired – (only for household type 1 and 2); 4) student (only for household type 1); 5) farmer (separate group regardless of household type and household net monthly income).
•Household net monthly income (4 classes): 1) < 1300 EUR; 2) 1300 to<2600 EUR; 3) 2600 to < 3600 EUR; 4) > 3600 EUR;
Sampling unit is the private household at main residence. Households are selected from address databases (i.e. from household budget survey).
In the first step, the households are contacted by mail and asked to send back their confirmation of participation. On that confirmation they are also asked to report the number of persons in the household and particularly the number of persons aged 10 years and over. Then the household questionnaire and the correct number of individual questionnaires are sent to the household. All persons aged 10 years and over living in the household are the selected respondents. There is no second step of sampling within the households.

Denmark

The sample is formed by a random selection from the CPR. When the sample has been drawn, the individuals selected are contacted by mail, in which they are asked to participate in the survey. If it is impossible to find a telephone number, the respondent is asked to contact Statistics Denmark with a number by which he or she can be reached.

Estonia

The sample design is stratified systematic sampling of individuals whose households are included into sample. Different inclusion probabilities of households are taken into account during calculation design weights.
Stratification is made by place of residence. The 15 counties of Estonia and Tallinn are divided into four strata according to the population size (I – Tallinn, II – four bigger counties, III – ten smaller counties, IV – Hiiu county) and different inclusion probabilities are used in stratas, the highest being for Hiiu county.
All household members aged 15-74 are interviewed for labour force survey and for ICT.

France

22479 addresses were selected in the local residence tax database, with a underlying stratification by department (NUTS3) and municipality.
Thanks to these addresses, 7580 phone numbers were selected in the official French telephone directory "Pages Blanches" for the telephone interview.
The other 14 899 addresses were selected for the internet and paper.

Greece

The sample of households for the ICT survey of the year 2011 was consisted of a sub-sample of the sample used in the EU-SILC survey of the years 2009-2010.
The multi-stage area sampling was adopted for the survey. The primary sampling units are the areas (one or more unified city blocks) participating the EU-SILC survey of the years 2009-2010. The secondary sampling units are the households of the EU-SILC survey 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 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 90 strata (stratum= Region x degree of urbanization) is equal to $f = \frac{1}{\lambda} = \frac{n}{N} \cong 1{,}73$ ‰, where n =6.500 is the total sample size of households and N =3.746.946 is the total number of households belonging to the target population, for the 2nd quarter of the year 2010. The quantity N is a projection, based on the time trend of the Labour Force Survey data of the period 2001-2010.

The number of the sampling households in each of the 90 strata (let h) was defined by applying the proportional allocation as follows:

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

Stages of probability sampling

The sample of households for the ICT survey of the year 2011 was selected of the sample used in the EU-SILC survey. Although the sample of households for the ICT survey of the year 2011 was selected of the sample used in the EU-SILC survey, 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:

1st 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:

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

where:

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

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:

$\frac{n_{hi}}{N_{hi}}$ (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 $\frac{1}{\lambda} \cong$ 1,7‰ in each stratum: $E\left(\frac{n_h}{N_h}\right) = \frac{1}{\lambda}$ (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 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 relation (3), it is deducted that: $\frac{1}{\lambda} = \frac{n_h}{N_h} \Rightarrow \lambda = \frac{N_h}{n_h}$ (7)

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

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

The initial selection probabilities of the sampling households of 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 EU-SILC and ICT do not coincide. The measures of 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 2011 was selected of the sample used in the National Survey of Income and Living Conditions (EU-SILC of the years 2009 and 2010), the following measures were taken for improving the representativeness:
a.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 2010
b.The 2nd stage selection probabilities of the households were modified taking into consideration the updated register of the households in the primary sampling units.
c.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 2010.
After the application of the above measures, the sampling households for the ICT have no the same probability of selections (1st and 2nd) with the sampling households of the EU-SILC survey, after changing the selection probabilities of the EU-SILC households, in order the probabilities of ICT households to be determined on the updated target population.

Spain

TYPE OF SAMPLING
An independent sample is designed in each autonomous community (NUTS 2) to represent it. The sampling type used is a three-staged stratified sampling. Primary sampling units (PSU) are census sections (geographical areas). Secondary sampling units (SSU) are the main family households and tertiary sampling units (TSU) are people aged 16 or over. The stratification criterion used for PSU's is the size of the municipalities to which the section belongs. There are a maximun of six strata in each NUTS2.

SAMPLE ALLOCATION
The sample was distributed among autonomous communities using a compromise between uniform and proportional to size allocation. Allocation among strata is proportional to their size, but ensuring that the number of sections per stratum in each autonomous community is a multiple of four.

SELECTION OF THE SAMPLE
The selection of first-stage units in each stratum was made using probability proportional to the size of each section. In a second stage, the households were selected by means of systematic sampling with random start and equal selection probabilities for each household in the section; so this procedure leads us to obtain self-weighted samples of households in each stratum.
In a third-stage, and within each household, a person aged 16 or over was chosen with equal probability using the Kish random method. All children aged from 10 to 15 are interviewed.

Hungary

The ICT sample in 2005 was a sub-sample of the sample of the micro-census. The micro-census sample was designed to provide reliable estimates of the main demographic indicators for the 176 General Electoral Districts (GEDs) of the country. The GEDs were roughly of the same size, the average being 24,000 in terms of dwellings. Each GED has a 2 % sample of its own, resulting in a self-weighting 2 % overall sample of the country. Within each GED localities were stratified by size in terms of the number of dwellings. Some GEDs are towns or segments of major cities (type 1), other GEDs consist of a number of small localities (type 2) GEDs of type 1 have 2 % systematic sample of dwellings, those of type 2 have two-stage stratified samples of dwellings; the PSUs (primary sampling units) are localities, selected with pps.

When selecting the ICT sample from the micro-census sample, an effort was made to keep it close to a self-weighting sample with 10000 dwellings in 374 strata. In the selected dwellings each household and one individual per household with ICT-user(s) were observed.

According to the rotational design, in 2011, 6538 former households remained in the sample, roughly 2000 households were dropped out and 3462 new dwellings were selected, resulted in a sample of size 10000.

Ireland

For the QNHS, a two-stage sample design is used. This comprises a first stage sample of 2,600 blocks (or small areas) selected at county level to proportionately represent eight strata reflecting population density. Each block was selected to contain, on average, 75 dwellings and the sample of blocks is fixed for a period of about five years. In the second

stage of sample design, 15 households are surveyed in each block to give a total quarterly sample design of 39,000 households.
Households are asked to take part in the survey for five consecutive quarters and are then replaced by other households in the same block. Thus, one fifth of the households in the survey are replaced each quarter and the QNHS sample involves an overlap of 80% between consecutive quarters and 20% between the same quarter in consecutive years. The ICT household survey was conducted on a sub-sample (40%) of the main QNHS sample.

Italy

The survey is based on a two stage stratified sample. At first stage, about 800 Municipalities are selected, stratified by region and type of municipality (metropolitan areas and other municipalities by size). Households are second stage sampling units, randomly (systematically) selected from administrative lists of each sampled municipality. All members of selected households are interviewed

Latvia

Stratified two-stage sampling (systematic sampling with inclusion probabilities proportional to unit size at the first stage, simple random sampling at the second stage) was used.
The stratification was made depending on degree of urbanisation of area Riga (the capital city), eight largest cities, other cities and rural areas forms four strata. The code of administrative territories was used to stratify. The population census counting areas were used as primary sampling units (PSU) at the first stage. PSUs were selected by systematic sampling with inclusion probabilities proportional to population size (number of households) of PSUs.
Households were used as secondary sampling units (SSU). Simple random sampling was used to select SSUs in each sampled PSU.
The total sample size (number of households) was approximately defined depending on the resources available and desirable result of the survey. The initial sample size was proportional to population size in each stratum. The initial sample size was adjusted according to response rates in each stratum to get the final sample size in each stratum.

Lithuania

A stratified sample design with a simple random sample was used in strata. The entire Lithuanian territory was divided into 7 non-overlapping groups – strata. The 5 largest cities of Lithuania, small and medium towns and rural areas were divided into separate strata. From every stratum, a simple random sample of individuals aged 16 and older was drown using the Residents' Register. A household, which lives at the selected person's address, was surveyed. The sample size of households in every stratum is proportional to the number of the population aged 16 and over in them.

Luxembourg

The national sample was constructed in 2 stages: 1rst level: selection of households by RDD and last birthday rule. 2nd level: selection of respondent in the household; 1 individual per household is selected.
Quota sampling: sex by age, region, nationality and professional activity. In other words: there is a stratification on age (16-18 years, 19-24 years, 25-34 years, 35-49 years, 50-64 years and 65-74 years), sex by age (with the same age categories), sex (male, female), region (city, centre, south, north (including west), east), nationality (Luxembourgish, Portuguese, other) and professional activity (active, not active).
For the individuals a weighting procedure for these quota variables was applied. For the households a weighting procedure for "household composition" was applied.

Malta

A sample of 2,000 individuals aged 16 to 74 was selected using systematic random sampling from a population database held and maintained by NSO. In an attempt to minimise response burden, individuals who had been selected to participate in the ICT survey over the last three years were filtered out before sampling.

The Netherlands

Sampling method: For the ICT survey a stratified two-stage sample is used. In the first stage municipalities are selected. Large municipalities being selected with a probability of 1. All the other municipalities are selected at random, taking into account the size of the municipalities. In the second stage the number of individuals, as determined for every municipality, are selected. Individuals aged 12 to 74 year are selected

Poland

A two-stage sampling scheme was adopted with stratification on the first stage. ASPs (Area Survey Points) have been stratified within each of 16 voivodships in Poland by type of locality. Big cities usually constitute separate strata, whereas the strata in rural areas are composed of ASPs comprising neighbouring poviats in voivodship. There are 96 strata (of which 31 covering rural areas).
The number of the ASPs selected to the sample from a given stratum was proportional to the number of dwellings in that stratum. The probability of ASP selection was proportional to the estimated number of dwellings in the ASP.
Beginning from 2006 a scheme of rotation has been applied in sampling design in order to improve comparability of data over time. Six dwellings from each of 675 ASPs selected for 2010 survey round constituted a half of sample for 2011 survey, which gave 4050 dwellings.
In order to chose the second half of the sample 675 "new" ASP were selected by means of procedure of systematic sampling after random ordering of the units (ASPs selected with PPS using Hartley-Rao method).
On the second stage 6 dwellings were selected from each "new" ASP by means of simple random sampling method.
Gross sample size for ICT usage survey 2011 counts: (6 x 675 "old" ASP) + (6 x 675 "new" ASP) = 8100 dwellings.

Portugal

The ICT survey is representative for Portugal (Mainland), Azores and Madeira. For this propose, 539 geographical areas were selected from the MS. Depending on the region a fixed number of dwellings were selected in each area.
A rotation system comprising four waves is being used. Dwellings are kept in the sample for four consecutive years before being replaced by an identical number of dwellings in the same area. One-fourth of the sample is replaced each year.
The original sample size for this survey (12460 dwellings) was calculated in 2009 by the time of a methodological redefinition. However, according to information collected in 2009 and 2010, dwellings that are out-of-scope, i.e. not occupied (vacant, secondary use, business, etc.) or have refused to answer the survey two consecutive years, were not followed in 2011. Thereafter, the sample size in 2011 comprises 10541 dwellings.
A sub-sample of dwellings was selected for CATI with the size of 4079 dwellings. This sub-sample is made of dwellings that had, at least once, being surveyed by CAPI and have given a telephone for contact. The goal is to facilitate the gradual transition to telephone interview in order to decrease the number of face-to-face interviews.
As a dwelling sample, households and individuals are listed at the time of the interview. In the case of more than one household in the dwelling only the main household is considered (should be living at the dwelling at the reference period – at the moment of the interview and at 31st March - and should have at least one individual between 16 and 74 years old).
After listing all the eligible individuals in the main household the Kish method is used to select the respondents: one individual aged between 16 to 74 years, if there is any.
Proxy interviews are not allowed. When the interview to the selected respondent is not possible it is assumed as a (total) non-response.

Romania

TIC sampling design is founded on a two-stage sampling technique.
In the first stage, a stratified random sample of 780 areas, Primary Sampling Units (PSU's), was designed after the 2002 census, using as stratification criteria the residence area and county and selected based on a systematic selection algorithm. The primary sampling unit, corresponding to the selection of the master sample, is a group of census section. The including probabilities from the first stage were proportional calculated with the size, expressed in number of permanent dwellings.
In the second stage, dwellings are systematically selected from the initial sample of PSU's: 12 from urban PSU's and 14 from rural PSU's. The secondary (ultimate) sampling unit, corresponding to the selection of the survey sample, has been the dwellings. All households within each sampling unit are included.
Stratification concerns only the first stage sampling. There are 88 strata and the criteria used being the area where a certain PSU is located (urban or rural area) and county (NUTS 3 level).

Slovenia

Two-stage, stratified sample design was employed in the survey. The sample is implicitly stratified according to the size and type of settlements (6 classes) and according to the statistical region (12 regions).
At the first stage 313 sampling units are selected with the probability proportional to size (PPS with replacement), and at the second stage 8 persons aged between 10 and 74 years are selected. Thus 2504 persons were selected. The person who is supposed to be interviewed is defined by the selected person.
The probability of selection of a person is calculated by the formula: $\frac{N_a}{\sum_{Slo} N_a} \cdot \frac{8}{N_a}$, where

N_a is the number of persons aged between 10 and 74 years in the sampling unit.
The selected person in the household is interviewed.

Slovak Republic

Surveyed households were selected so that they were representative for each of 8 regions of SR (NUTS3). There were from 463 to 668 households per region depending on the total number of households in the particular region. The sample was created individually for each region in two stages. Firstly, the municipalities representing five size categories were selected (on the base of population number). Secondly, in these municipalities there were selected randomly the households representing five household categories (on the base of their members number). The final sample size was 4500 households. One individual in each household was selected by lot to answer the questions in the second part of questionnaire. This selection took into account the prescribed quotas for the individuals regarding their sex, age category, education and social activity. In case there were not an eligible individual in the household another household with the same number of members were selected in the particular municipality. The quotas were defined according to the official population number as of 31 December 2010.
Survey was cover all required age categories, i.e. individuals of the age from 16 to 74 (sample size 4 000) as well as the marginal categories from 12 to 15, 75 and more (sample size 500). In case of the category from 12 to 15 the representative less were ensured, the results for 75 and more category were only informative regarding the last experience.
The interviewer visited the households and realised the face-to-face survey.
The breakdown of households was following:

•1 adult without children	– a1	= 997
•1 adult and children	– a1_dch	= 85
•2 adult without children	– a2	= 861
•2 adult and children	– a2_dch	= 885
•3 and more adult without children	– a_ge3	= 1 190
•3 and more adult and children	– a_ge3_dch	= 562
	TOTAL	= 4 500

The households are selected inside of each strata applying random sampling.

Finland

Systematic random sampling is used to extract the sample from the updated Central Population Register. The sorting system of the sampling frame is based on geographical population density. The target area is the whole country, and the respondents represent 16-74 years old population according to age, gender, province, and native language. Sample size is based on the knowledge of response rates in respective surveys earlier.

Sweden

Until 2010, the whole sample has been drawn from the Labour Force Survey (LFS survey). The LFS sample has, in turn, been drawn from the Total Population Register (TPR) kept by Statistics Sweden. The ICT survey has been conducted through telephone calls at the end of the LFS interview, i.e. after the LFS questions. However, in 2012, the aim is to completely separate the surveys and make separate phone calls for the ICT survey. The main reason for this is to improve the quality. Experiences from other telephone surveys indicate that a separation increases the response rate. A separation also enables us to use a smaller number of interviewers that are
experts in the field. We have provided them with an introductory course regarding concepts used in the survey. We therefore expect the quality of the responses to be higher. One further advantage of a separation is that it enables to adjust the methodology to the survey in terms of for example stratification.

In order to be able to link the results and minimize negative effects of potential breaks in times series, in 2011, we have collected data in both ways. Therefore 4 200 units were asked the ICT questions at the end of the LFS interview like in previous years (except retired people of 65 years of age or older that get only the ICT questions) and 1 750 units in separate phone calls.

Sample and Frame
Two different samples were drawn, one from the LFS survey like in previous years and one directly from the TPR. The TPR obtains notifications of births, deaths, migrations etc. daily from the Tax Authorities.

The LFS-sample is drawn at the end of the last quarter every year to cover the coming year's need of sample persons. When the sample is drawn it is stratified according to county, sex, and age. In this way 144 strata are constructed from

each stratum a simple random sample is drawn. The sample size is approximately proportional to the size of the stratum.

The LFS sample consists of three separate samples, one for each month in the quarter. Each of the samples, which consists of about 29 500 people, is rotated in such a way that an eighth leaves the sample between two survey cycles. For every sample this occurs every three month. Persons in sample are interviewed once a quarter with a total of eight interviews during a two-year period, after which they leave the sample.

The separated ICT-sample was drawn from the TPR at the beginning of March 2011 and it was stratified according to sex and age. In this way 12 strata were constructed and from each stratum a systematic random sample was drawn. The sample size is directly proportional to the size of the stratum.

United Kingdom

A sample of 67 postal sectors were selected and stratified by: region, the proportion of households renting from local authorities; and the proportion in which the household reference person is in Socio-Economic Group 1-5 or 13 (i.e. a professional, employer or manager). The postal sectors were selected with probability proportionate to size and, within each sector, 30 addresses (delivery points) were selected randomly. If addresses contained more than one household, the interviewer used a standard ONS procedure to randomly select just one household. Within households with more than one adult member, just one person aged 16 or over was selected with the use of random number tables. The interviewers endeavoured to interview that person – proxy interviews were not taken.
At the time of sampling, the age of the respondent is not known, nor the ages of the other members of the household. Therefore, the gross sample size in 6.3 includes ineligible units.

Iceland

Random sampling without replacement is used to select individuals. 2100 individuals aged 16-74 are randomly selected. Each individual then represents his household.

Norway

The individuals are randomly selected from SPR – the statistical demographic population database, classified by municipality, household and individual.
Statistics Norway draws a master sample of between 4500 – 5000 individuals where all age groups are represented. This sample also includes household members. Then the individuals' age by 31 December 2010 are calculated.
The final sample consists of 2000 individuals drawn from the master sample. In some cases the final sample are adjusted due to gender and age groups, in order to be representative for the Norwegian population.

Croatia

Total size of the sample was 6500 households. Sampling design deployed was the two-stage, stratified sample. Two-stage stratification was used by following two variables - region and size of settlements. Variable region consists of 6 geographical areas where each of them is defined by a group of existing counties (basic administrative unit in Croatia, whose territorial division and organisation is defined formally by law, in total there is 21 county).
Variable size of settlements consists of 4 size categories – up to 2.000 habitants, 2.001- 10.000 habitants, 10.001 – 100.000 habitants and 100.000 habitants or more. Allocation of the sample by strata was done proportional to size (number of households) of stratum.
At the first stage, households were randomly selected from the strata using sampling frame described in chapter 5. At the second stage, when the household is contacted, individual (aged between 16 and 74 years) as a final unit of selection was randomly selected following the procedure described in TCB ("Troldahl-Carter-Bryant") method of random selection. Every individual was from a different household.

Republic of Macedonia

The sample of households is two stage stratified cluster random sample. Stratification is on the 8 NUTS-3 regions and both rural and urban settlements, which means there are 16 stratums. Allocation of clusters is proportionally of population in the stratums. It is chosen 285 clusters and 7 households from each one; in total it is 1995 households. The mode of choice: the clusters are random chosen in the 16 strata, and from them households are randomly picked. A person from the household, aged 15-74, who was first to have a birthday, was interviewed.

Turkey

COVERAGE:

All households within the boundaries of Turkey are covered except the institutional population (population living in dormitories, hospitals, jails, rest homes and soldiers).

ESTIMATION LEVEL:
Estimation level of the survey is "Total of Turkey", "Total of Urban-Rural of Turkey" and "Nuts1" levels separately.

SAMPLING METHOD:
The sampling method of the survey is two stage stratified cluster sampling.

First stage; All areas in the frame are divided into clusters (blocks) containing approximately 100 households. Blocks are selected by systematically. However villages not having municipalities are assumed as a single block without considering the village size. Selection of such blocks (villages) is based on probability proportional to size (PPS).

Second stage; Addresses are determined from selected clusters (blocks) by using systematic selection method.

Primary Sampling Units:

533 selected blocks from rural areas
382 selected blocks from urban areas
Total : 915 selected blocks

Final Sampling Units:

3820 selected households (10 households from each clusters) for rural areas.
8528 households (16 households from each clusters) for urban areas.
Total : 12348 households.

SAMPLING FRAME :

Updated National Address Database linked with Address Based Population Registration System.
STRATIFICATION CRITERION
URBAN - RURAL : All the residential places having more than 20001 population are defined as
NUTS1 : Nomenclature of Territorial Units for Statistics

II.2.3.1 STRATIFICATION

The recommendation is to use a stratified sample of individuals or households with the aim to form groups (or layers) of units characterised, in relation to the variables subject of the survey, by maximum homogeneity within the layers and maximum heterogeneity between the layers. Achieving this goal in statistical terms means precision of estimates, or a reduction in sampling errors on a part with the sample quantity.

Each country should use the stratification variables according to what is most efficient to that country with particular attention to the demographic size of the localities.

II.2.3.2 SAMPLE SIZE

Calculation of sample sizes should take into account that this is a survey with multiple objectives. It has to ensure representative results for all the estimates produced. In particular calculation of sample size should take into account that each statistics have to be tabulated by age, sex, education level, employment situation, geographical location and type of locality.

As budgets are limited, the design of study involves making trade-offs along various dimensions. Larger samples make it possible to analyse sub-groups in depth but every interview increases survey costs.

On the basis of the previous considerations, it is suggested to adopt a mixed view, based on both cost and organisational criteria and on an evaluation of the sample errors of the main estimates on a national level and with reference to each of the territorial domains and to each of the breakdown variables of interest.

The calculation of sample sizes should be based on precision requirements (see above). On this basis countries should decide on sample design and calculate the sample sizes in order to receive estimates with sufficient quality and within possible budgetary constraints.

The table below gives an overview of the sample sizes reported by the national statistical institutes in the Final Reports i.e. Quality Reports for the 2011 surveys. Note that, due to national programs (e.g. production of regional estimates) and different units (household versus individual), the countries can't always be compared. For most EU countries, the final or net sample size was between 3000 and 6000 elements. The response and non-response as well as the quality (in terms of standard error) are discussed in more detail in the next chapters.

Table: Overview of the sample sizes and response rates (2011 survey)

Source: Quality reports for the Community survey on ICT usage in households and by individuals (2011)

	Unit	Gross sample size	Ineligible: out-of-scope	Other ineligible	Number of eligible elements	Non-contact	Refusal	Inability to respond	Rejected interviews	Other non-response	Net sample size	Unit response rate
Belgium	HH	14.619	233	456	13.930	1.390	5.057	53	325	1.116	5.989	43,0%
Bulgaria	IND	13.230	0	87	13.101	638	695	16	12		11.740	90,0%
Czech Republic	HH	6.847	552	643	5.652	104	1.312	46	:	:	4.190	74,1%
Denmark	IND	6.395	0	131	6.264	1.371	683	66	0	0	4.144	66,2%
Germany	HH	12.000	570	:	11.340	:	:	:	:	:	10.767	94,9%
Estonia	IND	4.373			4.373	29	69				4.275	97,8%
Greece	IND	6.500	20		6.480	1.325	352				4.803	74,1%
Spain	IND	24.972	2.185	2.854	20.944	2.591	2.349	197	6		15.801	75,4%
France	HH	22.479	1098	3210	18271	9214	613	80	276	0	5.080	27,8%
Ireland	IND	8.721	0	380	8.341	0	1.597	0	852	0	5892	70,6%
Cyprus	IND	4.858	1.032	160	3.666	9	5	0	0	0	3.652	99,6%
Italy	IND	52.349	6.964	1.463	45.385	3.845	3.344			1.772	36.423	80,3%
Latvia	HH	5.498	77	182	5.239	884	328	5	0	145	3.877	74,0%
Lithuania	HH	7.000	572	211	6.708	563	495	10	0	90	5.550	82,7%
Luxembourg	IND	20.999	2.488	1.340	17.171	7.472	7.846	181	0	153	1.519	8,8%
Hungary	HH	10.041	848	612	8.581	813	654	40	0	35	7.039	82,0%
Malta	IND	2.000	18	22	1.960	425	105	0	0	0	1.430	73,0%
Netherlands	IND	10.419	:	3.288	7.131	247	1.061	187	0	1.497	4.139	58,0%
Austria	IND	11.278	0	0	11.278	0	6.302	18	0	0	4.960	44,0%
Poland	IND	16.806	4.607	0	12.199	321	420	20	0	96	11.342	93,0%
Portugal	IND	10.541	948	654	8.939	1.116	156	138	0	354	7.175	80,3%
Romania	HH	10.095	1.140	413	8.542	246	190			189	7.917	92,7%
Slovenia	IND	2.326	8	9	2.309	248	344	23	0	13	1.681	72,8%
Slovak Republic	IND	4.171	0	0	4.171	0	0	0	0	0	4.171	100,0%
Finland	IND	4.300	0	8	4.292	234	458	104	0	681	2.815	65,6%
Sweden	IND	5.963	0	33	5.930	924	2.010	126	0	135	2.735	46,1%
United Kingdom	IND	5.608		550	5.058	445	1.604				3.009	59,5%
Iceland	IND	2.100	0	10	2.090	306	138	32	0	0	1.614	76,9%
Norway	IND	1.927		13	1.914	385	369	69		1	1.090	56,9%
Croatia	HH	6.500	410	81	6.009	922	1.910	56	0	95	3.026	50,4%
FYR of Macedonia	HH	1.995	4	107	1.884	245	61	7		23	1.548	82,0%
Turkey	HH	12.348	391	843	11.114	746	22	24		86	10.236	92,1%

Regarding the 2011 survey and EU-27, the net sample sizes of households and individuals according to Eurostat's database is 149331 households and 214580 individuals.

II.2.3.3 WEIGHTING – GROSSING UP METHODS

The weighting factors are to be calculated taking into account in particular the probability of selection and external data relating to the distribution of the population being surveyed, where such external data are held to be sufficiently reliable.

As the sampling design used differs strongly across countries, it is difficult to present 'fit-all' guidelines. Moreover, the weighting procedures / grossing up methods are usually determined by the sampling design used. The discussion is more of a theoretical nature and goes beyond the scope of this manual.

Where more advanced methods for dealing with unit non-response are not feasible, it is advised to correct for unit non-response by adjusting the grossing up weights. Ideally, auxiliary information such as socio-economic differences should be taken into account.

II.2.4. **Survey type**

Data collection method

Face-to-face interviews, telephone interviews and postal surveys are all possible techniques of collecting data. The socio-demographic characteristics which can be found in registers need not be collected in the survey.

Face-to-face interviews are recommended especially where telephone penetration is low. Interviews face-to-face or by telephone are preferred to postal surveys because interaction between the respondent and the interviewer can be important for the understanding of the questions and the answers but care should be taken to ensure a low rate of proxy response. When using telephone interviews it is important to include mobile and ex-directory users as well as fixed link users in the population sampled. Postal surveys might be used to lower survey costs when budgets are limited.

The table below indicates that most countries collect the data via a face-to-face interview, although telephone interviewing is also a common technique. Postal surveys and web surveys are used by only few countries.

Independent versus embedded survey

For practical reasons, an important number of countries have embedded the ICT usage survey into an existing social survey. The main advantages of such approach are a cost-reduction (e.g. travel time for interviewers) and the fact that certain variables need to be collected only once (this is especially the case for background characteristics such as educational level or employment situation). Further, linking the survey to an existing survey may give additional analytical possibilities (e.g. when embedded in the EU-SILC, the ICT usage pattern can be linked to the individual's or the household's living conditions).

However, although the filtering and routing limits the length of the ICT usage survey, it can be very burdensome for the respondents to be expected to give 50 yes/no answers following an already long interview for the survey vehicle, e.g. the Labour Force Survey. As the ICT usage questions will usually be inserted after the questions of the main survey, this can have unwanted negative effects on the response rates and the reliability of the answers.

The overview table indicates that a majority of the countries has a separate survey, although an important number of countries link the ICT usage survey to an existing survey. In this case, the ICT usage survey is usually embedded in another social survey such as the Labour Force Survey or general-purpose surveys (omnibus, micro-census, etc.).

Table: Overview of the data collection methods (2011 survey)

Source: Metadata reports for the Community survey on ICT usage in households and by individuals (2011)

Belgium	Self-administered mail survey-via drop-off by LFS-interviewer-with possibility to answer via web-application
Bulgaria	Face-to-face interviews
Czech Republic	Face-to-face - CAPI
Denmark	Telephone interviews - CATI (52%) and web-survey (48%)
Germany	Self-administered mail survey
Estonia	Face-to-face - CAPI
Greece	Telephone interviews
Spain	Face-to-face CAPI and telephone interviews CATI
France	Telephone interview, self administered internet, CATI, PAPI/CAWI and paper survey
Ireland	Face-to-face - CAPI (Blaise)
Italy	Face-to-face, PAPI, based on self administrate questionnaire
Cyprus	Face-to-face - CAPI (Blaise)
Latvia	Face-to-face interviews and telephone interviews
Lithuania	Face-to-face
Luxembourg	Telephone interviews - CATI
Hungary	Face-to-face
Malta	Face-to-face interviews and telephone interviews
Netherlands	Telephone interviews - CATI
Austria	Telephone interviews - CATI (Blaise)
Poland	Face-to-face - CAPI
Portugal	Face-to-face - CAPI and telephone interviews - CATI
Romania	Face-to-face
Slovenia	Face-to-face (70%) and telephone interviews (30%)
Slovak Republic	Face-to-face
Finland	Telephone interviews - CATI
Sweden	Telephone interviews
United Kingdom	Face-to-face and telephone interviews
Iceland	Telephone interviews - CATI
Norway	Telephone interviews - CATI
Croatia	Telephone interviews - CATI
FYR of Macedonia	Face-to-face
Turkey	Computer aided personal interview-CAPI

Table: Overview of the survey vehicles (2011 survey)

Source: Quality reports for the Community survey on ICT usage in households and by individuals (2011)

Austria	Labour Force Survey
Belgium	Labour Force Survey
Bulgaria	Stand-alone survey
Cyprus	Stand-alone survey
Czech Republic	Labour Force Survey
Denmark	Embedded in other national survey
Estonia	Labour Force Survey
Finland	Stand-alone survey
France	Stand-alone survey
Germany	Stand-alone survey
Greece	Stand-alone survey
Hungary	Stand-alone survey
Ireland	Labour Force Survey (QNHS)
Italy	Embedded in multipurpose social survey
Latvia	Stand-alone survey
Lithuania	Stand-alone survey
Luxembourg	Stand-alone survey
Malta	Stand-alone survey
Netherlands	Stand-alone survey
Poland	Stand-alone survey
Portugal	Stand-alone survey
Romania	Stand-alone survey
Slovak Republic	Stand-alone survey
Slovenia	Stand-alone survey
Spain	Stand-alone survey
Sweden	Stand-alone and Labour Force Survey
United Kingdom	Embedded in Omnibus Survey
Iceland	Stand-alone survey
Norway	Embedded in Omnibus survey
Croatia	Stand-alone survey
FYR of Macedonia	Stand-alone survey
Turkey	Stand-alone survey

Mandatory versus voluntary survey

Voluntary surveys are usually cheaper, quicker and easier to manage. A mandatory survey implies to make several attempts to contact the respondent or to send several reminders. This process usually makes the collection period longer as one need to wait a long time for all responses. The advantage of a mandatory survey is that your response rate is much higher, reducing the risk of having serious non-response bias.

But a voluntary survey can settle this argument by increasing the sample size and sending reminder letters to the respondents.

In practice, the ICT usage surveys became mandatory as of 2006 in all EU countries following the legal acts (Regulation 808/2004 and Regulation 1099/2005).

II.2.5. **Questionnaire – Data collection tool**

To enhance the comparability across countries, Eurostat provides a model questionnaire to the NSIs. It is recommended to adopt the model questionnaire for the national survey. However, for practical reasons (for instance linked to the survey vehicle), this may not be appropriate. In this case, the statistical institute should nevertheless make sure that the routing and the filtering is followed.

Eurostat designs the model questionnaire in close cooperation with the national statistical institutes, the main users and in coordination with other organisations such as the OECD. The results of previous surveys are taken into account, for instance to assess the relevance of certain questions or items, or to assess whether a variable needs to be collected on an annual basis. For reasons linked to the annually renewed implementing measures for Regulation 808/2004, the model questionnaire needs to be finalised 9 months before the survey takes place. In practice, this means that the questionnaire reaches its final status in March-April of the year preceding (T-1) the reference year (T). The obvious disadvantage is that 'last-minute' policy needs can't be included or that the results of the T-2 survey can't be fully taken into account. The advantage is that this gives the NSIs the necessary time to translate the questionnaire and implement the survey.

The model questionnaire for the survey on ICT usage in households and by individuals can be found in Annex II.3.1.

For a more general discussion of questionnaire design, the reader can consult the *Australian Bureau of Statistics Forms Design Standards Manual*:

http://www.sch.abs.gov.au/SCH/A1610103.NSF/Survey+Design?OpenView

II.2.6. **Quality control systems**

Quality control systems are of course country-specific as most statistical institutes have standard procedures and guidelines for plausibility checks or logic tests of datasets.

Such controls can be executed on-line, at the moment of the data capture by the interviewer or the data entry in the statistical institute, or after the data entry process (a program checks the data and prints the errors to be checked or corrected). On-line tests have the advantage that the errors can be corrected immediately (a lot of errors will be simple data-entry errors, f.i. typing 17773 instead of 1773), a disadvantage is that one needs data-entry staff that is familiar with the concepts in the survey

(in case they have to decide if a certain value will be accepted or not). The latter case of course doesn't apply with CATI or CAPI where the interviewer immediately enters the data.

Below, some of the most common errors or problems are briefly discussed.

Measurement error

There are a number of sources of measurement error: survey instruments (questionnaire), the respondent, the information system, the mode of data collection, the interviewer. This paragraph focuses on the latter, other types can back below. The first and probably most important step in the statistical process is the data collection or field work. One can design excellent questionnaires, excellent capture tools, excellent imputation methods and excellent data analyses, but as with any process the adagium 'Garbage In Garbage Out' also applies to statistics. If the interviewer is not well-trained, one can't expect input data of a good quality. A more important problem is usually the reliability of the interviewers. The interviewers may have an interest in finishing the interviewers in the shortest possible time. This can lead to interviews where certain questions are skipped (e.g. the person looks poor, let's fill in he's unemployed) or interviews that never took place (e.g. the person is 70 years old, let's put down he never used a computer nor Internet). Therefore the sensibilation of interviewers (which can include threats of non-payment or exclusion for future interview rounds) and the follow-up is extremely important. The follow-up can be a check for coherence and consistency of the answers or follow-up phone calls to verify with the respondents whether the interview has actually taken place or whether the interview was of an acceptable quality (e.g. the respondent confirms that the interview only took 4 minutes, but on the questionnaire answers to 60 questions were recorded).

Invalid response

Relatively unimportant in the ICT usage survey as most answers are limited to Yes or No. However, it is possible that several items were ticked in questions where only one answer is expected.

Relationship error

Comparing the answers across the survey can reveal inconsistencies between the answers. The routing and filtering should normally guarantee that respondents are not presented with questions they can't answer. It is nevertheless possible that e.g. an individual aged 18 indicates higher educational level or that an unemployed person indicates 'place of work' as a location where the Internet has been accessed. In the first case, this is most probably an absolute error. In the latter case the combination is possible if the respondent lost his job only very recently (meaning he could have accessed the Internet at a place of work during the last 3 months).

Compulsory question left unanswered

Again, the routing and filtering should avoid this type of errors. CATI or CAPI programs usually improve the quality of the data capture in this respect. With traditional interviews or self-administered mail surveys, this error is more important.

Suspicious values

This can occur when the individual responses are compared with the average scores within a stratum and unexpected answers are found. In practice it will however be difficult to determine whether it concerns a mistake or an outlier.

In terms of quality of the survey *as such*, the methodology and outcomes of the survey can be benchmarked against other surveys:

Representativeness

It can be useful to do an *ex-post* check of the representativeness of the sample, e.g. does the sample have a representative age distribution, is there some variability in the occupational and educational codes?

Year-to-year comparison at aggregate level

Comparing the results for the current year with the previous survey can also reveal quality problems where the growth is outside the range of the expected growth (e.g. the proportion of individuals using

the Internet decreases …). In such cases, it is of course possible that the problem stems from the previous survey exercise …

For this purpose, it can be interesting to produce some simple tabulations of the survey results.

Coherence or consistency with other surveys

The results can be compared with results from related survey or studies. However, in case inconsistent results are observed, it is not always easy to identify which survey gave the 'wrong' results.

II.2.7. Data processing

This chapter mainly discusses the treatment of non-response and accuracy measures. Although the grossing-up methods can be considered as a part of the *data processing*, this topic is discussed above in section II.2.3.3.

II.2.7.1 MISCLASSIFICATION TREATMENT

Not applicable to the household survey.

II.2.7.2 NON-RESPONSE TREATMENT[28]

Introduction

An important source of non-sampling error in surveys is the effect of non-response on the survey results. Non-response can be defined as the failure to obtain complete measurements on the (eligible) survey sample. The extent of non-response varies from partial non-response (failure to answer just one or some questions) to total non-response.

The latter case occurs when the interviewer was either unable to contact the respondent, no member of the household was able to provide the information, the respondent refused to participate in the survey or not enough information was collected in the interview (i.e. the response is too incomplete to be useful). This type of non-response is called unit non-response (see II.2.7.2.1): the sample unit does not provide any of the data required by the survey. Unit

[28] References for this chapter:

- Government Statistical Service (1997), *Report of the Task Force on Imputation – June 1996*. Government Statistical Service Methodology Series No. 3. London: Office for National Statistics.
- Statistics Canada (2003), *Household Internet Use Survey, 2002 – Microdata User Guide*.
- European Commission (2003), *Household Budget Surveys in the EU: Methodology and recommendations for harmonisation 2003*. Luxembourg: Office for Official Publications of the European Communities.
- National Centre for Social Research (1999), *Item non-response* (different contributions). Survey Methods Newsletter Volume 10 No.2 1999.
- Adjusting for Missing Data, Mohadjer, L., Choudhry, H., Studies of welfare populations: Data Collection and research Issues, 2001.
- Introduction to Survey Quality, Biemer, P., Lyberg, L., Wiley series in survey methodology, 2003.
- Weighting for non-response, Lynn, P., Survey and Statistical Computing, 1996.
- Non-response in the Norwegian Business Tendency Survey, Wang, J., Statistics Norway, 2004.

non-response is generally handled by adjusting the weight of the households and/or individuals that responded to the survey to compensate for those that did not respond.

Partial non-response or item non-response (see II.2.7.2.2) occurs when the respondents did not answer all questions because they did not understand or misinterpreted a question, refused to answer a question or could not recall the requested information. Item non-response is generally dealt with by imputation.

Effect of non-response on the quality of the data

Non-response – unit as well as item non-response – can seriously affect the quality of the data collected in a survey. Firstly, the characteristics (or answering pattern) of the non-respondents can be different from those collected among the sample units who did provide eligible answers. If such difference is systematic, serious bias can be introduced in the survey results. Secondly, the reduction of the sample size (overall or for certain questions) will increase the variance of the estimates. Thirdly, non-response can have an impact on the total cost of a survey exercise. Not only because a larger initial sample may be necessary, but also because of higher unit costs of the last few percentages of respondents (due to multiple visits). Finally, non-response can be an indicator of poor overall quality of the survey and thus create an image or confidence problem.

Minimising non-response

As prevention is always better than cure, attention should be given to avoiding non-response rather than treating non-response. The number (and timing) of reminder letters or call backs, the length of the fieldwork period, the survey technique(s) used, the length of the survey (i.e. the response burden), the use and structure of advance letters, the dissemination of previous results or the mandatory nature of the survey can all have an impact on the number of non-contacts or refusals. The length, design and complexity of the questionnaire, the interviewer's training or the interviewer's style can have an impact on the item non-response.

As this issue is common to all surveys, it will not be discussed in detail in this manual. This chapter is mainly focussed on the treatment of non-response in the specific context of the ICT usage survey.

II.2.7.2.1. UNIT NON-RESPONSE

Introduction

Unit non-response is defined as households/persons that are included in the sample but that have not participated to the survey and for which information consequently is missing for all the questionnaire variables.

Types of non-respondents include:

- Non-contact
- Refusals
- Inability to respond
- Rejected interviews
- Ineligible: out-of-scope

•Other ineligible

•Other non-response

Unit non-response can introduce bias in the survey results especially in situations in which the non-responding units are not representative of those that responded. Non-response increases both the sampling error, by decreasing the sample size, and non-sampling errors.

An overview of the non-response patterns in the 2011 survey (based on the 2011 Final Quality Reports) can be found above (II.2.3.2, Sample size). It is clear that in many countries the unit non-response is important with rates of 25% and more, with *non-contacts* and *refusals* as the main types of non-response.

Weighting adjustment for unit non-response

The principal method for unit non-response adjustment is weighting. Most strategies for weighting for non-response involve dividing the respondents into a set of comprehensive and mutually exclusive groups, referred to as weighting classes. A weight is then applied to each class.

Weighting classes

In order to implement non-response adjustments, it is required to create weighting classes. It is desirable to divide the sample in "response homogeneity groups/classes". Within these classes the respond rates should be as homogeneous as possible, and the response rate should be different among the classes. Data used to form these classes must be available to both non-respondents and respondents. Usually it is possible to get information about demographical (age, gender, ethnicity), geographical (urban/rural, zip code) or socio-economical (employment, income) variables from administrative data.

More advanced methods for creating weighting classes are methods like classification based on a categorical search algorithm or a logistic regression model using auxiliary variables to estimate the probability of response.

Sample-Based Weighting Adjustment

In sample-based weighting adjustment the weight adjustment applied in each class, is equal to the reciprocal of the ratio of selected sample size to respondents within each of these classes (the inverse of the response rate within each class). This non-response adjustment factor should be multiplied with the initial base weight.

A simple example:

	Population (I)	Sample size (II)	Respondents (III)	Respondent with characteristic (IV)	Non-response adjustment Factor (V = II / III)	Initial Base Weight (I / II = VI)	Adjusted Base Weight (V*VI=VII)	Adjusted population estimate (=VIII)
Male	8 820 000	2 100	1 600	1 000	1.31	4 200	5 502	5 502 000
Female	9 020 000	2 200	1 750	1 200	1.26	4 100	5 166	6 199 200
Total	17 840 000	4 300	3 350	2 200				11 701 200

Alternative forms of sample-based weighting are that the weights are not inverse response rates, but estimated coefficients of a regression model (where survey response is the left-side variable). In this case, the weights are reciprocals of estimated response rate by the regression model.

Population-Based Weighting Adjustment

Population-based weighting adjustment requires population estimates and class membership of respondents. If there is no data available about the non-respondents, population-based adjustment still is possible since this uses external control counts for the population and not data from the sample. The method is used to correct simultaneously for both non-coverage and non-respondents. The method is used similar to the sample-based method.

In population-based adjustment (post-stratification adjustment) the classes are created based on variables, which are known both for respondents and for the population. Weights are then applied in proportion to the ratio of population to achieved sample, so that the sums of the adjusted weights are equal to population totals for certain classes of the population.

A two-step procedure of first adjusting for non-response (sample-based adjusting) and then adjusting to known population counts is a common method that is used. However, this procedure is the same as a population-based weighting adjustment if the weighting classes in the sample-based and the population-based weighting adjustment are equal.

If the strata used in the stratification are used as classes in the weighting adjustment, there is no need for the weighting adjustment. The adjusted weighting procedure is then equal to the final grossing up/weighting procedure.

II.2.7.2.2. ITEM NON-RESPONSE

Introduction

As already mentioned above, there are several reasons for the data being unavailable. These include the refusal to provide an answer, the inability to provide an answer, inadequate quality of the provided answer (e.g. implausible, incomplete, inconsistent with answers to other questions, etc.). It can be caused by either the respondent (e.g. refusal) or the interviewer (e.g. failure to record the answer adequately) but also by the survey design itself (e.g. ambiguous routing or filtering).

In case a particular questionnaire shows too many errors, or if too many data are missing, it can be assumed that the household/individual in question has not co-operated satisfactorily in the survey. Here, the best solution is probably to remove the household/individual from the database and adjust the weighting coefficients for the other households accordingly. In other words, sampling units with a very high item non-response can better be classified as total non-response or unit non-response (cf. the category rejected interviews in the reporting template in §2.8). It is however difficult to define a threshold as not all questions are equally important (i.e. having missing data on four crucial indicators can possibly be worse than missing data on eight questions of secondary importance).

In other cases where the household/individual has supplied high-quality information for most variables but for which data on other variables are missing, the missing data can be estimated by using appropriate imputation techniques. The imputed values are supplied in such a way as to

preserve the underlying structure of the data and to ensure that the resulting records will pass all required edits or plausibility checks. In other words, the objective is not to reproduce the true microdata value, but rather to establish internally consistent data records that yield good aggregate estimates.

Where, for a specific variable, the proportion of missing data in relation to the total number of households exceeds a certain threshold, it must be asked whether or not imputation is still appropriate as this variable can probably not be used for analytical purposes. High item non-response (e.g. more than 20% non-response) on a specific question across several countries, can be a good reason to consider dropping or modifying the question or variable.

It should be stressed that the detection and imputation of missing data is to be done by the national statistical offices as Eurostat only receives aggregated tables which do not allow imputations.

Why treating non-response in the survey on ICT usage by households and individuals?

Most methods of compensating for missing items implicitly or explicitly make the assumption that data are missing at random. That is, the probability of an item being missing does not depend of the value of the missing item.

In our survey on households' usage of information and communication technologies there are without any doubt systematic patterns in the occurrence of non-response. It is obvious that non-response may be higher among older respondents or lower educated respondents as they are more at risk of not understanding the questions. We can take this into account by imputing within strata or classes. But the risk of wrongly imputing the data of ICT users (who feel concerned and "happily" answer the questions) to non ICT users (who drop out because they consider themselves not concerned by the survey) remains when it is the research variable itself (e.g. internet use) which may be the critical factor for the willingness or ability to provide an answer.

The logical solution to this problem would be not to impute at all. However not imputing does not exist as there is always an implicit imputation.

This is obviously the case where a respondent is dropped: the status moves from item non-response to unit non-response, which will normally be solved by changing the extrapolation weight of the other respondents (and this basically comes down to a mean imputation …).

Where the statistician decides to preserve the collected data instead of imputing, the user will in most cases do make an implicit imputation of the "non-stated" cases. Let's take the following numerical example ("did you use internet during the past 3 months"):

Answering category	Number of respondents (grossed up)	Percentage
Yes	924 000	21,0%
No	1 980 000	45,0%
Non-stated (missing)	1 496 000	34,0%
Total	**4 400 000**	**100,0%**

The proportion of persons using the internet can be calculated in two ways:

- Number of Yes answers compared to total population: 924 000 / 4 400 000 = 21,0%
- Number of Yes answers compared to total respondents (Yes + No): 924 000 / 2 904 000 = 31,8%

In a realistic setting, the end-user will only be interested in the proportion of users, not in the percentage of non-users and definitely not in the number of Non-stated. If these Non-stated are mentioned separately, they will most probably be overlooked by the user or interpreted in different ways by different users – by accident or on purpose (which is not a good thing in terms of transparency and relevance of the official statistics).

Both results above suffer from silent or implicit imputation. In the first case, the Non-stated cases are treated as 'No' answers. This bears the risk of overestimating the number of non-users: whether one uses the internet is in any case not the only factor affecting response rates in household surveys.

In the second case, there is a risk of overestimating the number of users: the underlying distribution of 'Yes' and 'No' answers in the 'Non-stated' group is implicitly supposed to be identical to the distribution in the group of respondents while the incidence of non-use is expected to be higher among the non-respondents.

The true value (or better: most suitable estimate) is somewhere in between, but unknown.

The easy to implement mean imputation (see below) will normally give exactly the same proportion of internet users as the case where we only take into account the respondents , but it should be remembered that this method is the 'worst imputation scenario' where no additional information is taken into account. More intelligent methods would give more accurate estimates. E.g. in a postal survey the respective response pattern of the consecutive reception waves can help fine-tuning the imputation: if response to our questionnaire is related to ICT use, the questionnaires received after a first or second reminder letter will show lower internet use figures. Such digressive pattern could be introduced in the imputation process if we assume that non-response is the extreme case of a late reply.

How to treat non-response in the survey on ICT usage by households and individuals?

The choice of the imputation method is at the discretion of the national statistical institutes, but a number of options are briefly discussed in this section. As experiences with the application of different methods are still missing, putting forward one particular method is not desirable at this stage. A more detailed discussion of these and other methods is beyond the scope of this manual, but can be found in the extensive literature on the issue of dealing with non-response.

Deductive methods

These methods are rather related to heuristics than to modelling. They try to deduct the most logical answer using the available information for the household or individual. In general, such procedures will be part of the validation checks and not of the non-response treatment.

Example: the respondent did not state whether s/he uses internet but from his other answers, we know s/he doesn't use a computer. In this case, the most obvious value to impute for internet use would be No. Wrong imputations will occur if this specific person only uses internet by means of a mobile phone, TV, etc.

Imputing the mean or mode

This method consists of imputing missing values by the mean observed in the group of respondents in case of numerical variables or the mode in case of categorical or binary variables.

Instead of imputing the overall mean or mode, usually, the imputation will be done taking into account some background characteristics of the household: before imputing, the respondents are grouped into different classes according to background variables such as age group, gender and educational level; the breakdown variables can however also include study variables (e.g. computer users versus non users). Within each group, the class mean or class mode is imputed to cases with missing values. The classes may be different for each variable to be imputed.

The big advantage of this method is that it is very easy to implement and to explain. The main drawback is that it may compress the distribution of the survey variables (as the less popular items will probably never serve as a donor).

Example: in the class 'age group 45-54' x 'lower educational level' x 'male', 20% of the respondents appear to use internet (while 80% don't). The most popular answer – in other words, the mode – is No. Therefore, all non-respondents will be imputed with a No value (which means approximately 20% of those may be misclassified – to avoid this problem one can try the next technique).

Hot deck imputation

For each respondent with a missing value for a specific variable, this value will be imputed with the corresponding value from the previous respondent in the database with a valid response on this variable. Usually, this method is applied within classes in order to improve the quality of the imputations.

After fixing a starting value for each item and each class, each case is processed sequentially. If the case has a missing item, this is replaced by the imputation value from the relevant class. If the item is not missing, it replaces the stored initial imputation value for its class, and can be used for imputation of subsequent missing items.

Closely related to this sequential hot deck imputation, is hierarchical hot deck imputation. Here, a lot more imputation classes can be considered as the boundaries of the classes are not fixed. When no suitable donor is found at the finest level of the classification, classes can be collapsed into broader groups until a donor is found. Taking into account the relatively small sample sizes used in our ICT survey, the hierarchical hot deck imputation may be less appropriate than the sequential variant (there are probably not enough cases available to fill the high number of classes).

These methods are relatively simple to execute, assign real/existing values to a non-respondent and better respect the underlying distribution than the mean or mode imputation. On the other hand, the algorithms may be more difficult to program (especially in the case of hierarchical hot deck) and there is a risk that one donor value is used several times (in case successive cases all show a missing value).

Nearest-neighbour imputation

This method relies on being able to identify the distance between any two units based on some suitable distance measure.

Regression imputation

Imputing by regression methods is usually carried out one variable at the time. The methods require that the values of one or more auxiliary variables are known for both the complete cases on which the variable of interest is recorded and for the missing cases. A regression model is fitted that relates the variable of interest to the set of auxiliary variables. For categorical variables, particularly binary variables, logistic regression is usually used. In this case, the imputed value will usually be the value with the higher predicted probability. For categorical variables with many categories, this method is less suited. For binary variables, more complex methods such as discriminant analysis could also be used.

Indicating an optimal imputation method is difficult. The best method will depend on the data available, the nature of the survey and of course the nature of the question (binary, numerical, etc.).

A general rule of thumb is however that the more simple methods such as mode imputation within classes or sequential hot deck do not necessary give poorer results than the advanced methods using modelling. This argument will even be stronger when making a cost/benefit analysis (in terms of processing and programming time).

II.2.7.3 COMPUTING ACCURACY MEASURES

The use of a sample survey unavoidably leads to sampling error. To get an idea of the sampling error, it is advised to compute this error for certain indicators and report in the quality report. Such information is not only interesting to decide whether the indicator is reliable enough to be published, but can also help to determine whether the sample size should be increased in a next edition of the survey for certain strata where a large sampling error was diagnosed.

The sampling error reflects the fact that only a particular sample was surveyed rather than the entire population.

As accuracy measure, the standard error (the square root of the variance) of the estimator for proportions is recommended (see also Chapter 9 of the *Reporting template*, Annex II.3.3).[29]

[29] As mentioned in the original 2006 version of the manual, the (estimated) relative standard error – or (estimated) coefficient of variation (CV) – is the ratio of the square root of the variance of the estimator for the proportion to the expected value of the proportion. It is estimated by the ratio of the square root of the estimate of the sampling variance to the estimated value. To avoid that the level of the estimate when using the coefficient of variation has an extreme impact on the accuracy measure (e.g. a proportion of 1% leads to a denominator of 0,01 when calculating the CV, in other words, the standard deviation is multiplied by 100), the standard deviation (i.e. the square root of the sampling variance) should be used as a reference measure.

The estimated standard error (standard deviation), expressed by the square root of the estimate of the sampling variance, shall not exceed 2 percentage points of the overall proportions and shall not exceed 5 percentage 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.

The estimation of the sampling variance should ideally take into account the sampling design (e.g. the stratification).

Some guidelines for interpretation and suggested use of the standard deviation can also be found in section II.2.3 above where the sampling design is discussed.

The computed accuracy measures will also determine the reliability of the estimates. If the standard error for a certain proportion or absolute value is too high, the cell needs to be suppressed.

In case the computation of the standard error for all cells is not feasible, the number of respondents underlying the estimate can be used to decide whether a proportion can be published (e.g. if based on less than 10 respondents, the estimate is not published).

Unreliable data for any variable should be reported in the quality report and notes accompanying the transmitted microdata files.

After the transmission of microdata, Eurostat is calculating standard errors and confidence intervals for the variables of the households' and individuals survey based on information provided by the NSIs in the microdata files. See annex II.3.2, Transmission Format, for further details on the required information.

The template for the 2014 metadata/quality reports was updated. It requests specific details regarding sampling and sampling error calculation, see annex II.3.3.

Parameters that affect the sampling error are the estimate i.e. the actual proportion and the sample size. Standard errors reported in the last surveys have shown the overall proportions, for example for all individuals, are relatively accurate. But the situation is different when looking at the accuracy for the breakdowns.

The breakdown between men and women doesn’t really affect the quality of the estimates, because of the absence of a strong relation between Internet use and gender and hence relatively equal sample sizes for the two groups.

For the breakdown by age group, the situation is more problematic. In the older age groups, the sample Internet users is becoming very small (because of the filter in the questionnaire), which leads to unreliable estimates in most cases. These conclusions can be taken into account when designing the sample, e.g. by including relatively more respondents in the sample for strata where poor quality in terms of accuracy can be expected (such overrepresentation should of course be compensated for in the weighting). Unfortunately, cost restrictions may not always allow for larger samples and reducing the sample size (and accuracy?) in other strata may not be the optimal trade-off either.

II.2.8. **Survey execution reports**

Drawing up reports after the execution of the survey is not only a tool for a self-assessment of the work carried out by the statistical office and a detection of possible points where there is space of improvement. Survey execution reports are also a tool for assessing the comparability from one year to another or from one country to another.

The methodological reports for the ICT usage surveys coordinated by Eurostat should cover the following subjects:

a. General methodological information: reference period(s); survey period; survey vehicle, where applicable; survey type; pre-tests; methodological differences compared to previous data collection exercises.

b. Statistical unit(s), scope and target population: statistical unit; age groups covered; territorial coverage; target population for households and for individuals; non-target population for households and for individuals.

c. Information on the national questionnaire.

d. Sampling frame: name and description of the sampling frame or register used; known shortcomings of the sampling frame, if any.

e. Sampling design: sampling method; additional measures taken at the time of sampling design to improve representativeness.

f. Unit non-response at household and at individual level, by giving information on: gross sample size; number of out-of-scope cases, number of other ineligible cases; number of eligible elements; number of non-contacts; number of refusals; number of cases unable to respond; number of rejected interviews; other non-response; net sample size; unit response rate. Additionally, the report shall discuss the methods used for minimizing non-response as well as the methods used for dealing with unit non-response.

g. Item non-response: variables or items with item response rates below 90%; methods used for dealing with item non-response.

h. Grossing-up procedures for households.

i. Grossing-up procedures for individuals.

j. Sampling error: standard errors for a selected group of indicators or sub-indicators listed in the template to be provided by the Commission; the calculation method for the standard error; comments on reliability, representativeness and completeness of the data.

k. Problems encountered and lessons to be learnt.

l. The report should include the questionnaire in national language and, if available, in English. The statistical unit is the base type of the elements of a group (also called population) that we want to observe or analyse. The basic statistical operations of classification, aggregation and ordering are done on the statistical unit.

In cases where the recommendations provided in this methodological Manual are not followed, the report should clearly describe deviations, the impact of those on comparability a justification for the deviation and – where possible – plans to comply with the recommendations.

The *Reporting template* used for the European ICT usage surveys, including some instructions, can be found in Annex II.3.3.

The template covers both the Metadata report (Interim Report) and the Quality report (Final Report) referred to in Article 7(4) of Regulation 808/2004 (see Annexes III.1 and III.2).

II.3. ANNEXES

II.3.1. Model questionnaire

See document ***MM2014_Annex II.3.1_Model_questionnaire HH 2014 v 3.5.doc***

II.3.2. Transmission format

See document ***MM2014_AnnexII.3.2_Transmission format HH 2014.zip***

The format for microdata transmission has also been provided in the restricted Circabc domain.

Starting with the 2014 survey, NSIs only transmit microdata (no selected aggregate data).

The planned aggregation of variables and breakdowns computed by Eurostat from microdata transmitted by NSIs is included in the above zip file.

II.3.3. Reporting template

Note: the reporting template for the interim (metadata) & final (quality) reports refers to the 2014 ICT household survey.

See document ***MM2014_AnnexII.3.3_Reporting Template HH 2014.doc***

Part III

EU Legislation

III.1. REGULATION (EC) NO 808/2004 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL OF 21 APRIL 2004 CONCERNING COMMUNITY STATISTICS ON THE INFORMATION SOCIETY

Official Journal of the European Union, 30.04.2004, L143.

⇨ See document [***MM2014_AnnexIII.1_Reg808-2004.pdf*** (7 pages)]

Versions in other languages can be found on the EU's legislation server

http://eur-lex.europa.eu/en/index.htm

1. click 'Official Journal'
2. select 'Year' (=2004) and 'OJ Number' (=143)
3. select language (upper right, only the 11 official languages before 01/05/2004)
4. scroll down in the table of contents to the link to page 49.

III.2. COMMISSION REGULATION (EC) NO 859/2013 OF 5 SEPTEMBER 2013 IMPLEMENTING REGULATION (EC) NO 808/2004 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL CONCERNING COMMUNITY STATISTICS ON THE INFORMATION SOCIETY

Official Journal of the European Union, 06.09.2013, L238.

⇨ See for versions in all languages

http://new.eur-lex.europa.eu/legal-content/EN/TXT/?qid=1389870503357&uri=CELEX:32013R0859

or:

1. click 'Official Journal'
2. select 'Year' (=2013) and 'OJ Number' (=238)
3. select language (upper right, all official languages excepting Maltese)
4. scroll down in the table of contents.