Methodological Manual for **statistics** on the **Information Society**

Survey year 2007, v2.1

EUROPEAN
COMMISSION

eurostat

Methodological manual for statistics on the Information Society
implementation of Art. 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)

Published by Eurostat

Version 2.1, May 2007

Contact

Part I: Fernando Reis (Eurostat, Unit F6),
+352 4301 36568
fernando.reis (AT) ec.europa.eu

Part II: Heidi Seybert (Eurostat, Unit F6),
+352 4301 37416
Heidi.Seybert@ec.europa.eu

Table of contents

Preface

Information and Communications Technologies (ICT) are still the main drivers of the economy and of changes within society. Official statistics are indispensable for an informed understanding of the implications of the transformations underway. However, the right 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.

From 2002 the European Commission established annual Information Society surveys to benchmark the ICT-driven development in enterprises and by individuals. Eurostat developed two model surveys, one on enterprises, one on individuals, in close collaboration with Member States and the OECD, which have been regularly adapted to the changing needs of users and policy makers. These model surveys are composed of model questionnaires and 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: it allows flexibility to adapt the surveys to newly evolving needs by users and decision makers. Annual implementing measures such as Commission Regulation (EC) No 1099/2005 for the survey year 2006 are forming the basis for the Eurostat model surveys and will ensure harmonized data for all EU-25 Member States until 2010.

The current version of the methodological manual is drawn up as a tool for helping national statistical institutes (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 2006. Colleagues from the NSI and from OECD contributed substantially and helped Eurostat to finish it. 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. Certain areas are covered at length, others would need some additional input, so it is far from being exhaustive. Nevertheless we went public to provide this tool to those who are struggling with data collection and interviewing for the 2006 survey wave. Any proposal to the authors to improve the manual is welcome, also from interested readers outside the European Community, in view of reaching for harmonized information society statistics at global level.

Hartmut Buchow

Information Society Statistics

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 element to assure 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 "Community Survey on ICT usage and e-Commerce in Enterprises" is the enterprise, as defined in the Council Regulation (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 easy to attribute to the different establishments of an enterprise, and for this reason the enterprise is the statistical unit adopted.

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 Group 72 of NACE, localised in the EU".

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 practise 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 only and every element of the target population only once. However, most of the times 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 2.2.

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

- **Economic activity:**

  Enterprises classified in the following categories of NACE-Rev.1:

  - Section D – "Manufacturing";

  - Section F – "Construction";

  - Section G – "Wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods";

- Groups 55.1 and 55.2 – "Hotels" and "Camping sites and other provision of short-stay accommodation";

- Section I – "Transport, storage and communication";

- Section K – "Real estate, renting and business activities";

- Groups 92.1 and 92.2 – "Motion picture and video activities" and "Radio and television activities".

Optionally the target population can be extended to the following categories:

- Section E – "Electricity, gas and water supply";

- Groups from 55.3 to 55.5 inclusive;

- Groups from 92.3 to 92.7 inclusive; and

- Division 93 – "Other service activities".

● **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.

● **Geographic scope:**

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

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.

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 1099/2005 (see also *Part III – EU Legislation*).

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 "Individual" an attribute is his "Monthly Salary". We want to measure this attribute and can just use a variable to hold the monetary value of the monthly salary. But collecting the monetary value of an individual's monthly salary might not be feasible. In this case we can use another variable which is feasible, like a categorical variable with 4

different ranges of monthly salary, which information individuals are willing to give. 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 contain non-numeric information. They serve merely as labels or names for identifying special attributes of the statistical unit. An example is the Gender of employees with the values Male or Female. Sometimes qualitative variables can be rendered numeric by coding the non-numeric values. An important type of qualitative variables is that of binary 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 Salary or Experience (in years) of an employee.

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

For the 2007 enterprise Community Survey the observation variables to be collected are the following.

Question number in model questionnaire	Variable
A1	Computer usage
A2	Percentage of persons employed using computers at least once a week
A3a	ICT usage: wireless LAN
A3b	ICT usage: wire based LAN
A3c	ICT usage: Intranet
A3d	ICT usage: Extranet
A4	Existence of IT system to manage the placing and/or receipt of orders
A5a	Orders management IT sys. linked to internal sys. for re-ordering supplies
A5b	Orders management IT system linked to invoicing and payment systems
A5c	Orders management IT sys. linked to system for managing production, logistics
A5d	Orders management IT system linked to suppliers' business systems
A5e	Orders management IT system linked to customers' business systems
A6	Use of ERP software application
A7a	Use of CRM software to store, process and analyse information about clients
A7b	Use of CRM software to do data-mining
A8	Free / Open Source adoption at operating systems level
A9	Use of e-invoices in sending
A10	Use of e-invoices in receiving
A11	Use of advanced e-signatures (digital signatures)
B1	Access to Internet
B2	Percentage of persons employed using computers connected to the WWW
B3a	Internet connection: Traditional modem
B3b	Internet connection: ISDN
B3c	Internet connection: DSL
B3d	Internet connection: Other fixed Internet connection
B3e	Internet connection: Mobile connection
B4	Maximum download speed of Internet connection
B5a	Internet usage for banking and financial services
B5b	Internet usage for training and education
B5c	Internet usage for market monitoring
B5d	Internet usage for receiving digital products
B5e	Internet usage for obtaining after-sales services
B6	Internet usage for interaction with public authorities

B7a	Internet usage for obtaining information from public authorities' websites
B7b	Internet usage for obtaining forms from public authorities' websites
B7c	Internet usage for returning filled in forms to public authorities
B7d	Internet usage for submitting a proposal in an electronic tender system
B8	Existence of own website
B9a	Website for marketing own products
B9b	Website for facilitating access to catalogues and price lists
B9c	Website for providing after-sales support
C1	Have ordered via Internet
C2	Percentage of purchases resulting from orders placed via Internet
C3	Have received orders via Internet
C4	Percentage of turnover resulted from orders received via Internet
C5	Use of secure protocols (SSL/TLS) in receiving Internet orders
C6a	Internet e-sales by type of client: B2B and B2G
C6b	Internet e-sales by type of client: B2C
D1	Have ordered via non-Internet computer networks
D2	Percentage of purchases from orders placed via non-Internet computer networks
D3	Have received orders via non-Internet computer networks
D4	Percentage of turnover from orders received via non-Internet computer networks
E1	ICT specialist employment
E2	ICT specialists percentage or number
E3	ICT specialists recruitment
E4	ICT specialists hard-to-fill vacancies
E5a	ICT specialists hard-to-fill vacancies by reason: lack or low number
E5b	ICT specialists hard-to-fill vacancies by reason: lack of qualifications
E5c	ICT specialists hard-to-fill vacancies by reason: lack of work experience
E5d	ICT specialists hard-to-fill vacancies by reason: salary requests too high
E5e	ICT specialists hard-to-fill vacancies by reason: other reason
E6	ICT users recruitment
E7	ICT users hard-to-fill vacancies because of lack of skills in the use of ICT
E8a	Training provided: ICT specialists
E8b	Training provided: ICT users
E9	ICT specialists functions by external suppliers
E10a	ICT specialists functions by external suppliers abroad: own affiliates
E10b	ICT specialists functions by external suppliers abroad: other enterprises
E11a	ICT specialists functions by external suppliers abroad: management
E11b	ICT specialists functions by external suppliers abroad: development
E11c	ICT specialists functions by external suppliers abroad: operations
E11d	ICT specialists functions by external suppliers abroad: other
E12a	ICT specialists engaged by region: Other EU Member-States
E12b	ICT specialists engaged by region: Non-EU Countries
E13	ICT user functions by external suppliers
E14a	ICT user functions by external suppliers abroad: own affiliates
E14b	ICT user functions by external suppliers abroad: other enterprises
E15a	ICT user functions by external suppliers abroad: sales, marketing, customer services
E15b	ICT user functions by external suppliers abroad: R&D, product design and engineering
E15c	ICT user functions by external suppliers abroad: other
E16a	ICT users engaged by region: Other EU Member-States
E16b	ICT users engaged by region: Non-EU Countries
X1	Main economic activity of the enterprise
X2	Average number of persons employed
X3	Purchases of goods and services
X4	Turnover
X5	Location (Objective 1/ non-Objective 1 region)

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) or for subpopulations defined conditional on the answers of other study variables (e.g. Internet users versus non users).
E.g.:
- number of enterprises with Internet access
- number of enterprises with between 10 and 49 persons employed with Internet access
- number of enterprises having a LAN and an Intranet

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 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 the aggregated data to Eurostat, can be found in I.3.2 *Transmission format*.

I.1.6. Explanatory notes

The explanatory notes in this chapter refer to the questions in the 2007 model questionnaire (see I.3.1 *Model questionnaire*, p. 172). 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 MODULE A: ACCESS TO INFORMATION AND COMMUNICATION TECHNOLOGIES

A1: Did your enterprise use computers, during January 2007?

[**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. Normally the user is placed in front of it, behind the keyboard with the monitor placed on top.

- **Portable computers (laptop):** A portable computer (or laptop) can be defined as a computer that is small enough to carry. A portable computer (or laptop) can usually carry out the same tasks as a desktop computer. Includes notebooks but not handheld computers, although portable.

- **Handheld computers (palmtop):** 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.

All the above types of computers are in the scope of this question. On the other hand, electronic tills which are only used for monetary transactions (EPOS) and computer controlled machinery are not considered for this question.

Use of computers doesn't refer to ownership. The computers might belong to the enterprise, be leased or shared with another organisation.

In general, throughout this manual **use** never refers to ownership. Therefore, enterprises which outsource ICT functions are also considered and should answer "yes" to these questions.

A2: ***How many persons employed used computers at least once a week, during January 2007?***
or
Please indicate an estimate of the percentage of the number of persons employed used computers at least once a week, during January 2007.

[<u>Scope</u>: enterprises which use computers, i.e. A1 = Yes]

[<u>Type</u>: numerical, in real or percentage values]

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. It intends also to capture only the relevant use of computers by workers, setting a minimum level of frequency: at least once a week. Although there was in the past reference to the use of computers by persons employed in their normal work routine, this concept shouldn't be used as it would, for instance, include those which routine includes using a computer every three months when this wouldn't be a relevant use of computers by the enterprise personnel.

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 **real** or in **percentage** values. The variable in real values might be more appropriate for smaller enterprises, while the one in percentage terms indicated 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 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 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 to use a combination of both giving to the respondent the opportunity to choose to answer in real or percentage values.

Independently on how this variable is collected, in real 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 real value needs to be calculated afterwards multiplying the answer with that background variable to gross up the number of persons employed using computers.

A3: Did your enterprise have the following information and communication technologies, during January 2007?

[<u>Scope</u>: enterprises which use computers, i.e. A1 = Yes]

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

All enterprises with internal computer-mediated network (LAN or set of enterprise's LANs) infrastructure in January 2006. This question refers to the availability regardless property and purposes of the internal computer-mediated network infrastructure within the enterprise at whatever moment during January.

A LAN is a network for communication between computers confined to a single building or in closely located group of buildings, permitting users to exchange data, share a common printer or master a common computer, etc.

LAN is a proxy for internal computer network. The relevant concept to be measured is internal connectivity in the enterprise. This connectivity facilitates the sharing of information among different people in an organisation and it's a pre-condition for the automation and integration of business processes in the enterprise.

However, internal computer network is a concept which might be understood in very broad terms. Computer networking includes also the linking of computers to other devices, including telephones and personal digital assistants (PDA's), through for example USB or FireWire, and can include only two computer devices. Such networking is not very informative about the enterprise's ability to integrate its business processes and the concept used should narrower.

On the other hand, LANs are a basic and common way of connecting computers within a company size. And larger scale computer networks, at enterprise level, are generally made up of several LAN's connected or at least one LAN to which single computers at different sites connect. For these reasons, LAN is used to proxy the existence of a reasonable level of connectivity in the enterprise.

a) Wireless LAN

A wireless LAN is a local network which uses wireless technologies such as radio-frequency, infrared, microwave, or other types of electromagnetic or acoustic waves for the last internal link connecting employees' devices (computers, printers etc) to the business computer-mediated network. to transmit signals or data between stationary (fixed) points (connecting employees' devices such as computers, printers etc) and a LAN backbone line(s) within the enterprise's working premises.

"Backbone is a line or set of lines that local area networks connect to span distances efficiently (for example, between buildings)."

b) Wire based LAN

A wire based LAN uses wires, cables, or fibre optics for the last internal links to transmit signals or data between stationary (fixed) points (connecting employees' devices such as computers, printers etc) and a LAN backbone line(s) within the enterprise's working premises.

c) Intranet

Intranet refers to the use of Internet protocols (TCP/IP and others) to share and exchange company information in the same way as over the Internet (web site interface)

but with access restricted to the internal users via internal business computer-mediated network (LAN or set of enterprise's LANs) infrastructure, accessible only by authorised users (organisation's members, employees, or others with authorisation).

Typically the term intranet refers to the internal web site (an intranet's Web sites look and act just like any other Web sites). The same concepts and technologies of the World Wide Web such as web browsers and servers running on the internet protocol suite are used to build an intranet. Other internet protocols are commonly used as well, especially ftp and email. Intranet in general looks like a private version of the Internet (web sites). It may consist of many interlinked local area networks and also use leased lines in the wide area network. There does not necessarily have to be any access from the organisation's internal network to the internet itself (world wide web). However, usually larger enterprises allow users within their intranet to access the public Internet through firewall servers with a gateway through which all access takes place. Company security very often has the ability to screen messages in both directions so that company security is maintained. This means that organisations that allow their staff internet access can normally determine which internet web sites are being viewed, block access to specific sites they don't want them to see (such as pornographic sex, sites), and even trace offenders who persistently attempt to view them. They can also block certain types of web content (such as objects) which they consider a particular security risk."

When part of an intranet is made accessible to customers, partners, suppliers, or others outside the company, that part becomes part of an extranet.

c) Extranet

A secure extension of an Intranet that allows external users to access some parts of an organisation's Intranet.

A4: ***Did your enterprise have in use, in January 2007, any software application to manage the placing and/or receipt of orders?***
If Yes,
A5: ***Did that software application to manage orders link automatically with any of the following, as of January 2007?***

[<u>Scope</u>: enterprises which use computers, i.e. A1 = Yes]

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

The purpose of these variables is to provide a measure of the integration of business processes. Implicitly, we are identifying that integration with automatic links. That is, something changes in the IT system supporting a business process and automatically (i.e. without human intervention) that change is reflected in the IT system supporting another business process. The next issue then is which business processes we are interested in (as we can't cover all business processes in the enterprises, and as they actually differ between enterprises from different economic activities). One solution is to start with two business processes that exist in every enterprise: selling and buying. (That's why we start with a question on the existence of software application for managing orders, received and placed.)

As the final purpose of these questions is to identify the existence of automatic links between different business processes, the definition of the starting point (software application for managing orders) should be as broad as possible.

At the same time, the definition of software application is not restricted to any specific software used or its level of sophistication. The level of sophistication of a software application or a system is not directly correlated to its effectiveness or efficiency, and therefore correlated to a positive impact in the enterprise, which ultimately that's what we want to identify.

So, in conclusion, even an excel worksheet or an access database where orders are inserted manually can be considered as a software application for managing orders.

The term "IT system" in these questions is not used anymore and changed to "software application". In terms of measurement they are equivalent and the latter term is more specific and easily identifiable by the respondent. An "IT system" consists of a system composed of hardware and software. However, we can solely refer to software as it always requires the existence of hardware.

A6: Did your enterprise have in use, in January 2007, an ERP software package to share information on sales and purchases with other internal functional areas (for example, finance, planning, marketing, etc.)?

[<u>Scope</u>: enterprises which use computers, i.e. A1 = Yes]

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

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.

ERP systems have the following characteristics:

1. ERP systems are software designed for a client 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, built on purpose for 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 show that ERP is easy to ask. However, it is not necessarily easy to understand by the respondents.

It is argued that if the respondent doesn't know what ERP is, then he's not using it (just like for most ICT related questions). However, this opinion is not consensual. It depends on what one specifically mean 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.

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.

As such, this question is not a substitute for questions A4 and A5, although they are very much related.

Questions A4/A5 measure the integration of business processes (internal and external) looking at the link between the front-office (customers) and other 3 specific business functions: invoicing, production and re-ordering. Question A6 is intended to measure the adoption of a specific type of software packages.

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.

Therefore, the approach here is to measure the integration of business processes with both questions A4/A5 and question A6.

The question on ERP software should follow questions A4 and A5 on the integration of business processes. The question on ERP adoption shouldn't be filtered by questions A4 or A5, because any change to the specific items of question A5 or to the question A4 would make the results non comparable between different years.[1]

The measurement of integration of business processes should be based on transactions, i.e., from the front office to the back office, just like it is done in question A9. Therefore, question A6 refers to sales and purchases.

---

[1] The "software application to manage orders" in question A4 is a concept wider then "ERP software" in question A6, because it doesn't refer to any integration (only in A5 that is checked), while in an ERP software integration is implicit. It is also wider because enterprises can use other software applications, besides ERP, to manage orders. Therefore, we wouldn't expect an enterprise to tick "yes" in A6 if it ticked "no" in A4. However, although questions A4 and A5 are at the moment the best questions in the enterprise community survey to measure internal integration of business processes, it is believed that they will need to be revised in the future. So question A6 should not be filtered by question A4 in order not to provoque a break in the series, in case question A4 is changed. In case an enterprise ticks "no" in A4 and "yes" in A6, this should be taken as a warning that question A4 was not well understood by the respondent and the answer should be checked.

A7: Did your enterprise have in use, in January 2007, any software application for managing information about clients (so called CRM) that allows it to:

a) Capture, store and make available to other business functions the information about its clients?
b) Make analysis of the information about clients for marketing purposes (setting prices, make sales promotion, choose distribution channels, etc.)?

[<u>Scope</u>: enterprises which use computers, i.e. A1 = Yes]

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

CRM stands for Customer Relationship Management and 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 ICTs. 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.
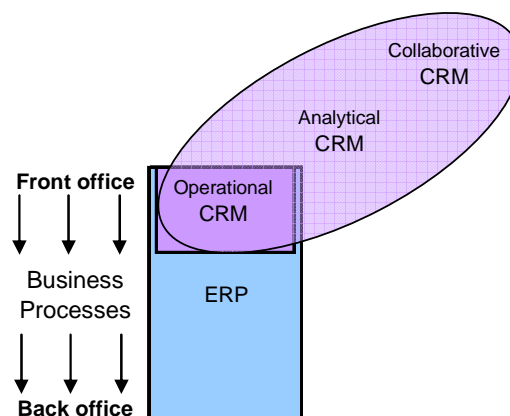
Although CRM is more than adoption of ICTs it is based on an intensive use of these technologies. It is believed that the adoption of CRM software improves the marketing functional area of the enterprises, by improving customer service and customer relationship. Improvement comes, for instance, from 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 well established 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 splat in two, one focusing on operational CRM and another on analytical CRM (items a and b of the question).

A8: Did your enterprise have in use, in January 2007, third party free or open source operating systems, such as Linux ? (i.e. with its source code available, no copyright cost, and the possibility to modify and/or (re)distribute it)

[Scope: enterprises which use 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 percentage of enterprises that use open-source software at systems level in their business activity. The adoption of open source is measured in this question at the level of systems software because it seems to be more feasible to collect than at any software level, as the IT department might not be aware of the use of general open source software by users throughout the organisation.

Although "at systems level" basically means operating systems and servers, "server" is a difficult concept to use in the data collection, as it can be understood in very different ways. Therefore, only reference to "operating systems" is made in the question.

Linux is an example of system level software which is open source. The use of an example helps to clarify the question and improve the accuracy of the answer. However, it is very probable that the inclusion or non inclusion of the example in the question has a significant impact on the results. Therefore, in order to assure the comparability of the results among different countries, they all have to use the example in the question.

The reason for measuring the use of open source is on one hand because it reduces the cost of IT infrastructure of the enterprise and on the other hand because it provides interoperability. By reducing the cost of the IT infrastructure, open source is believed to promote / facilitate the adoption of e-business. For this reason, the measurement of open-source is most relevant at systems level.

Both concepts of "free software" and "open source software" are included in this measurement. Although very similar, software licenses under one of the definitions can be

non compliant with the other, while the benefits just described above come from both free and open source software.

Another reason for measuring the adoption of open-source is that there is an interest of policy makers on the use of open-source as a business model, i.e., on the development of business model as a profitable activity (selling consulting and training services on the developed open-source software, developing a more advanced version of the software which needs to be purchased, etc.). From that point of view, in order to understand the potential of the "open-source business model", it is important to know how high the adoption of open-source is. Nevertheless, the use of the information from this variable for this purpose is limited, because it would require the collection of information on the adoption of open source not only at systems level, but in general.

As the two key definitions for this measurement are those of open source and systems level software, they are explained next.

Definition of systems level software

Practical computer systems divide software into three major classes: system software, application software and programming software, although the distinction is somewhat arbitrary, and often blurred.

- **System software** helps run the computer hardware and computer system. It includes operating systems, device drivers, diagnostic tools, servers, windowing systems, utilities and more.

- **Application software** allows a user to accomplish one or more specific tasks. Typical applications include office suites, business software, educational software, databases and computer games. Most application software has a graphical user interface (GUI).

- **Programming software** usually provides some useful tools to help programmer to write computer programs and software using different programming language in a more convenient way. The tools include text editor, compiler, interpreter, linker, debugger, and so on. Integrated development environment (IDE) merges those tools in a software bundle, and programmer may not need to type a lot of commands for compiling, interpreter, debugging, tracing, and etc., because IDE mostly has a GUI.

Definition of open source software

Open source software refers to computer software under an open source license. An open-source license is a copyright license for computer software that makes the source code available under terms which allow for modification and redistribution without having to pay the original author. Such licenses may have additional restrictions such as a requirement to preserve the name of the authors and the copyright statement within the code. One popular (and sometimes considered normative) set of open source licenses are those approved by the Open Source Initiative (OSI) based on their Open Source Definition (OSD).

The **Open Source Initiative** is a non-profit organization dedicated to promoting open source software, namely through promoting its **Open Source Definition**. This Open Source Definition is used by the OSI to determine whether or not a software license can be considered open source. Under the Open Source Definition, licenses must meet ten conditions in order to be considered open source licenses:

1. Free Redistribution: the software can be freely given away or sold.

2. Source Code: the source code must either be included or freely obtainable.

3. Derived Works: redistribution of modifications must be allowed.

4. Integrity of the author's source code: licenses may require that modifications are redistributed only as patches.

5. No discrimination against persons or groups: no-one can be locked out.

6. No Discrimination against fields of endeavour: commercial users cannot be excluded.

7. Distribution of license: The rights attached to the program must apply to all to whom the program is redistributed without the need for execution of an additional license by those parties.

8. License must not be specific to a product: the program cannot be licensed only as part of a larger distribution.

9. License must not restrict other software: the license cannot insist that any other software it is distributed with must also be open source.

10. License must be technology-neutral: no click-wrap licenses or other medium-specific ways of accepting the license must be required.

It should be noted that merely opening the source code publicly does not mean that the code can be called 'Open Source', as the term is normally used as defined above. Indeed, it does not give the rights normally associated with open source software such as the freedom to adapt, test or redistribute. For example, the 'Shared Source Initiative' of Microsoft requires an agreement to be signed before gaining access to the source and furthermore the agreement does not confer any of the open source software rights.

There are also shared source licenses which have some similarities with open source, but a number of critical differences make such licenses incompatible with the Open Source Definition. Some software licenses define an open standard basis and may or may not be similar to open source, like some versions of Solaris and PGP.

Related to the Open Source Definition is the **Free Software** definition by the Free Software Foundation, which attempts to capture what is required for a program license to qualify as being free-libre software. In practice, licenses which meet the open source definition almost always also meet the Free software definition. All licenses reported to meet the free software definition as of 2005 also meet the open source definition.

Free software is software that meets the following criteria:

1. users are free to use the program for any purpose;

2. users are free to examine the software to see how it works;

3. users are free to redistribute the program to others;

4. users are free to improve the program.

Although it is not explicitly referred to, access to the source code is a precondition for fulfilling these criteria.

A related concept is that of **Copyleft**. Copylefting licences grant the freedom of free or open source software subject to one condition: if the amended code is redistributed, it must be

under the same licence. This is the most controversial aspect of the GNU General Public License (GPL), but guarantees the rapid expansion of the community around the code developed under this licence. It also prevents any appropriation of code for lock-in purposes which would be possible with more permissive kind of licences. Copylefting license examples are, e.g., GPL, L-GPL and the W3C Software Notice and Licence.
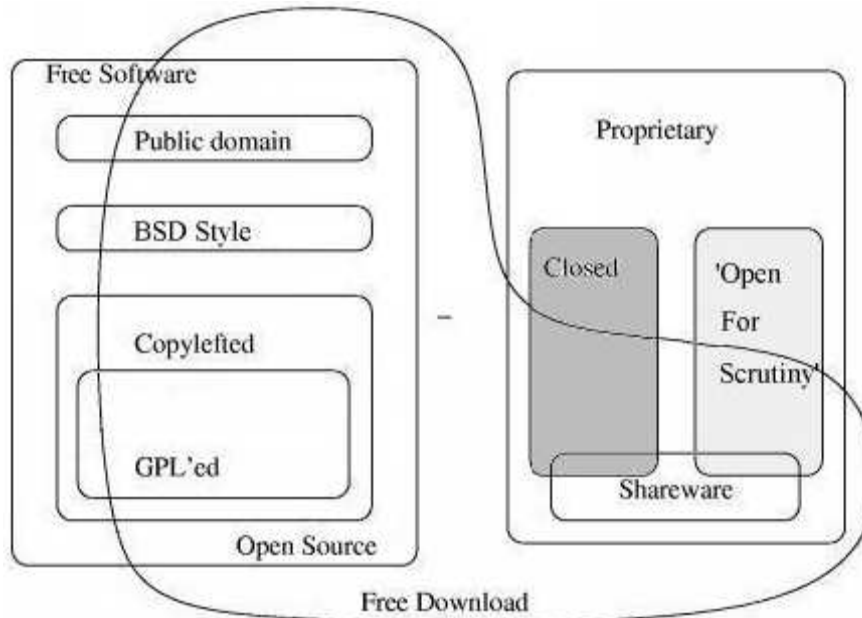
Software in the public domain (that is, with no copyright license at all), meets the criteria for free or open source software as long as all source code is made available.

Free or open source software can be developed to work standalone or to be integrated with existing proprietary software.

The following list presents the software licenses which, as of 2005, were approved by OSI and therefore follow the open source definition.

- Academic Free License
- Adaptive Public License (APL)
- Apache Software License
- Apple Public Source License
- Artistic License
- Attribution Assurance Licenses
- BSD License
- Computer Associates Trusted Open Source License
- Common Development and Distribution License
- Common Public License
- End-User License Agreement (EULA)
- EU DataGrid Software License
- Eclipse Public License
- Educational Community License
- Eiffel Forum License
- Entessa Public License

- Fair License
- Frameworx License
- GNU General Public License (GPL)
- GNU Lesser General Public License (LGPL)
- Historical Permission Notice and Disclaimer
- IBM Public License
- Intel Open Source License
- Jabber Open Source License
- Lucent Public License
- MIT License
- MITRE Collaborative Virtual Workspace License (CVW License)
- Motosoto License
- Mozilla Public License 1.0 (MPL)
- Mozilla Public License 1.1 (MPL 1.1)
- NASA Open Source Agreement
- Naumen Public License
- NetHack General Public License
- Nokia Open Source License
- OCLC Research Public License
- Open Group Test Suite License
- Open Software License
- PHP License
- Python License
- Python Software Foundation License
- Qt Public License (QPL)
- RealNetworks Public Source License
- Reciprocal Public License
- Ricoh Source Code Public License
- Sleepycat License
- Sun Industry Standards Source License (SISSL)
- Sun Public License (SPL)
- Sybase Open Watcom Public License
- University of Illinois/NCSA Open Source License
- Vovida Software License v. 1.0
- W3C License
- wxWindows Library License
- X.Net License
- zlib-libpng license
- Zope Public License

Examples of licenses which include the access to the source code but which do not follow the open source definition are:

- HESSLA
- Solaris
- PGP

A9: Was your enterprise, in January 2007, sending e-invoices in a digital format which allows its automatic processing?

A10: Was your enterprise, in January 2007, receiving e-invoices in a digital format which allows its automatic processing?

[<u>Scope</u>: enterprises which use computers, i.e. A1 = Yes]

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

An **e-invoice** is an invoice[2] where all data is in digital format and it can be processed automatically. A distinctive feature of the definition of e-invoice adopted in this manual is automation[3]. E-invoices need to offer the possibility to be automatically produced by the seller, transferred and processed on reception by the buyer.

e-Invoices are a form of electronic integration of business processes between different enterprises. Its measurement is important for the more general measurement of e-business.

An invoice sent by fax, even if done in an automatic way, is not an e-invoice. It is in an electronic format, but it is transmitted as a picture and therefore the information is not in a suitable form to be extracted and reused automatically.

An invoice sent by e-mail can be an e-invoice or not. If the information of the invoice is sent in a pre-defined format which allows the receiver to process it automatically then it is an e-invoice. However, if the message includes the information of an invoice manually typed and not following a pre-defined format then it is not an e-invoice. If the invoice is a file of its picture (or bitmap) attached to the e-mail message then, also in this case, it is not an e-invoice, because the information cannot be used automatically.

In general, any invoice sent electronically in a bitmap format, in which the only information being transmitted is that of a picture (of an invoice in this case) are not e-invoices. Here, we consider in digital format only information which is transmitted in a form which allows its automatic processing.

It is important to note that what defines an e-invoice is its format which allows its automatic processing and not if it is actually treated automatically. There are two reasons to define it in this way, a conceptual and a practical one. Conceptually, it is desirable that an invoice is considered as an e-invoice depending on its own characteristics and not on how it is treated. Otherwise, an invoice could be an e-invoice when sent by the seller, if produced automatically, and not be an e-invoice when received by the buyer, if not treated automatically. This also means that, in practical terms, when a seller sends an invoice in a digital format, he doesn't know if it will be processed automatically by the receiver. If the definition depended on how a digital invoice is treated then he wouldn't know if it is an e-invoice or not.

The definition of e-invoice is not restricted by the format used, which might be XML, EDI or other similar format.

---

[2] "*Invoice is a commercial document issued by a seller to a buyer, indicating the products, quantities and agreed prices for products or services that the seller has already provided the buyer with. An invoice indicates that, unless paid in advance, payment is due by the buyer to the seller, according to the agreed terms.*" (http://en.wikipedia.org/wiki/Invoice)

[3] The discussion on the legal definition of e-invoice is still going on. While it is clear that an e-invoice is an invoice where all data is in digital format, the requirement that it is in a format that allows its automatic processing is not so clear-cut.

An invoice should not be confused with other commercial documents related to transactions, as for example the receipt.[4]

When collecting this variable, the question should not refer explicitly to the transmission of the document from the seller to the buyer, because the use of e-invoices can include third parties, like service providers or banks.

The measurement of the adoption of e-invoices is split between the reception of e-invoices and the sending of e-invoices. The two questions measure two different things. Sending invoices in digital format says something about the level of ICT use in the enterprise. It requires the existence of IT systems in the enterprise to make and send them. Receiving is a more passive role. It might even happen that internally after reception they are printed and treated as if they were originally in paper.

A11: Was your enterprise, in January 2007, using a digital signature in any message sent, i.e. using encryption methods that assure the authenticity and integrity of the message (uniquely linked to and capable of identifying the signatory and where any subsequent change to the message is detectable)?

[<u>Scope</u>: enterprises which use computers, i.e. A1 = Yes]

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

This question aims at measuring the adoption of digital signatures by enterprises, also referred to as advanced electronic signature[5]. One important distinction to note is between the more general concept of electronic signature and the narrower one of digital signature.

Electronic signature refers to data in electronic form which are attached to or logically associated with other electronic data and which serve as a method of authentication constituting in that manner a signature. A digital signature is an electronic signature which uses cryptographic means to add non-repudiation and message integrity. Non-repudiation means that the message is authentic and is uniquely linked to and capable of identifying the signatory. Integrity means that any eventual subsequent change to the data will be detected if tested for using the signature.

A signature in paper which is then digitised and sent together by e-mail is not a digital signature. It is even questionable that it is an electronic signature. Even if considered an electronic signature, it isn't considered a digital signature in this manual because it can't assure that the data transmitted was really created by the signatory and it has no way to assure the integrity of the message. Similarly, a signed paper document sent by fax does not constitute the use of a digital signature.

Examples of electronic signatures which are not digital signatures are PIN codes, chip cards, retina scanners and fingerprint readers. Although these devices allow the authentication of signatories, they don't provide data integrity assurance.

Some web pages and software EULAs (end user license agreements) claim that various electronic actions are legally binding signatures, and so are instances of electronic signature.

---

[4] "*A receipt is a document made by a merchant, landlord, vendor, or other creditor in favour of a customer acknowledging having received an amount of money for some purpose.*" (http://en.wikipedia.org/wiki/Receipt)
[5] Advanced electronic signature is the term used in the Directive 1999/93/EC of the European Parliament and of the Council on a Community framework for electronic signatures.

For example, a web page might announce that, by accessing the site at all, you have agreed to a certain set of terms and conditions. However, these mechanisms are not digital signatures.

It is important to understand that the use of cryptographic techniques in digital signatures are much more than an error checking technique akin to checksum algorithms, or even high reliability error detection and correction algorithms. These can offer no assurance that the text has not been changed. In addition, no message integrity protocols include error correction, for to do so would destroy the tampering detection feature.

Popular electronic signature standards include the OpenPGP standard supported by PGP and GnuPG, and some of the S/MIME standards. All current cryptographic digital signature schemes require that the recipient have a way to obtain the sender's public key with assurances of some kind that the public key and sender identity belong together, and message integrity measures which assure that neither the attestation nor the value of the public key can be surreptitiously changed. However, a secure channel is not required to transmit the data or message.

Council Directive 2001/115/EC (the "e-invoice" Directive) includes as one of the conditions to make the acceptation of e-invoices obligatory the use of procedures that assure the authenticity and integrity of its contents, of which digital signature is one of the methods. The adoption of digital signatures is also relevant in a broader sense as it is also a measure to increase security in the online environment and it applies to all electronic documents circulating electronically between enterprises.

The requirement that the e-signature has a legal binding status is not considered as a relevant feature for measurement. What is important is the adoption of this authentication mechanism.

In the scope of EU legislation the digital signature is referred to as "advanced electronic signature". However, the wording "digital signature" should be used in the question as supposedly it is the term used most often by IT professionals.

Nevertheless, for the construction of the question, the wording of the Directive 1999/93/EC, when it refers to message integrity and authenticity, should be used (article 2, point 2).

I.1.6.2 MODULE B: USE OF INTERNET

B1. Did your enterprise have access to Internet, during January 2007?

[Scope: enterprises which use 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 availability regardless of property and purposes of the connection at whatever moment during January 2006. The term "Internet access" means having an external connection to the Internet through an Internet access 'service' provider (ISP).

The technologies used to connect to the Internet are physical based or wireless. So we can distinguish wire based connections from wireless connections. Another important distinction is between fixed connections and mobile connections. Access via fixed line refers to the connection between to stationary points, while mobile connections are performed via mobile phone networks. The distinctions wire/wireless and fixed/mobile are not the same, as wireless connections can be used to connect two fixed points, as is the case of satellite connections. Another breakdown of the Internet access is between temporary (dial-up on demand) and permanent connection (always on-line).

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 protocols (Internet protocol suit).

The Internet is made up of thousands of smaller commercial, academic, 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 following Internet Protocol (IP) based networks and systems: WWW; Extranet over the Internet; EDI over the Internet; Internet accessed by mobile phones; Internet email."

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 two to be defined.

An Internet access 'service' provider (ISP) is a company that provides individuals and other companies 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. The consumer then has access to the Internet for an ulimited number of hours, although the speed at which this data is transferred varies widely and depends on the type of the Internet connection.

B2: How many persons employed used computers connected to the World Wide Web at least once a week, during January 2007?

or

Please indicate an estimate of the percentage of the number of persons employed used computers connected to the World Wide Web at least once a week, during January 2007.

[<u>Scope</u>: enterprises with Internet access, i.e. B1 = Yes]

[<u>Type</u>: numerical, in real 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. In principal, any person employed in an enterprise with access to the Internet can access it from a computer in the enterprise. However, here only persons employed to whom it was granted access are counted. The criteria in this variable to identify Internet access is WWW access. The reason for its usage 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 it's very already very common in Europe and it doesn't provide the same potentialities in terms of access to information as the WWW.

Just like with persons employed using computers (question A2), this variable can be collected in **real** 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 real or percentage values.

Independently on how this variable is collected, in real or percentage values, the background variable "number of persons employed" is needed for grossing up the overall percentage of persons employed using computers.

B3. Did your enterprise have the following types of external connection to the Internet, during January 2007?

[<u>Scope</u>: enterprises with Internet access, i.e. B1 = Yes]

[<u>Type</u>: binary (*Yes/No*)]

By external 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 access service provider "ISP". <u>The last mile</u> is the final leg of delivering communications connectivity to the enterprise.

Type of external Internet connection:

a) Traditional Modem (dial-up access over normal telephone line)

Dial-up is a temporary connection to the Internet via an analogue (standard) modem and standard telephone line (public switched telephone network PSTN), which requires that the modem dial a phone number when Internet access is needed (to dial the Internet service provider's node to establish a modem-to-modem link, which is then routed to the internet).

Dial-up pertains to a telephone connection in a system of many lines shared by many users. A dial-up connection is established and maintained for limited time duration. A

dial-up connection can be initiated manually or automatically by the computer's modem or other device. This type of the Internet connection is capable of carrying up to 56 kilobits per second (Kbit/s) and is considered as the typical example of the low capacity 'speed' connection (narrowband).

b) ISDN connection

ISDN (Integrated Services Digital Network) connection is a temporary connection to the Internet using a type of circuit switched telephone network system (a set of CCITT/ITU standards), designed to allow digital (as opposed to analogue) transmission of voice and data over ordinary telephone copper wires (enables digital transmission over the public switched telephone network), resulting in better quality and higher speeds, than available with analogue systems.

Enterprises that install an ISDN adapter (in place of a telephone modem) receives up to 128 Kbps compared with the maximum 56 Kbps rate of an analogue (standard) modem connection. ISDN services can simultaneously transmit voice, data and video. ISDN is also considered as a low capacity 'speed' connection (narrowband).

This category includes Basic Rate Interface (BRI) consisting of two B channels, each with bandwidth of 64 Kbit/s, and one D channel with a bandwidth of 16 Kbit/s.

This category does not include Primary-Rate Interface, a type of ISDN service designed for larger organizations. PRI includes 23 B-channels (30 in Europe) and one D-Channel. PRI service is generally transmitted through an E1 line in Europe. This type of connection belongs under alternative e).

c) DSL (xDSL, ADSL, SDSL etc) connection

xDSL (Digital Subscriber Line) refers to a family of a high-bandwidth (broadband), local loop technologies that provide a digital permanent Internet connection over the copper wires of the local telephone network.

Asymmetric Digital Subscriber Line (ADSL) where more bandwidth is allocated to download than upload and High Rate Digital Subscriber Line (HDSL) are considered dominant DSL technologies. Typically, individual connections will provide from 1.544 Mbps to 512 Kbps downstream and about 128 Kbps upstream. A DSL line can carry both data and voice signals and the data part of the line is continuously connected. DSL connection is considered as one of the high capacity 'speed' permanent 'fixed' Internet connection (broadband).

d) Other fixed internet connection

Other high capacity 'speed' fixed (wire or wireless) connection includes the following types of the Internet connection: Cable modem 'cable TV network connection'; High capacity leased lines 'Frame Relay, ATM, Digital Multiplex'; Ethernet LANs connection; Optical fibre connection; Satellite connection; Wi-fi connection, other FWA connections etc..

<u>Cable modem</u> 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 and receive data at about 1.5 Mbps. It is considered as one of the high capacity 'speed' permanent 'fixed' Internet connection (broadband).

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 Mbps. 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.

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 E-1 and E-3) lines are used for the Internet connection via a leased lines. Higher speeds are available on alternative interfaces.

Frame relay is an efficient data transmission technique used to send digital information quickly and cheaply to one or many destinations from one point. It is a packet-switching protocol for connecting devices on a Wide Area Network (WAN). Frame Relay networks in the U.S. support data transfer rates at T-1 (1.544 Mbps) and T-3 (45 Mbps) speeds. In fact, you can think of Frame Relay as a way of utilizing existing T-1 and T-3 lines owned by a service provider. Most telephone companies now provide Frame Relay service for customers who want connections at 56 Kbps to T-1 speeds. (In Europe, Frame Relay speeds vary from 64 Kbps to 2 Mbps. Frame relay is being displaced by ATM and native IP based products, including IP virtual private networks.

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 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 capable of speeds of up to 11 Mbps. 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 phone networks, which are included in alternative e) instead.

e) Mobile connection

This variable includes connections to the Internet via mobile phone networks and includes connections via low capacity 'speed' mobile phone networks (HSCSD, GPRS etc) - mobile

Internet narrowband - and connections via high capacity 'speed' mobile phone networks (UMTS 'W-CDMA'; EDGE 'EGPRS'; CDMA 2000x etc) - mobile Internet broadband.

A Connection via **low capacity 'speed' mobile phone networks** is an access to the Internet using a long range wireless transmission of the mobile network technologies as High-Speed Circuit-Switched Data (HSCSD) or General Packet Radio Service (GPRS) that is sometimes called as 2,5 mobile generation technology (2,5 G).

Access to the Internet via mobile phone networks should be considered as low capacity 'speed' mobile connection (narrowband) if it is being equal to, or greater than 256 kbit/s, as the sum of the capacity in both directions (download or upload).

HSCSD is a development of Circuit Switched Data, the original data transmission mechanism of the GSM mobile phone system. As with the CSD channel the allocation is done in circuit switched mode. The difference comes from the ability to use different coding methods and even multiple time slots to increase data throughput. HSCSD is a temporary mobile connection.

2.5G is a stepping stone between 2G and 3G cellular (mobile) wireless technologies. The term "second and a half generation" is used to describe 2G-systems that have implemented a packet switched domain in addition to the circuit switched domain. While the terms "2G" and "3G" are officially defined, "2.5G" is not. 2.5G provides some of the benefits of 3G (e.g. it is packet-switched) and can use some of the existing 2G infrastructure in GSM and CDMA networks. The most commonly known 2.5G technique is GPRS.

GPRS is a 2.5G mobile standard typically adopted by GSM operators as a migration step towards 3G (W-CDMA). GPRS is based on packet-switched technology enabling high-speed data transmission.

Examples of some most common spread standards of low capacity 'speed' mobile network connection and their bandwidth (can be different between the countries and also within the same country):

	download	upload
GPRS 4+1	57.6 kbit/s	14.4 kbit/s
GPRS 3+2	43.2 kbit/s	28.8 kbit/s
CSD	9.6 kbit/s	9.6 kbit/s
HSCSD	28.8 kbit/s	14.4 kbit/s
HSCSD	43.2 kbit/s	14.4 kbit/s

A connection via **high capacity 'speed' mobile phone networks** is an access to the Internet using a long range wireless transmission of the 3rd generation (3G) 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) or some other high capacity mobile technologies based on the GPRS as EDGE (Enhanced Data rates for Global Evolution - EGPRS) etc.

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.).

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 kbps and up to 2 Mbps 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.

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. The standard is based on the GSM standard and uses TDMA multiplexing technology.

Narrowband is defined by the type of the Internet connection and includes: 1) Dial-up (connection via standard telephone line 'PSTN using analogue modem); and 2) ISDN (Integrated Services Digital Network)

Broadband is defined by the type of the Internet connection and includes: 1) xDSL (Digital Subscriber Line) technologies as ADSL, HDSL, SDSL, VDSL that form core broadband; and 2) Other high capacity 'speed' fixed (wire or wireless) connection (Cable modem 'cable TV network connection'; Leased lines 'Frame Relay, ATM, Digital Multiplex'; Ethernet LANs connection; Optical fibre connection; Satellite connection; Wi-fi connection etc) – extendend fixed wire based and wireless broadband.

B4. Did your enterprise use the Internet for the following purposes, during January 2007?

[Scope: enterprises with Internet access, i.e. B1 = Yes]

[Type: one single answer needed, i.e. *Tick only one*]

This variable intends to measure the adoption of the Internet by the enterprise as a user or consumer of several services. These services may be paid or not. This does not include the provision of these services in the Internet. For example, for "banking and financial services" this variable measures the use of the Internet to use the services provided by banks and not the use of the Internet by banks to provide financial services.

This variable includes the use of the Internet in general and is not restricted to the World Wide Web.

The following Internet uses are measured:

a) Banking and financial services

This includes the use of web-banking, the consultation of financial information (like exchange rates, interest rates, etc.) and the use of the Internet for automatic data interchange between the enterprise and other financial organisations (for instance through the use of XML or EDIFACT).

b) Training and education

This includes the use of online education programs by the employees.

c) Market monitoring (e.g. prices)

This includes the use of the World Wide Web to look for information on current or potential suppliers, clients or competitors.

B5. Did your enterprise use the Internet for interaction with public authorities, during 2006?

[<u>Scope</u>: enterprises with Internet access, i.e. B1 = Yes]

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

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 authorities refer to both public services and administration activities. While 'administration' refers to obligation and rights one has as a citizen or as an enterprise in activity in the country (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.

B6. Did your enterprise interact with public authorities in the following ways, during 2006?

[<u>Scope</u>: enterprises having interacted with public authorities, i.e. B6 = Yes]

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

a) For obtaining information

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 to facilitate respondents to recall public authority web sites that they have used. For example respondents may not realise some public information sites are run by public authorities, e.g. a meteorological office offering weather forecast services.

b) For obtaining forms, e.g. tax forms

Includes downloading official forms 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) For returning filled in forms, e.g. provision of statistical information to public authorities

Includes filled in forms sent via internet only, either . Forms downloaded, printed, filled in and sent by post should not be included in this category.

d) Submitted a proposal in an electronic tender system (e-procurement)

This item includes the submission of proposals in Internet based systems (either based on web interfaces or any other architecture). It excludes the sending of proposals by normal manually typed email messages.

This question makes use of the conceptual framework for the levels of interaction between enterprises or citizens and public authorities, used by DG INFSO. 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). The first three levels are considered in the enterprise survey, as the last one has proved to be very difficult to collect.

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 *b*).

B7. Did your enterprise have a Web Site / Home Page, during January 2007?

[<u>Scope</u>: enterprises with Internet access, i.e. B1 = Yes]

[<u>Type</u>: 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. It includes not only the existence of a website which is located in servers which belong to the enterprise or are located at one of the enterprise's sites, but also third party web-sites (e.g. one of the group of enterprises to which it belongs).

However, it doesn't include any web presence 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.

B8. Did the Web Site of your enterprise provide the following facilities, during January 2007?

[<u>Scope</u>: enterprises with a website, i.e. B8 = Yes]

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

a) Marketing the enterprise's products

It includes providing information on the enterprise's products in its website.

b) Facilitating access to product catalogues and price lists

It includes the provision of product catalogues, including prices, in the enterprise's website.

c) Providing after sales support

It includes the use of the website to provide support to clients, for example with a web-form where clients can pose questions.

I.1.6.3   MODULE C: E-COMMERCE VIA INTERNET

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. According to this definition:

> An electronic transaction is the sale or purchase of goods or services, whether between businesses, households, individuals or private organisations, conducted over computer-mediated networks. The goods and services are ordered over those networks, but the payment and the ultimate delivery of the good or service may be conducted on or off-line.

> **Source:** OECD Guide to Measuring the Information Society

This module covers only orders which are transmitted via Internet, while orders transmitted via other computer networks are covered in module D.

One important difference in this module compared with most of the questionnaire is that the reference period is the previous calendar year, instead of January of the survey year. Flux economic variables like turnover and purchases, the main variables to measure in e-commerce, need to be measured for a period longer than 1 month. In order to keep comparability with the main business statistics a calendar year is taken as reference period.

The concept of Internet is the same as the one taken in module B. That is, it refers to the public open worldwide network known as Internet and not to the use of the Internet protocol in closed networks.

e-Commerce consists on a business process which is conducted via computer networks, in this case the Internet. The role of the computer network is very important. It doesn't include the use of the Internet merely as a human-to-human communication tool. For this reason conventional e-mails manually typed by humans are excluded from this definition. In this case the computer and the network are being used only as a communication device, just like a telephone and its network. However, if the e-mail system is just used for the transmission of an automatic 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 orders transmitted from a human to another human in a manual manner through the Internet are not considered e-commerce. The typical ways in which e-commerce orders are placed via the Internet is: 1) through a web-site, which typically characterises business-to-consumer e-commerce; 2) through automatic data interchange, i.e. using XML, which characterises business-to-business e-commerce; and 3) through B2B e-marketplaces. Nevertheless, any other automatic use of the Internet to transmit orders is included.

C1. Did your enterprise order products/services via the Internet, during 2006 (excluding manually typed e-mails)?

[<u>Scope</u>: enterprises with Internet access, i.e. B1 = Yes]

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

C2. Please indicate for 2006 the percentage of the Internet orders in relation to the total purchases (in monetary terms, excluding VAT).

or
Please state the value of the purchases resulted from orders placed via Internet (in monetary terms, excluding VAT), in 2006.
> *If you can't provide this value,*
> ### Please indicate an estimate of the percentage of the total purchases resulted from orders placed via Internet, in 2006.

[<u>Scope</u>: enterprises which have ordered via Internet, i.e. C1 = Yes]

[<u>Type</u>: numerical]

The value of the purchases made over the Internet can be collected in 3 different ways. The first one, with the lowest level of measurement, is to collect it in terms of classes of percentage of the total purchases. This is the minimum level of measurement required for this variable. The reason is that in several countries the collection of any other information on this value is not feasible. In this case, the 5 response categories used are:

- Less than 1%

- 1% or more and less than 5%

- 5% or more and less than 10%

- 10% or more and less than 25%

- 25% or more

The main disadvantage of collecting the information in this way is that it doesn't allow to aggregate the value of Internet purchases.

As the aggregated value of Internet purchases is one indicator of policy and general interest, alternative ways of collecting this data are recommended when such is possible. Internet purchases should in alternative be collected with a numerical value, either in real terms or in percentage terms.

Internet purchases collected in percentage values

Please indicate an estimate of the percentage of the total purchases resulted from orders placed via Internet, in 2005.	%

Collecting Internet purchases in percentage values has the advantage of allowing us to obtain an approximate estimate from respondents which don't have in the enterprise records which can provide this value. Field experience has shown that enterprise purchases are done in a decentralised way. 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 Internet purchases.

However, 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%. The problem is that this question doesn't allow the respondent to provide such an answer. The answer in this case would have to be either 0% or 1%.

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%. This has to be taken into account when the results are tabulated in percentage classes. If 1% is at the limit of the classes used (e.g. [0% - 1% [; [1% - 5% [), then in which class the 1% is included will have a significant impact on the results.

Another disadvantage of this method is that for big enterprises a value of less than 1% can still be a significant amount. If collected in percentages, Internet purchases of less than 1% can either not be taken into account if the respondent answers 0% or be over-estimated. When answers from big enterprises are involved this can bring a significant instability to the results.

Internet purchases collected in real values

Please state the value of the purchases resulted from orders placed via Internet (in monetary terms, excluding VAT), in 2005.	*(National Currency)*

Therefore, the preferable way to collect Internet purchases is in real values. Unfortunately, for a large number of enterprises this value is not available and to make educated estimates, percentage values are better then real values.

Internet purchases collected in real and percentage values

Please state the value of the purchases resulted from orders placed via Internet (in monetary terms, excluding VAT), in 2005.	*(National Currency)*
If you can't provide this value, **Please indicate an estimate of the percentage of the total purchases resulted from orders placed via Internet, in 2005.**	%

Therefore, currently the recommended method is to give the alternative to the enterprise to answer in real values if it has the information or, in case such information is not available, ask for an estimation of its percentage in relation to the turnover.

This intends first to reflect what the practice in several national surveys is already: some countries ask for the real 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 real 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 grossing-up of results will be different depending on how enterprise's Internet purchases are collected. If collected in percentage classes, then only the proportion of enterprises per class can be grossed-up. If collected in real or percentage values then real or the overall percentage can be grossed-up. Eurostat practice has been to get the grossed-up value of the percentage, i.e. the overall percentage of purchases made via Internet.

When grossing-up, the values collected as percentages need to be weighted by the total purchases of goods and services (variable X2). Please refer to chapter I.2.3.3 for more information on the weighting.

C3. Did your enterprise receive orders via the internet, during 2006 (excluding manually typed e-mails)?

[<u>Scope</u>: enterprises with Internet access, i.e. B1 = Yes]

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

C4. Please state the value of the turnover resulted from orders received via Internet (in monetary terms, excluding VAT), in 2006.

If you can't provide this value,
Please indicate an estimate of the percentage of the total turnover resulted from orders received via Internet, in 2006.

[<u>Scope</u>: enterprises which have received orders via Internet, i.e. C3 = Yes]

[<u>Type</u>: numerical]

Contrary to purchases, sales are usually centralised in enterprises. For this reason, Internet sales are collected in value (percentage or real) and not in percentage classes like it is the case for Internet purchases.

The problematic of collecting Internet sales in real or percentage terms are the same as for collecting Internet purchases. As in that case the recommended method is to collect it giving the possibility to the respondent to answer in real values and if that is not possible to provide an educated estimate in percentage of the total turnover.

C5. Was your enterprise using a secure protocol, such as SSL and TLS, for the reception of orders via Internet, in January 2007?

[<u>Scope</u>: enterprises which have received orders via Internet, i.e. C3 = Yes]

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

The measurement of the adoption of secure protocols is related to the measurement of the security level of the online environment for e-commerce over internet. A low level of security functions as a barrier to the further development of e-commerce.

A protocol is a set of rules for exchanging information between two or more computers over a network. Knowing the protocol and having access to the network allows in principal third parties to retrieve and read the messages exchanged. In order to provide secure communications over open networks, e.g. the internet, secure protocols use cryptographic mechanisms to assure that only the recipient of the message can read it and to authenticate the sender.

Secure Sockets Layer (SSL) and Transport Layer Security (TLS), its successor, are cryptographic protocols which provide secure communications on the Internet. SSL provides endpoint authentication and communications privacy over the Internet using cryptography. In typical use, only the server is authenticated (i.e. its identity is ensured) while the client remains unauthenticated. The protocols allow client/server applications to communicate in a way designed to prevent eavesdropping, tampering, and message forgery.

Common applications of the SSL/TLS secure protocols are https and S-HTTP (Secure hypertext transfer protocol). Web browsers use the HTTP protocol to communicate with web servers, sending and receiving information without encrypting it. For sensitive transactions, such as Internet e-commerce, the browser and server need to encrypt this information. The https and S-HTTP were both defined to address this need.

https is a scheme equivalent to the http scheme, originally intended to be used with the HTTP protocol, but with added encryption layer. Instead of using plain text communication, the session data is encrypted using either a version of the SSL (Secure Socket Layer) protocol or the TLS (Transport Layer Security) protocol, thus ensuring reasonable protection from eavesdroppers.

The relevant filter for this question is internet e-commerce, because it is only concerned with security in open networks for e-commerce. Additionally, it should be directed only to enterprises receiving orders via internet. The use of a secure protocol such as SSL/TLS requires an active role of both the computer which sends the message (e.g. order) and the one which receives it (they both need to be prepared to use the protocol). However, the "human" decision to use it lies with the enterprise receiving the orders, while the buyer, e.g. in the case of its use to receive orders via the web-site, will be simply using a browser SSL/TLS enabled, without having to do anything to establish that secure communication.

The question should make reference to internet and not the web-site, because the use of secure protocols is not restricted to the World Wide Web. Also in automated data exchange over the internet the use of secure protocols is fundamental to provide security to e-commerce transactions.

I.1.6.4 MODULE D: E-COMMERCE VIA COMPUTER NETWORKS OTHER THAN INTERNET

D1. *Did your enterprise order products/services via external computer networks other than Internet, during 2006?*

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

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

D2. *Please indicate for 2006 the percentage of the orders placed or transmitted via computer networks other than Internet, in relation to the total purchases (in monetary terms, excluding VAT).*

   or
   Please state the value of the purchases resulted from orders placed via computer networks other than Internet (in monetary terms, excluding VAT), in 2006.
   If you can't provide this value,
   Please indicate an estimate of the percentage of the total purchases resulted from orders placed via computer networks other than Internet, in 2006.

[Scope: enterprises which have ordered via other computer networks than Internet, i.e. D1 = Yes]

[Type: numerical]

Just like with orders placed via Internet, the value of the purchases made over computer networks other than Internet can be collected in 3 different ways: 1) in terms of classes of percentage of the total purchases; 2) in percentage value; and 3) in real value. The advantages and disadvantages of collecting this information in the several ways are the same as in the case of Internet purchases. Please refer to the explanatory note of question C2 for more in-depth information.

When collected in classes of percentages, the 5 response categories used are:

- Less than 1%

- 1% or more and less than 25%

- 25% or more and less than 50%

- 50% or more and less than 75%

- 75% or more

The classes are not the same as the ones used for Internet purchases, because although other computer networks are used by fewer enterprises, they are used more intensively.

Just like Internet purchases, currently the recommended method is to give the alternative to the enterprise to answer in real values if it has the information or, in case such information is not available, ask for a estimation of its percentage in relation to the turnover.

Equally, the grossing-up of results will be different depending on how enterprise's Internet purchases are collected. If collected in percentage classes, then only the proportion of enterprises per class can be grossed-up. If collected in real or percentage values then real or the overall percentage can be grossed-up. Eurostat practice has been to get the grossed-up value of the percentage, i.e. the overall percentage of purchases made via Internet.

When grossing-up, the values collected as percentages need to be weighted by the total purchases of goods and services (variable X2). Please refer to chapter I.2.3.3 for more information on the weighting.

D3. *Did your enterprise receive orders via external computer networks other than Internet, during 2006?*

[<u>Scope</u>: enterprises which use computers, i.e. A1 = Yes]

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

D4. *Please state the value of the turnover resulted from orders received via computer networks other than Internet (in monetary terms, excluding VAT), in 2006.*

If you can't provide this value,
Please indicate an estimate of the percentage of the total turnover resulted from orders received via computer networks other than Internet, in 2006.

[<u>Scope</u>: enterprises which have received orders via computer networks other than Internet, i.e. D3 = Yes]

[<u>Type</u>: numerical]

Contrary to purchases, sales are usually centralised in enterprises. For this reason, sales via computer networks other than Internet are collected in value (percentage or real terms) and not in percentage classes like it is the case for the purchases.

The problematic of collecting sales in real or percentage terms are the same as for collecting purchases. As in that case the recommended method is to collect it giving the possibility to the respondent to answer in real values and if that is not possible to provide an educated estimate in percentage of the total turnover.

I.1.6.5 MODULE E: E-SKILLS – ICT COMPETENCE IN THE ENTERPRISE UNIT AND THE DEMAND FOR ICT SKILLS

E1. *Did your enterprise employ ICT/IT specialists, in January 2007?*

    [<u>Scope</u>: enterprises which use computers, i.e. A1 = Yes]

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

E2: *How many ICT/IT specialists were employed by your enterprise, during January 2007?*
or
Please indicate an estimate of the percentage of the number of ICT/IT specialists in relation to the total number of persons employed, during January 2007.

    [<u>Scope</u>: enterprises which employ ICT/IT specialists, i.e. E1 = Yes]

    [<u>Type</u>: numerical, in real or percentage values]

The purpose of these two variables is to measure the incidence of ICT specialists' skills.

Two main types of e-skills can be distinguished. ICT specialists' skills include specifying, designing, developing, installing, operating, supporting, maintaining, managing, evaluating and researching ICT systems. ICT users skills 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 industry.

ICT specialists, or IT specialists, have the capability to specify, design, develop, install, operate, support, maintain, manage, evaluate and research ICT and ICT systems. ICT is the main job. It refers to the following ISCO-88 classification codes:

- 1236 Computing services managers
- 2131 Computer systems designers, analysts and programmers
- 2139 Computing professionals not elsewhere classified
- 2144 Electronics and telecommunications engineers
- 3114 Electronics and telecommunications engineering technicians
- 3121 Computer assistants
- 3122 Computer equipment operators
- 3132 Broadcasting and telecommunications equipment operators

E3. Did your enterprise recruit or try to recruit personnel for jobs requiring ICT specialist skills, during 2006?

[<u>Scope</u>: enterprises which use computers, i.e. A1 = Yes]

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

E4. Did your enterprise have hard-to-fill vacancies for jobs requiring ICT specialist skills, during 2006?

[<u>Scope</u>: enterprises which tried to recruit personnel for jobs requiring ICT specialist skills, i.e. E3 = Yes]

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

The aim of this question is to identify shortages in the labour market concerning ICT specialist skills.

E5. What do you believe were the main reasons of having hard-to-fill vacancies for jobs requiring ICT specialist skills, during 2006?

[<u>Scope</u>: enterprises which had hard-to-fill vacancies for jobs requiring ICT specialist skills, i.e. E4 = Yes]

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

For enterprise having had hard-to-fill vacancies for jobs requiring ICT specialist skills, the following reasons are colleted.

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

This item includes all those situations where the 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

This item includes those situations where the 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. In this item the objective is to identify cases where the qualifications of the candidates to the job vacancies 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. What distinguishes item a) from item b) is that the first one refers to quantity (not having people applying to the vacancy) and the second one 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

This item includes those situations where the 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

This item includes all those situations where the enterprise found difficulties in recruiting personnel with ICT specialist 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.

E6. Did your enterprise recruit or try to recruit personnel for jobs requiring skills in the use of ICT, during 2006?

[<u>Scope</u>: enterprises which use computers, i.e. A1 = Yes]

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

Skills in the use of ICT include those capabilities enabling the effective use of common, generic software tools (basic user skills) or advanced, often sector-specific, software tools (advanced user skills). ICT is an important tool for the job and is used to produce work output and/or is used intensively at work (in day-to-day activities). Common generic software tools refer in many cases simply to office tools. However, it also includes specific tools like, e.g., CAD.

E7. Did your enterprise have hard-to-fill vacancies due to applicants' lack of skills in the use of ICT, during 2006?

[<u>Scope</u>: enterprises which tried to recruit personnel for jobs requiring skills in the use of ICT, i.e. E6 = Yes]

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

This question includes all those situations where the enterprise tried to recruit personnel with ICT user skills, or skills in the use of ICT, and found difficulties in doing so because it was difficult to find a candidate with the right knowledge in the use of ICT which the enterprise needs.

E8. Did your enterprise provide training to develop or upgrade ICT related skills of your personnel, during 2006?

[<u>Scope</u>: enterprises which use computers, i.e. A1 = Yes]

[<u>Type</u>: 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 both ICT user skills and ICT specialist skills.

E9. Were any ICT functions requiring ICT/IT specialists performed by external suppliers (fully or partly), during 2006?

Definition External Suppliers: Other enterprises including foreign enterprises/legal entities. It includes enterprises belonging to the same enterprise group but excludes affiliates of the enterprise.

[<u>Scope</u>: enterprises which use computers, i.e. A1 = Yes]

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

The purpose of this question is to measure the proportion of enterprises outsourcing ICT functions of the enterprise.

In this manual in order to measure the outsourcing of ICT functions, we identify those situations where the enterprise has any of its operations or processes requiring ICT specialists performed outside the company by an independent third party, either in the country or abroad.

The term used in the question to refer to independent third party is "external supplier", which is defined as other enterprises, even if making part of the same enterprise group, and excluding affiliates. The basic idea used to identify outsourcing here is that of control. We have outsourcing if the tasks are performed by an entity over which the enterprise has no control. Affiliates of an enterprise are normally considered to be other enterprises or entities of which it owns more than 50% of the capital. Therefore, they are not considered in this question as an external supplier. Enterprises belonging to the same group and specialised in providing ICT services to the group might have autonomy or not. However, for simplicity and in order not to complicate too much the question, tasks performed by other enterprises of the same group is considered always as outsourcing.

E10. Were any ICT functions requiring ICT/IT specialists performed by suppliers in a foreign country (fully or partly), during 2006?

Definition Suppliers in a foreign country: Suppliers in a foreign country or foreign suppliers includes 1) foreign affiliates, usually legal entities, established by the enterprise (internal suppliers from abroad), and 2) other foreign enterprises (external suppliers from abroad).

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

[Type: one single answer needed, i.e. *Tick only one; filter question*]

The purpose of this question is to measure the proportion of enterprises off-shoring ICT functions of the enterprise.

In this manual in order to measure the off-shoring of ICT functions, we identify those situations where the enterprise has any of its operations or processes requiring ICT specialists performed in a foreign country, either by an independent third party or by an affiliate of the enterprise. It therefore includes outsourcing to an independent third party which is placed in another country and also having the tasks performed by an affiliate of the company which is located outside the country.

This question has 3 response alternatives, splitting the case where the suppliers abroad are foreign affiliates and where they are other foreign enterprises.

E11. Which ICT functions were performed by suppliers' ICT/IT specialists in a foreign country, during 2006?

[Scope: enterprises which had ICT functions performed by suppliers in a foreign country, i.e. E10 = Yes]

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

For those enterprises off-shoring ICT, information is collected on the off-shoring of several ICT functions.

a) ICT management (includes e-business and ICT systems management)

This item includes those processes related to the planning, organization and control of ICT resources of the enterprise.

b) ICT development and implementation

This item includes business software development, programming, web development, database development, communication network development and systems integration and installation.

c) ICT operations

This item includes technical support, user help and support, network administration, web administration and database administration.

d) Other ICT functions

This item includes all other tasks usually performed by ICT specialists.

E12. From which of the following geographical regions did your enterprise engage suppliers' ICT/IT specialists, during 2006?

[<u>Scope</u>: enterprises which had ICT functions performed by suppliers in a foreign country, i.e. E9 = Yes]

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

For those enterprises off-shoring ICT, information is collected on destination of the off-shoring: the European Union and the rest of the world.

For the 2007 community survey all 27 EU Member, as of 1st of January 2007 are considered: Belgium, Bulgaria, Czech Republic, Denmark, Germany, Estonia, Greece, Spain, France, Ireland, Italy, Cyprus, Latvia, Lithuania, Luxembourg, Hungary, Malta, Netherlands, Austria, Poland, Portugal, Romania, Slovenia, Slovakia, Finland, Sweden and United Kingdom.

E13. Were any business functions requiring users of ICT performed by external suppliers (fully or partly), during 2006?

[<u>Scope</u>: enterprises which use computers, i.e. A1 = Yes]

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

The purpose of this question is to measure the proportion of enterprises outsourcing business functions for which ICT user skills are important.

Just like in question E9, the outsourcing of these business functions is measured identifying those situations where the enterprise has any of its operations or processes requiring ICT user skills performed outside the company by an independent third party, either in the country or abroad.

E14. Were any business functions requiring users of ICT performed by suppliers in a foreign country (fully or partly), during 2006?

[<u>Scope</u>: enterprises which use computers, i.e. A1 = Yes]

[<u>Type</u>: one single answer needed, i.e. *Tick only one*; filter question]

The purpose of this question is to measure the proportion of enterprises off-shoring business functions requiring ICT user skills.

Just like in question E10, in order to measure the off-shoring of business functions requiring ICT user skills, we identify those situations where the enterprise has any of its operations or processes requiring ICT user skills performed in a foreign country, either by an independent third party or by an affiliate of the enterprise.

E15. Which business functions (non-ICT) were performed by suppliers' ICT users in a foreign country, during 2006?

[**Scope**: enterprises which had business functions requiring ICT user skills performed by suppliers in a foreign country, i.e. E14 = Yes]

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

For those enterprises off-shoring business functions requiring ICT user skills, information is collected on the off-shoring of several of these functions.

a) Sales and marketing, customer services

b) Research and development, product design and engineering

c) Other (non-ICT) business functions

E16. Please indicate the geographical regions from where you engaged business services requiring ICT users, during 2006.

[**Scope**: enterprises which had business functions requiring ICT user skills performed by suppliers in a foreign country, i.e. E14 = Yes]

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

For those enterprises off-shoring business functions requiring ICT user skills, information is collected on destination of the off-shoring: the European Union and the rest of the world.

For the 2007 community survey all 27 EU Member, as of 1st of January 2007 are considered: Belgium, Bulgaria, Czech Republic, Denmark, Germany, Estonia, Greece, Spain, France, Ireland, Italy, Cyprus, Latvia, Lithuania, Luxembourg, Hungary, Malta, Netherlands, Austria, Poland, Portugal, Romania, Slovenia, Slovakia, Finland, Sweden and United Kingdom.

I.1.6.6   MODULE X: BACKGROUND INFORMATION

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

Second, 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 number of persons employed is also used to weight the qualitative variables.[6]

Third, 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.[7]

The background variables 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 Rev. 1.1, first at the highest level of classification and then at more detailed levels ("top-down" method).

The nomenclature NACE Rev. 1.1 is available in Eurostat's RAMON database:

> http://europa.eu.int/comm/eurostat/ramon

The main economic activity of the enterprise should be classified by NACE Rev. 1.1 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 groupings
1	DA + DB + DC + DD + DE	Manufacture of products based on: food, beverages, tobacco, textile, leather, wood, pulp and paper; publishing and printing
2	DF + DG + DH	Manufacture of coke, refined petroleum products, chemical products, man-made fibres, rubber and plastics
3	DI + DJ	Manufacture of other non-metallic mineral products, basic metals and fabricated metal products
4	DK + DL + DM + DN	Manufacture of machinery and electrical, optical and transport equipment and other manufacturing n.e.c.

---

[6] For more information on the weighting refer to chapter I.2.3.3.
[7] For more information on the sampling refer to chapter I.2.3.1.

5	45	Construction
6	50	Sale, maintenance and repair of motor vehicles and motorcycles; retail sale of automotive fuel
7	51	Wholesale trade and commission trade, except of motor vehicles and motorcycles
8	52	Retail trade and repair of personal and household goods
9	55.1 + 55.2	Hotels and other provision of short-stay accommodation
10	60 + 61 + 62 + 63	Transport and storage
11	64	Post and telecommunications
12	72	Computer and related activities
13	70 + 71 + 73 + 74	Real estate, renting and business activities except computer activities
14	92.1 + 92.2	Motion picture, video, radio and television activities
F1	65.12 + 65.22	Monetary intermediation and other credit granting, except central banking
F2	66.01 + 66.03	Insurance, except compulsory social security

Optional:

17	22	Publishing, printing and reproduction of recorded media
18	40 + 41	Electricity, gas and water supply
19	55.3 + 55.4 + 55.5	Restaurants, bars, canteens and catering
20	92.3 to 92.7	News agencies, sporting and libraries, archives, museums and other cultural and recreational activities
21	93	Textile washing and dry-cleaning, hairdressing and other beauty treatment, funeral activities, physical well-being activities, etc.
F3	67.12 + 67.13 + 67.2	Activities auxiliary to financial intermediation, except administration of financial markets

These NACE categories are grouped together for dissemination purposes into several aggregates organised in 3 hierarchal levels. At the first level there are 3 categories distinguishing "Manufacturing, Energy and Construction", "Non-financial services" and "Financial Services". At the second level, activities are grouped at the Section level of NACE Rev. 1.1, making 9 categories. Level 3 is the most detailed one, composed of the 21 NACE groupings described above.

Level 0	Level 1	Level 2	Level 3
Total	Manufacturing, Energy and Construction	Section D	DA+DB+DC+DD+DE
			DF+DG+DH
			DI+DJ
			DK+DL+DM+DN
		Section E	40 + 41
		Section F	45
	Non-financial Services	Section G	50
			51
			52
		Section H	55.1 + 55.2
			55.3 - 55.5
		Section I	60 - 63
			64
		Section K	72
			70 +71 +73 +74
		Section O	92.1 + 92.2
			92.3 - 92.7
			93
	Financial Services	Section J	65.12 + 65.22
			66.01 + 66.03
			67.12 + 67.13 + 67.2

Legend:
- Optional category
- Non-optional category
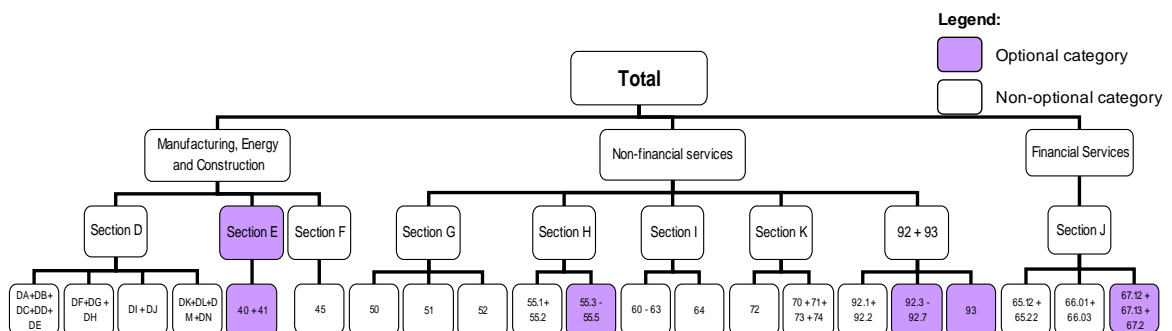
X2. Average number of persons employed

[<u>Scope</u>: all enterprises]

[<u>Type</u>: 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.
>
> **Source:** Structural Business Statistics Regulation (variable 16 11 0)

The number of persons employed should be measured as the yearly average during the previous calendar year. However, if the seasonal effect is not significant, the last day of the previous calendar year can be used.

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 hierarchal way into 3 levels. The first level distinguishes micro-enterprises from the others. The second and thirds levels further breakdown these two groups.

Size Categories (persons employed)		
1	10 to 49	Small enterprises
2	50 to 249	Medium enterprises
3	250 or more	Big enterprises
Optional:		
4	Less than 5	Small micro-enterprises
5	5 to 9	Big micro-enterprises

Level 0	Level 1	Level 2	Level 3
Total	Micro-Enterprises	Small Micro-Enterprises	Small Micro - Enterprises
		Big Micro-Enterprises	Big Micro - Enterprises
	Small, Medium and Large Enterprises	Small and Medium Enterprises	Small Enterprises
			Medium Enterprises
		Large Enterprises	Large Enterprises

Legend:

Optional category

Non-optional category
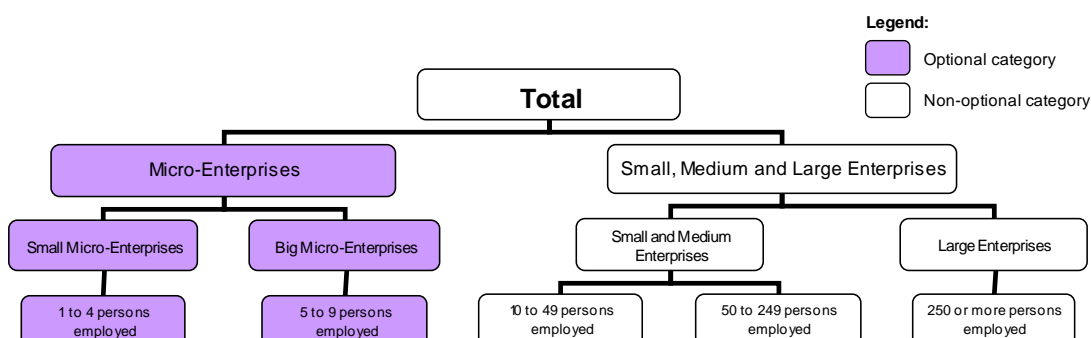
X3. Total purchases of goods and services

[Scope: all enterprises]

[Type: numerical]

This variable is not to be collected for the financial enterprises, i.e., enterprises which main economic activity is classified in Section J of NACE Rev. 1.1, because the concept is only applicable to non-financial economic activities.

This variable is needed to weight the percentage of purchases resulted from orders placed via Internet and other computer networks. However, if these variables are collected in classes, instead of percentage (or real) values, then this background variable is not necessary.

As for the number of persons employed and turnover, 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 excluding deductible VAT and other deductible taxes linked directly to turnover.

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.

Specific calculation methods are needed for NACE Rev.1 classes 66.01 and 66.03

Source: Structural Business Statistics Regulation (variable 13 11 0)

X4. Total turnover

[Scope: all enterprises]

[Type: numerical]

Just like the total purchases of goods and services, this variable is not to be collected for the financial enterprises, i.e. enterprises which main economic activity is classified in Section J of NACE Rev. 1.1, because the concept is only applicable to non-financial economic activities.

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

> 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.

> Turnover includes all duties and taxes on the goods or services invoiced by the unit with the exception of the VAT invoiced by the unit vis-à-vis its customer and other similar deductible taxes directly linked to turnover.

> 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 extraordinary income in company accounts is excluded from turnover. Operating subsidies received from public authorities or the institutions of the European Union are also excluded.

> For NACE Rev. 1 classes 66.01 and 66.03, the corresponding title of this characteristic is 'Gross premiums written'.

> **Source:** Structural Business Statistics Regulation (variable 12 11 0)

X5. Location (Convergence/ non-Convergence region)

[Scope: all enterprises]

[Type: categorical]

The "Convergence objective" 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.

For 2007, the list of objective 1 regions (valid until end 2006) is substituted by the new list of convergence regions (including related phasing out regions) eligible for funding from the Structural Funds.

The NUTS regions (statistical demarcation) eligible for support from the Structural Funds under the convergence objective (ex-objective 1) - including phasing-out regions, has been decided for the whole period 2007-2013. No update of this list is foreseen in the Structural Funds Regulation. The list is shown below.
Please note that the list is different for the one used in the 2006 survey.

Countries *entirely* composed of Convergence (ex-Objective 1) regions:

Bulgaria

Estonia (the whole country counts as one single region at NUTS2)

Ireland (Southern and Eastern under transitional support)

Latvia (the whole country counts as one single region at NUTS2)

Lithuania (the whole country counts as one single region at NUTS2)

Malta (the whole country counts as one single region at NUTS2)

Poland

Romania

Slovenia (the whole country counts as one single region at NUTS2)

Countries partially composed of Convergence (ex-Objective 1) regions:

Belgium: Hainaut

Czech Republic: Střední Čechy, Jihozápad, Severozápad, Severovýchod, Jihovýchod, Střední Morava, Moravskoslezsko (i.e. the whole country except Praha)

Germany: Brandenburg Nord-Ost, Brandenburg Süd-West, Mecklenburg-Vorpommern, Lüneburg, Chemnitz, Dresden, Leipzig, Dessau, Halle, Magdeburg, Thüringen

Greece: Anatoliki Makedonia, Thraki, Kentriki Makedonia, Dytiki Makedonia, Thessalia, Ipeiros, Ionia Nisia, Dytiki Ellada, Peloponnisos, Attiki, Voreio Aigaio, Kriti

Spain: Galicia, Principado de Asturias, Castilla-La Mancha, Extremadura, Andalucía, Región de Murcia, Ceuta, Melilla

France: Guadeloupe, Martinique, French Guyana, Réunion

Italy: Campania, Puglia, Basilicata, Calabria, Sicilia

Hungary: Közép-Dunántúl, Nyugat-Dunántúl, Dél-Dunántúl, Észak-Magyarország, Észak-Alföld, Dél-Alföld

Austria: Burgenland

Portugal: Norte, Algarve, Centro, Alentejo, R. A. Açores

Slovakia: Západné Slovensko, Stredné Slovensko, Východné Slovensko (i.e. the whole country except Bratislavsky kraj)

United Kingdom: West Wales and the Valleys, Cornwall and Isles of Scilly, Highlands and Islands

Countries with *no* Convergence (ex-Objective 1) regions:

Denmark (the whole country counts as one single region at NUTS2)

Cyprus (the whole country counts as one single region at NUTS2)

Luxembourg (whole country counts as one single region at NUTS2)

The Netherlands

Finland

The list of convergence regions was published in 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"

I.2. PRODUCTION METHODOLOGY

I.2.1. Timetable – Survey period

The survey period of the Community Survey on ICT Usage and e-Commerce in Enterprises is the first quarter of calendar year t. The variables in the questionnaire are predominantly of a qualitative nature (yes-or-no-questions or tick-box-questions). The answers to these questions refer to the situation in January of year t. The answers to the quantitative questions - for example the percentage of total turnover realised by sales over the Internet - refer to the calendar year t-1.

In order to minimize the recall-bias of the respondent the collection should occur just after the reference period, after January during the first quarter. However, some countries have difficulties in collecting monetary data (turnover and purchases) right in 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 periods (2005 survey)
Source: Final reports for the Community Survey on ICT Usage and e-Commerce in Enterprises (2005)

Belgium	Second half of March 2005 – Mai 2005
Czech Republic	Second quarter 2005
Denmark	February-April 2005.
Germany	First quarter 2005
Estonia	The questionnaires were sent out in March 2005. The deadline was the 1 of May 2005
Greece	The survey started on March 2005.
Spain	Data collection was carried out on the second quarter 2005
Ireland	April 2005
Italy	April-September 2005.
Cyprus	Data collection period: 28/1/2005 – 18/3/2005
Latvia	Survey period was May 2005
Lithuania	First quarter 2005
Luxembourg	April-May 2005
Hungary	First quarter of 2005.
Malta	February to March 2005 as planned.
Netherlands	February 2005: sending out the questionnaires, general survey
Austria	Started on 2nd February 2005; Last questionnaire received on 23rd May 2005
Poland	From 1st April till 30th April
Portugal	Data collection started in April and finished in November.
Slovenia	7th March 2005 – 31st May 2005
Slovak Republic	Questionnaires sent in March. Required date of questionnaire return was April 2005
Finland	Mailing started beginning of February; Most responses received in May at the latest.
Sweden	Questionnaires sent 23rd March 2005. Data collection lasted until the beginning of July.
United Kingdom	The survey was conducted during 2005
Norway	Paper questionnaires sent and web questionnaire available 9 March 2005. Returned paper questionnaires no longer accepted, Web questionnaire taken off Internet 21 May 2005.

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 which main economic activity is agriculture (outside the scope of the Community Survey) in the sampling frame: if this sector presents 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 ICTs, the overall indicators may be relatively overestimated for such country.

The sample for the Community Survey should be drawn from the business register in the different Member-States as defined in Council Regulation (EEC) No 2186/93. Part of this register is the activity code at the four-digit (class) level of NACE Rev.1, the size measured by the number of persons employed and the geographical location code (territorial units) 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 meso-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 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 2005 survey.

Table: Overview of the sampling frames (2005 survey)
Source: Quoted from the Final reports for the Community survey on ICT usage and e-commerce in enterprises (2005)

Belgium	Frame population is derived from SBS frame population: elimination of enterprises with < 5 persons employed or NACE different from ICT-frame. Some SBS reporting units coupled if they belong to the same enterprise in the business register.
Czech Republic	The frame population is the same as the one for the Structural Business Survey (SBS).
Denmark	The frame population corresponds with SBS.
Germany	The sampling base of our survey is the active enterprise register.
Estonia	The frame is the same as for short term business statistics as main questions are about January 2005. The same frame is also used for SBS, the only difference is that enterprises, which were active only part of 2004 and are not active in 2005 are included in 2004 SBS survey but not in the ICT survey.
Greece	The sampling frame used was based on data coming from the Business Register of NSSG. This Business Register was set up from data coming from tax authorities and social security funds. Moreover, this register was updated with data obtained by recent statistical surveys of NSSG. The frame population is the one used for the SBS.

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.
Ireland	The CSO Central Business Register provides the frame population. The Central Business Register is also used by the SBS surveys.
Cyprus	The frame population is the Registration of Establishments 2000 which is updated on the basis of the Employment Survey and other resources so as to take into account changes in the structure of employment by industry, occupation and size of establishment. The frame population is used for the Structural Business Statistics too.
Latvia	The statistical business register holds information on all enterprises that are registered in the administrative Enterprise Register of the Republic of Latvia. In the sampling frame only economically active enterprises from the business register are included and it is the same as used for the Structural Business Statistics.
Lithuania	The statistical unit of survey are economically active enterprise as defined in the Statistical Business Register.
Luxembourg	For the sample design, the enterprises register was used.
Hungary	The frame population is almost the same as the one used for the Structural Business Statistics. The underlying register concerning the target population is the Business Register of Enterprises. The observation and sampling units consist of those enterprises that have an obligatory data provider status in December 2004 concerning the following surveys: "Simplified monthly report on the industrial activities", "Simplified monthly report on the construction activities", "Simplified monthly report on the agricultural, trading and service activities". 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 chosen sample elements concerning the above mentioned surveys.
Malta	The frame population is taken from the Business Register section of the 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 1-12-2004.
Austria	Frame population is the enterprises in the National Business Register. SBS use the same source for sampling.
Poland	-
Portugal	The frame population of this survey is the reference universe of the Structural Business Statistics, which is taken from the national business registers.
Slovenia	The sampling frame was made by using the data from the Business Register of the Republic of Slovenia and the data from some other surveys.
Slovak Republic	Frame population was based on the same Register of statistical units as in case of structural business statistics.
Finland	The frame population is drawn from the production database of Structural Business Statistics.
Sweden	The sampling frame is the Swedish Business Register.
United Kingdom	Whole UK economy taken from the ONS business register, based on SBS.
Norway	The frame population is based on the Structural Business Statistics.

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 coefficient of variation (or relative standard error) should not exceed 2% for the overall proportions and should not exceed 4% for the proportions relating to the different subgroups of the population where these subgroups constitute at least 5% 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.

Table: Overview of the sampling design and sampling methods (2005 survey)

Source: Quoted from the Final reports for the Community Survey on ICT Usage and e-Commerce in Enterprises (2005)

Belgium	Stratified random sample: ▪ Stratification by Activity (17 categories), Size class (4 categories) and NUTS1 region (3 categories), resulting in 204 strata. ▪ Exhaustive selection in strata with less than 5 enterprises. Oversampling of smaller regions. Sample selection = SRS algorithm in SPSS. No procedures for the coordination of non-overlapping with other samples.
Czech Republic	Stratified random sampling for enterprises with 5-249 and census for 250 and more employees with respect to Business Register was used. Number of enterprises and number of employees was used for stratification. GREG was used for sample design. The sample was designed with no reference to any other survey. Sample size was designed to enable accurate, reliable and representative results for breakdown into four size groups, 32 NACE aggregates groups and two territory groups and for main variables as a cross-tabulated.
Denmark	The sample is a stratified random sample made by use of register-based data. Strata was made by: ▪ employment (10-19, 20-49, 50-99, 100+) and ▪ activity (Nace, 2-digit)
Germany	Germany used a two-stage stratified random sample. The stratified variables were federal states (first stage), economic activities (NACE) and number of employees (second stage).
Estonia	The data are collected by sample survey, using stratified simple random sampling. The sample is stratified by economic activity and by number of persons employed. By number of persons employed, the survey includes all enterprises with 50 and more persons employed. Stratified simple random sampling is used for enterprises 10 to 49 persons employed. For data processing the enterprises were divided by number of persons employed into following size groups: 10-19, 20-49; 50-99; 100-249; 250-499; 500-999; 1000+ 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 structure of business survey. The number of strata is 263.
Greece	The sampling method applied was the simple stratified random sampling, employing the enterprise as the survey unit. The enterprises with 10 or more employees included in the survey were stratified as following: ▪ By geographical region – NUTSI ▪ By aggregations of NACE Rev.1 subsections, as defined in the relevant contract, within each geographical region ▪ By size class of the enterprise. In each one of the major strata (major stratum = Geography x Economic Activity), the enterprises were stratified into L=7 size classes, according to their size, determined by their average annual number of employees in the business register, as follows. <table><tr><td>Size Class</td><td>Number of employees</td></tr><tr><td>Class 1</td><td>10-19</td></tr><tr><td>Class 2</td><td>20-49</td></tr><tr><td>Class 3</td><td>50-99</td></tr></table>

Class 4	100-249
Class 5	250-499
Class 6	500-999
Class 7	1000+

In one geographical region, for one economic activity, let h be one size class. So h will take the following values: h=1,2,…,7. In the initial sampling design, the size classes with h=3,4,5,6,7 were census strata (take-all), because they contain the large size units and the element population variance is high. In each stratum (let h), a sample of nh enterprises was selected from the total number of enterprises Nh with equal selection probabilities, by applying the systematic sampling method. In the size classes with h=1,2, the distribution of sampling units was conducted by applying the Neyman technique (optional allocation). The final sample size (survey in Nace Groups except Nace J) was 4.668 enterprises (sampling rate = 26,01%), of which the 1.787 enterprises belong to the first two size classes (they were surveyed on a sample basis) and the rest 2.881 enterprises belong to the census size classes.

The whole sample size was defined so that the results to reflect the true situation of the total population units with a sufficient degree of representativeness.

Finally, concerning the aggregations 65.12+65.22, 66.01+66.03 and 67.12+67.13+67.2 (separate survey for Nace J), these were also exhaustively surveyed (118 enterprises).

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, 10 to 49, to 199, 200 or more

- **main activity:** following the aggregation of NACE classification as follow:
  GENERAL QUESTIONNAIRE:
  1)DA+DB+DC+DD+DE
  2)DF+DG+DH
  3)DI+DJ
  4)DK+DL+DM+DN
  5)E (40+41)
  6)F (45)
  7)G (50)
  8)G (51)
  9)G (52)
  10)H (55.1+55.2)
  11)I (60+61+62+63)
  12)I (64)
  13)K (72)
  14)K (70+71+73+74)
  15)O (92.1+92.2)
  FINANCIAL SECTOR:
  1)J (65.12+65.22)
  2)J (66.01+66.03)
  3)J (67.12+67.13+67.2)

- **region:** all Spanish regions (19) are covered (except for the pilot survey referred to enterprises with less than 10 employees).

The total number of initial strata is 1.710 (18*5*19), including empty strata.

Ireland

Stratified random sample. NACE (2 digit level) and employment size class (10-49, 50-249 and 250+) were the variables that were used to stratify.

In total there were 136 strata.

Where possible, to assist reduce the response burden on small and medium size enterprises, we try to minimise the possibility of these enterprises being included in both the SBS and ICT surveys.

Cyprus

The sampling method used was stratified random sampling. Two variables were used for stratification, NACE group and size. There were 14 NACE groups and 3 size groups.

Due to the small number of large and medium enterprises it has been decided to use all those enterprises in the sample. Half of the small enterprises were selected systematically with probability proportional to size from a list ordered in ascending order by enterprise size.

Sixty additional enterprises were included in the sample in April 2005, in order to update the sample and include almost all enterprises from smaller NACE groups. These enterprises belong to NACE groups: 23-25, 50, 64, 72, and 92.1-92.2.

The size of the sample, i.e. 1240 enterprises, could satisfactorily serve the desirable and acceptable confidence limits of the survey results, taking always into consideration time and cost constraints.

The total number of all financial enterprises falling in the 3 NACE groups, for the size classes 10-49, 50-249, 250+ is small (113 enterprises), thus all enterprises are used.

Latvia	The sample is made as stratified simple random sample. The stratification is made according to the main line of economic activity of enterprise (according NACE breakdown shown in general outlines of ICT survey), the territorial location of enterprise (6 regions) and enterprise size group (5 groups). Size groups were defined by turnover, number of persons employed, value added tax paid, balance and other variables. The final number of strata was 597 (considering also other NACE activities for national needs).
	The sample allocation is made using proportional allocation. There are different proportions for each size groups of enterprises. The proportions depend on population size of strata. The proportions are higher for strata with smaller population size.
	The sample for this survey is made as sub-sample from Structural Business Statistics survey. The sampling fraction for the biggest enterprises was 100%.
Lithuania	It was used stratified one-stage random sampling.
	The number of employees was used to stratify population. Categories of variables:
	▪ number of employees: 10-49; 50-249; 250 and more.
	The final number of strata – 54.
	The Neyman optimal allocation was used for sample selection and determination of sample size.
	If stratum size is 20 or less then all enterprises in this stratum are included.
	All financial enterprises are covered in the survey.
Luxembourg	A census have been made.
Hungary	The sample selection was done with a stratified sampling scheme. The strata are formed in the following way.
	Concerning the NACE-codes all the 2-digit level branches are being treated as a different stratum. We break down some 2-digit level branches to smaller sub-branches, if there is a need for a separate treatment considering the importance or the size of these areas. This is being done by this way.
	Branch Sub-branches
	14 14.21 and others
	22 22.22 and others
	28 28.11 and others
	29 29.24 and others
	36 36.14 and others
	37 37.10 and others
	45 45.1, 45.21, 45.25, 45.22-45.24, 45.3, 45.4, 45.5
	50 50.10 and others
	51 51.70 and others
	55 55.1-55.2, 55.30, others from 55.3-55.5
	60 60.24 and others
	63 63.40 and others
	80 80.42 and others
	Concerning the number of employees the following categories are being treated separately: 40 (20-49 employees), 30 (10-19 employees), 22 (5-9 employees).
	The capitol Budapest and the provinces are forming different strata.
	The distribution of the sample among the above defined strata is being done by the following. By the cooperation with the informatics department we know the number of enterprises belonging to the sample frame in all the single strata. This information is available for the time periods January 2002 and January 2003 (in the latter case provisional data are available).
	The final sample sizes are determined by the well-known Neumann-type optimal allocation. For this procedure we make use of the information available for the previous time-period concerning the stratum level statistics for different variables.
	The process of sample selection is done by the following. 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 choosing those that possess the biggest values. The underlying concept of the distortion process can be described in this way. We want to include firstly those enterprises which were in the sample in 2002 but not in the sample between 1998 and 2000. Secondly, those enterprises which were not included in the sample neither in 2002 nor between 1998 and 2000. Thirdly those sample elements from 2002 which were in the sample between 1998 and 2000 too. Fourthly, those which were in the sample between 1998 and 2000 but not in the sample in 2002.
Malta	For size class FTE 50-249 and 250 and over, all enterprises were covered. For size class 10 to 49 FTE, 4 divisions were limited to 50 respondents.
	Regarding the micro-enterprises only, that is from 1-4 FTE and from 5-9 FTE size classes, this was approached by a quota structure of 50 maximum, but with aggregated NACE for Manufacturing D as a whole, and Motor, Wholesale, Retail Trade also taken as one cluster. See Excel file for the sampling scheme for this purpose.

Netherlands	The target-variables (of year t-1) used to allocate the sample are:
	▪ the number of enterprises receiving orders online (a qualitative variable);
	▪ - the number of fte's spend on developing own account software (a quantitative variable).
	The population/sample was divided into 56 NACE-aggregates and 7 groups of size-classes resulting in 392 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. By choosing a maximum number of 12 000 enterprises to be allocated a response percentage of 50 percent is anticipated. In principle the Neyman-allocation was used to allocate the 12 000 enterprises over the different strata, meaning that the allocation is inversely proportional to the variance of the target-variable in the relevant strata.
	A rotation-fraction of .5 was used, meaning that 50 percent of the sample – if possible – exists of enterprises which also in the year t-1 were included in the sample. A system to 'spread' the administrative burden for enterprises is used to avoid - if possible - that certain enterprises receive a lot of 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	Stratified random sample; 2 dimensions were used as stratification variables:
	▪ Main economic activity: **8 strata:** sections D, F, G, H (only 55.1 and 55.2), I, J, K, and 92 (only 92.1 and 92.2)
	▪ Size classes **3 strata:** 10-49 employed persons, 50-249 employed persons, more than 250 employed persons
	8 x 3 = 24 strata
	A full census was carried out among enterprises with 250 or more persons employed.
Poland	Firstly, some domains of economic activity and large enterprises (with more than 249 employees) were taken wholly to the sample. This gave 3633 units. Secondly to start rotation, enterprises from the last year sample that were present in the 2005 frame were also taken to the sample. This gave 4508 units . Thirdly some cross strata of economic activity and size that contained small number of enterprises were also taken wholly to the sample. They were 118 units. Thus we obtained 8259 units in the sample. The remaining units were allocated using method of equal precision in the domain (of economic activity) and Neyman allocation among the strata (size*won –administrative unit). This way 5265 of units were allocated. Finally just before starting the survey (in March) sample was corrected by removing some not active enterprises and adding new units because of some small changes in population.
Portugal	The sampling method used was the stratified random sample. Two variables were used to stratify the sample:
	▪ NACE Rev.1.1 in the following domains: DA, DB, DC, DD, DE, DF, DG, DH, DI, DJ, DK, DL, DM, DN, 45, 50, 51, 52, I (60-63), I (64), K (72), K (70+71+73+74), O (921+922);
	▪ Number of employed persons in the following domains: 1-4; 5-9; 10-49; 50-249; 250 and more.
	For large enterprises a census is adopted.
Slovenia	Design of the sample was stratified systematic sample. Activity group and size class were used as stratification variables.
	In addition, implicit stratification by 5–digit NACE group was used.
	No procedures for coordinated samples were applied.
Slovak Republic	The sample was created on the base of the same methods, which were used in other surveys organised by SO SR (short-term surveys, structural business surveys). In comparison with other surveys the boundary between exhaustive and sample survey was changed in the direction upwards.
	In this survey on enterprises 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.
	Two-stage sampling was realised:
	NACE
	Number of employees
	In the sample the level of NUTS3 was applied as well. In spite of that due to the great number of strata this criterion was not used for grossing-up.
Finland	Stratified random sample (in enterprises with at least 100 employees a census).
	The variables for stratification were the size of the enterprise (as number of persons employed) and economic activity.
	▪ The categories of size were 5: 5-9, 10-19, 20-49, 50-99 and 100+.
	▪ The categories of economic activity were 26: as NACE classes 15-19, 20, 21, 22, 23, 24-25, 26, 27-28, 29-33, 34-35, 36, 37, 40, 45, 50, 51, 52, 55.1-55.2, 55.3-55.5, 60-63, 64, 70-71, 72, 73-74, 92.1+92.2+92.4+92.71, 93.
	All together this is 130 strata.
	All enterprises in the sample frame employing at least one hundred persons were included in the

	inquiry. Sampling was applied to enterprises with five to 99 employees. As a starting point in constructing the sample, one half of the enterprises employing 50 to 99 persons, one fifth of the enterprises with 20 to 49 employees and ten per cent of those employing 5 to 9 and 10 to 19 persons were selected by size category from the strata. This preliminary sample structure was still modified by strata to include more enterprises where the number of enterprises would be small and analytical purposes relating to economic activity needed more enterprises to the sample. Also in strata where the number of enterprises was regarded large relative to the analytical needs the sample size was reduced. In financial sector all enterprises are covered, except in two smallest size groups in NACE 65120 a sample is drawn to cover 30 per cent of the enterprises with 10-19 employees and 40 per cent of the enterprises in enterprises with 20-49 employees.
Sweden	Stratified random sampling is used, where size and economic activity are the variables used for stratification. Size : 10-19, 20-49, 50-99, 100-199, 200-249, 250-499, 500- Economic activity : 15-22, 23-25, 26-28, 29-37, 40-41, 45, 50, 51, 52, 55.1-55.2, 55.3-55.5, 60-63, 64, 65.12 and 65.22, 66.01 and 66.03, 67.12-67.13 and 67.2, 65.11 and 65.21 and 65.23 and 66.02 and 67.11, 70-71 and 73-74, 72, 92.1-92.2, 93 (NACE) The final number of strata (containing at least one enterprise) : 149 The sample size is determined by the precision 0.07 and "worst case" variances (that's why the precision is as low as 0.07). As allocation method, we used optimum allocation (Neyman allocation). The sample is positively coordinated with the sample drawn for the Survey of ICT Usage and e-Commerce in Enterprises 2004.
United Kingdom	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 can have information more readily available in the case of complaints, and know where to position new surveys on the line. However, the larger the business the more significant it is and therefore the more ONS questionnaires it receives. To minimise burden on the 10-49 employment businesses we set our PRN start point to exclude any businesses selected by the Annual Business Survey. 0-9 employment businesses are only selected for 1 year in 3 for any ONS survey. The following strata are used based on reported employment, held on our business register (IDBR). 0-9 random sampling 10- 49 random sampling 50 - 249 random sampling 250 - 999 random sampling 1000+ Fully enumerated strata
Norway	The sample was drawn from all enterprises with at least five employees in the Central Register of Establishments and Enterprises. It was stratified by industry and size of the enterprise measured by employment. • Enterprises with 5-9 employees the sample was 5 per cent, • Enterprises with 10-19 employees the sample was 7,5 per cent, • Enterprises with 20-49 employees the sample was 15 per cent, • Enterprises with 50-99 employees the sample was 50 per cent, • Enterprises with 100-249 employees the sample was 75 per cent, and • Enterprises with at least 250 employees all enterprises were included in the sample. The same sample percentages were used within all of the selected industries.

I.2.3.1 <u>STRATIFICATION</u>

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. The *"Location of the Enterprise"* (variable X5) can also used as a stratification variable (if an objective 1 / non-objective 1 breakdown is provided, or any other geographical breakdown, than the stratification by this variable is necessary). These

variables, according to the Council Regulation (EEC) N° 2186/93 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, size class and location 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 which casually end up in some NACE category or size class 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.1.6.6). 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 possibly 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/or 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 and 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 Community Survey (48 strata). The "x", in normal text, represents the minimum level of stratification for the optional scope of the survey (an additional 57 strata).

				Total				
				Micro-Enterprises		Small, Medium and Large Enterprises		
				Small Micro-Enterprises	Big Micro-Enterprises	Small and Medium Enterprises		Large Enterprises
				Small Micro - Enterprises	Big Micro - Enterprises	Small Enterprises	Medium Enterprises	Large 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	Manufacturing, Energy and Construction	Section D	DA+DB+DC+DD+DE	x	x	**x**	**x**	**x**
			DF+DG+DH	x	x	**x**	**x**	**x**
			DI+DJ	x	x	**x**	**x**	**x**
			DK+DL+DM+DN	x	x	**x**	**x**	**x**
		Section E	40 + 41	x	x	x	x	x
		Section F	45	x	x	**x**	**x**	**x**
	Non-financial Services	Section G	50	x	x	**x**	**x**	**x**
			51	x	x	**x**	**x**	**x**
			52	x	x	**x**	**x**	**x**
		Section H	55.1 + 55.2	x	x	**x**	**x**	**x**
			55.3 - 55.5	x	x	x	x	x
		Section I	60 - 63	x	x	**x**	**x**	**x**
			64	x	x	**x**	**x**	**x**
		Section K	72	x	x	**x**	**x**	**x**
			70 +71 +73 +74	x	x	**x**	**x**	**x**
		Section O	92.1 + 92.2	x	x	**x**	**x**	**x**
			92.3 - 92.7	x	x	x	x	x
			93	x	x	x	x	x
	Financial Services	Section J	65.12 + 65.22	x	x	**x**	**x**	**x**
			66.01 + 66.03	x	x	**x**	**x**	**x**
			67.12 + 67.13 + 67.2	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, size class and geographical location.

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 then for variables like for example production value.

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 percent 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 Final Reports for the 2005 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. 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 coefficient of variation) are discussed in more detail in the next chapters.

Table: Overview of the sample sizes and response rates (2005 survey)

Source: Final reports for the Community survey on ICT usage and e-commerce in enterprises (2005)

	Gross sample size	Ineligible (out-of-scope)	Number of eligible elements	Non-response	Net sample size	Unit response rate	Usable sample rate
Belgium	7,529	72	7,457	2,768	4,689	62.9%	62.3%
Czech Republic	10,325	529	9,796	2,281	7,515	76.7%	72.8%
Denmark	4,261	:	:	:	:	:	:
Germany	39,229	28	39,201	21,170	18,031	46.0%	46.0%
Estonia	3,440	0	3,440	568	2,872	83.5%	83.5%
Greece	4,786	209	4,577	2,044	2,533	55.3%	52.9%
Spain	20,467	2,255	18,212	2,812	15,400	84.6%	75.2%
Ireland	5,135	56	5,079	2,455	2,624	51.7%	51.1%
Italy	32,919	1,354	31,565	14,072	17,493	55.4%	53.1%
Cyprus	1,353	28	1,325	136	1,189	89.7%	87.9%
Latvia	7,398	7	7,391	823	6,568	88.9%	88.8%
Lithuania	2,512	65	2,447	152	2,295	93.8%	91.4%
Luxembourg	2,472	71	2,401	1,240	1,161	48.4%	47.0%
Hungary	9,677	0	9,677	2,412	7,265	75.1%	75.1%
Malta	753	3	750	15	735	98.0%	97.6%
Netherlands	8,339	0	8,339	2,416	5,923	71.0%	71.0%
Austria	4,700	86	4,614	1,416	3,198	69.3%	68.0%
Poland	14,904	429	14,475	2,703	11,772	81.3%	79.0%
Portugal	6,068	403	5,665	1,356	4,309	76.1%	71.0%
Slovenia	1,996	17	1,979	369	1,610	81.4%	80.7%
Slovak Republic	4,684	69	4,615	632	3,983	86.3%	85.0%
Finland	4,854	167	4,687	1,169	3,518	75.1%	72.5%
Sweden	4,000	39	3,961	1,232	2,729	68.9%	68.2%
United Kingdom	11,951	2,498	9,453	1,551	7,902	83.6%	66.1%
Norway	3,298	0	3,298	735	2,563	77.7%	77.7%

Note: Usable sample rate equals net sample size divided by gross sample size.

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 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. *"Do you use Internet?"*; and
2) Quantitative variables, e.g. *"Please indicate the percentage of persons employed with access to the Internet."*.

The enterprise ICT Community Survey is dominated by <u>qualitative questions</u> 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.

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 <u>quantitative questions</u> are in general collected in percentage terms, e.g. percentage of persons employed using computers, percentage of turnover resulted from sales received via Internet. 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.).

When the quantitative variables are collected in real 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:

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

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

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 real 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:

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

For each stratum:

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

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

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

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 real 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 real values, "number of persons employed, for enterprises with Internet access".

One way of doing it for the real 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:

$$Y = \sum Y_h \text{ ,}$$

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

where Y_h is the estimated real 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:

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

Which simplifies to:

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

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 real values, it refers to the numerator only: "number of persons employed with access to Internet".

One way of doing it for the real 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:

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

where Y_h is the estimated real 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:

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

Which simplifies to:

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

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.

I.2.4. Survey type

Data collection method

Face-to-face interviews, telephone interviews and postal 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.

Personal interviews (face to face)

Strengths:

Interviews are a far more personal form of research than questionnaires. 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 publicity-listed telephone numbers. Some don't have telephone. People often don't like the intrusion of a call to their homes. Telephone interviews have to be relatively short or people will feel imposed upon. The inability to use visual aids is also a weakness.

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 increase 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. The table below presents the information on the data collection methods used in the 2005 national surveys.

Table: Overview of the data collection methods (2005 survey)
Source: Final reports for the Community survey on ICT usage and e-commerce in enterprises (2005)

Belgium	Stand-alone survey. Answer was voluntary. Self-administered postal mail survey.
Czech Republic	Stand-alone survey. Answer was mandatory. Postal (self-administered mail out and mail or e-mail back) survey Electronic questionnaire (computer assisted program available on the web site) for electronic completion was developed.
Denmark	Stand-alone survey. Answer was mandatory. Self-administered mail survey. Postal questionnaire.
Germany	Stand-alone survey. Answer was voluntary. Postal mail survey.
Estonia	Stand-alone survey. Answer was mandatory. Mail survey and the possibility for enterprises to get the questionnaire from Statistical Office homepage. Only in case of problematic answers, respondents were contacted by phone.
Greece	Stand-alone survey. Answer was mandatory.

	The survey types used were: • Face to face interviews (in Athens and in Thessaloniki) • Self-administered mail survey
Spain	Stand-alone survey. Answer was mandatory and enterprises can be penalized in case of no response. Firstly, questionnaires are sent by mail to the enterprises in the sample. Moreover, enterprises may phone to a free-charge telephone number to get clarification of the questions and to be helped with the filling out of the questionnaire Also, questionnaires might be filled in over the Internet. In all cases there is a system of reminders, including letters and phone calls.
Ireland	Stand-alone survey. Answer was voluntary. The survey is a self-administered mail survey
Cyprus	Stand-alone survey. Answer was mandatory. Face-to-face interview. The questionnaire was sent by mail to the IT manager of the enterprise just before the beginning of the data collection period. The IT manager was informed, from the covering letter, that an enumerator would contact him in order to arrange a meeting for the completion of the questionnaire.
Latvia	Stand-alone survey. Answer was mandatory. It was a mail survey where enterprises also were able to submit the filled in questionnaires electronically by e-mail and to fill questionnaires using our on-line submission system via Internet (in CSB web-site).
Lithuania	Stand-alone survey. Answer was mandatory. Postal mail surveys
Luxembourg	Stand-alone survey. Answer was voluntary. Self administered.
Hungary	Stand-alone survey. Answer was mandatory. Self-administered mail survey
Malta	Stand-alone survey. Answer was mandatory. Face-to-face interviews.
Netherlands	Stand-alone survey. Answer was voluntary. The questionnaires are sent out by mail.
Austria	Stand-alone survey. Answer was voluntary. A paper questionnaire with an attached glossary and a stamped addressed envelope for returning the questionnaire to Statistics Austria was sent to the respondents – together with an accompanying letter explaining the political background of the survey and providing basic instructions. Additionally, enterprises could report data by an electronic questionnaire via Internet (writable pdf-file). A download of the questionnaire and the glossary was also possible. A telephone hot line for the respondents has been established. Two written reminders were sent out and 1.300 enterprises were contacted by phone ("motivation calls" by our in-house call centre).
Poland	Stand-alone survey. Answer was mandatory. Self-administered mail survey
Portugal	Stand-alone survey. Answer was mandatory. This survey combines two techniques of data collection: postal and electronic. The enterprises can send the information by post or download the questionnaire from the NSI site and send it through the system created for the effect.
Slovenia	Stand-alone survey. Answer was mandatory. Self-administered mail survey – the questionnaires were sent to the manager of the company by mail.
Slovak Republic	Stand-alone survey. Answer was mandatory. The questionnaires as well as the reminders were sent to the enterprises by post. Some respondents were also reminded by phone.
Finland	Stand-alone survey. Answer was mandatory. Mailed survey with possibility to respond electronically in web (Financial sector not in web)

Sweden	Stand-alone survey. Answer was voluntary. Postal mail survey.
United Kingdom	Stand-alone survey. Answer was mandatory. Self-administered postal mail survey.
Norway	Stand-alone survey. Answer was voluntary. Self-administered mail survey in combination with web questionnaire. All sampled enterprises were able to report either via paper questionnaire or web questionnaire.

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.

The overview table above indicates that in 2005 all EU national ICT usage surveys were stand-alone surveys.

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 nearly 100%. But a voluntary survey can settle this argument by increasing the sample size and sending reminder letters to the respondents.

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 theses 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 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 preceeding (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 Community 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 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.

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 in section II.2.3.3.

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).

Recommendation in the presence of misclassification of enterprises

There are different possible options available in the presence 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 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 strata 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[8]

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 (§2.7.2.1): 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 (§2.7.2.2) 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 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.

---

[8] 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.

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.

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 (poststratification 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.

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 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 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 previous year is not recommended as the use of ICT can change over time. It is e.g. not obvious that an enterprise that sold goods or services via the Internet during year t-1 still did 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 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:

1) 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.

2) 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.

I.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. 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. 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 (see also Chapter 9 of the *Reporting template*, Annex I.3.3). The estimation of the sampling variance should ideally take into account the sampling design (e.g. the stratification).

Some guidelines for interpretation and use of the coefficients of variation can be found in section I.2.3 where the sampling design is discussed.

The computed accuracy measures will also determine the reliability of the estimates. If the coefficient of variation for a certain proportion or absolute value is too high, the cell needs to be suppressed.

To avoid that the level of the estimate 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) could be used as a reference measure.

In case the computation of the standard deviation or coefficient of variation 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).

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:

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; economic activities covered; size coverage in terms of number of persons employed; and territorial coverage.

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, 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 cases unable to respond; other non-response; net sample size; unit response rate. Additionally, the report 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 response rates below 90%; methods used for dealing with item non-response.

h. Grossing-up procedures.

i. Sampling error: the report indicates the coefficient of variation (relative standard error) for a selected group of indicators or sub-indicators listed in the template.

j. Problems encountered and lessons to be learnt.

k. The report includes the questionnaire in national language and, if available, in English.

The *Reporting template* used for the European enterprise ICT usage surveys, including some instructions, can be found in Annex I.3.3.
(see document MM2007_AnnexI.3.3_ReportingTemplateENT2007.doc (20 pages))

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 Part III.1) and in Annex II, chapter 6 of Regulation 1031/2006 (see Part **Error! Reference source not found.**).

I.3. ANNEXES

I.3.1. Model questionnaire

<table>
<tr><td colspan="2" align="center">COMMUNITY SURVEY ON ICT USAGE AND E-COMMERCE IN ENTERPRISES

2007

Model Questionnaire (Version 3.2 of 19 April 2007)</td></tr>
</table>

(Questions relating to the *i*2010 Benchmarking Indicators are marked with an asterisk *)

<table>
<tr><td colspan="4">Module A: General information about ICT systems</td></tr>
<tr><td>A1.</td><td>Did your enterprise use computers, during January 2007?
(Filter question)</td><td>Yes ☐</td><td>No ☐ → Go to X</td></tr>
<tr><td>A2.</td><td>How many persons employed used computers at least once a week, during January 2007? - Optional

If you can't provide this value,
Please indicate an estimate of the percentage of the number of persons employed used computers at least once a week, during January 2007. - Optional</td><td colspan="2"><i>(Number)</i>

_____ %</td></tr>
<tr><td>A3.*</td><td colspan="3">Did your enterprise have the following information and communication technologies, during January 2007?</td></tr>
<tr><td></td><td></td><td align="center">Yes</td><td align="center">No</td></tr>
<tr><td></td><td>a) Wireless LAN</td><td align="center">☐</td><td align="center">☐</td></tr>
<tr><td></td><td>b) Wire based LAN</td><td align="center">☐</td><td align="center">☐</td></tr>
<tr><td></td><td>c) Intranet</td><td align="center">☐</td><td align="center">☐</td></tr>
<tr><td></td><td>d) Extranet</td><td align="center">☐</td><td align="center">☐</td></tr>
<tr><td>A4.*</td><td>Did your enterprise have in use, in January 2007, any software application to manage the placing and/or receipt of orders?
(Filter question)</td><td>Yes ☐</td><td>No ☐ → Go to A6</td></tr>
<tr><td>A5.*</td><td colspan="3">Did that software application to manage orders link automatically with any of the following, as of January 2007?</td></tr>
<tr><td></td><td></td><td align="center">Yes</td><td align="center">No</td></tr>
<tr><td></td><td>a) Internal system for re-ordering replacement supplies</td><td align="center">☐</td><td align="center">☐</td></tr>
<tr><td></td><td>b) Invoicing and payment systems</td><td align="center">☐</td><td align="center">☐</td></tr>
<tr><td></td><td>c) Your system for managing production, logistics or service operations</td><td align="center">☐</td><td align="center">☐</td></tr>
<tr><td></td><td>d) Your suppliers' business systems (for suppliers outside your enterprise group)</td><td align="center">☐</td><td align="center">☐</td></tr>
<tr><td></td><td>e)Your customers' business systems (for customers outside your enterprise group)</td><td align="center">☐</td><td align="center">☐</td></tr>
<tr><td>A6.*</td><td>Did your enterprise have in use, in January 2007, an ERP software package to share information on sales and purchases with other internal functional areas (for example, finance, planning, marketing, etc.)?</td><td>Yes ☐ No ☐</td><td>Don't know ☐</td></tr>
<tr><td>A7.*</td><td colspan="3">Did your enterprise have in use, in January 2007, any software application for managing information about clients (so called CRM) that allows it to:</td></tr>
<tr><td></td><td></td><td align="center">Yes</td><td align="center">No</td></tr>
<tr><td></td><td>a) Capture, store and make available to other business functions the information about its clients?</td><td align="center">☐</td><td align="center">☐</td></tr>
<tr><td></td><td>b) Make analysis of the information about clients for marketing purposes (setting prices, make sales promotion, choose distribution channels, etc.)?</td><td align="center">☐</td><td align="center">☐</td></tr>
<tr><td>A8.*</td><td>Did your enterprise have in use, in January 2007, third party free or open source operating systems, such as Linux ? (i.e. with its source code available, no copyright cost, and the possibility to modify and/or (re)distribute it)</td><td>Yes ☐</td><td>No ☐</td></tr>
<tr><td>A9.*</td><td>Was your enterprise, in January 2007, <u>sending</u> e-invoices in a digital format which allows its automatic processing?</td><td>Yes ☐</td><td>No ☐</td></tr>
<tr><td>A10.*</td><td>Was your enterprise, in January 2007, <u>receiving</u> e-invoices in a digital format which allows its automatic processing?</td><td>Yes ☐</td><td>No ☐</td></tr>
<tr><td>A11.*</td><td>Was your enterprise, in January 2007, using a digital signature in any message sent, i.e. using encryption methods that assure the authenticity and integrity of the message (uniquely linked to and capable of identifying the signatory and where any subsequent change to the message is detectable)?</td><td>Yes ☐</td><td>No ☐</td></tr>
</table>

	Module B: Use of Internet		
	(asking enterprises with ICT)		
B1.	**Did your enterprise have access to Internet, during January 2007?** (Filter question)	Yes ☐	No ☐ → Go to D1
B2.*	**How many persons employed used computers connected to the World Wide Web at least once a week, during January 2007?**	(Number)	
	If you can't provide this value, **Please indicate an estimate of the percentage of the number of persons employed used computers connected to the World Wide Web at least once a week, during January 2007.**	└─┴─┴─┘ %	

B3.*	**Did your enterprise have the following types of external connection to the Internet, during January 2007?**	Yes	No
	a) Traditional Modem (dial-up access over normal telephone line)	☐	☐
	b) ISDN connection	☐	☐
	c) DSL (xDSL, ADSL, SDSL etc) connection	☐	☐
	d) Other fixed internet connection (e.g. cable, leased line (e.g. E1 or E3 at level 1 and ATM at level 2), Frame Relay, Metro-Ethernet, PLC - Powerline comunication, etc.)	☐	☐
	e) Mobile connection (e.g. e.g. analogue mobile phone, GSM, GPRS, UMTS, EDGE, CDMA2000 1xEVDO)	☐	☐

B4.	**Did your enterprise use the Internet for the following purposes, during January 2007?**		
	(as consumer of Internet services)	Yes	No
	a) Banking and financial services	☐	☐
	b) Training and education	☐	☐
	c) Market monitoring (e.g. prices)	☐	☐

B5.*	**Did your enterprise use the Internet for interaction with public authorities, during 2006?** (Filter question)	Yes ☐	No ☐ → Go to B8

B6.*	**Did your enterprise interact with public authorities in the following ways, during 2006?**	Yes	No
	a) For obtaining information	☐	☐
	b) For obtaining forms, e.g. tax forms	☐	☐
	c) For returning filled in forms, e.g. provision of statistical information to public authorities	☐	☐
	d) Submitted a proposal in an electronic tender system (e-procurement)	☐	☐

B7.	**Did your enterprise have a Web Site / Home Page, during January 2007?** (Filter question)	Yes ☐	No ☐ → Go to B10

B8.	**Did the Web Site of your enterprise provide the following facilities, during January 2007?**		
	(your enterprise as provider of Internet services)	Yes	No
	a) Marketing the enterprise's products	☐	☐
	b) Facilitating access to product catalogues and price lists	☐	☐
	c) Providing after sales support	☐	☐

	Module C: e-commerce via Internet		
	(asking enterprises with Internet access)		
	Orders placed via Internet (Purchases)		
C1.*	**Did your enterprise order products/services via the Internet, during 2006 (excluding manually typed e-mails)?** (Filter question)	Yes ☐	No ☐ → Go to C3
C2.*	**Please indicate for 2006 the percentage of the Internet orders in relation to the total purchases (in monetary terms, excluding VAT).**	Less than 1% ☐ 1% or more and less than 5% ☐ 5% or more and less than 10% ☐ 10% or more and less than 25% ☐	
	Alternative Question:	25% or more ☐	
	Please state the value of the purchases resulted from orders placed via Internet (in monetary terms, excluding VAT), in 2006.	*(National Currency)*	
	If you can't provide this value, **Please indicate an estimate of the percentage of the total purchases resulted from orders placed via Internet, in 2006.**	|_|_|_| %	
	Orders received via Internet (Sales)		
C3.*	**Did your enterprise receive orders via the internet, during 2006 (excluding manually typed e-mails)?** (Filter question)	Yes ☐	No ☐ → Go to D1
C4.*	**Please state the value of the turnover resulted from orders received via Internet (in monetary terms, excluding VAT), in 2006.**	*(National Currency)*	
	If you can't provide this value, **Please indicate an estimate of the percentage of the total turnover resulted from orders received via Internet, in 2006.**	|_|_|_| %	
C5.*	**Was your enterprise using a secure protocol, such as SSL and TLS, for the reception of orders via Internet, in January 2007?**	Yes ☐	No ☐

	Module D: E-commerce via external computer networks other than Internet		
	(asking enterprises with ICT)		
	Orders placed via external computer networks other than Internet (Purchases)		
D1.*	**Did your enterprise order products/services via external computer networks other than Internet, during 2006?** (Filter question)	Yes ☐	No ☐ → Go to D3
D2.*	**Please indicate for 2006 the percentage of the orders placed or transmitted via computer networks other than Internet, in relation to the total purchases (in monetary terms, excluding VAT).**	Less than 1% ☐ 1% or more and less than 25% ☐ 25% or more and less than 50% ☐ 50% or more and less than 75% ☐	
	Alternative Question:	75% or more ☐	
	Please state the value of the purchases resulted from orders placed via external computer networks other than Internet (in monetary terms, excluding VAT), in 2006.	*(National Currency)*	
	If you can't provide this value, **Please indicate an estimate of the percentage of the total purchases resulted from orders placed via external computer networks other than Internet, in 2006.**	|_|_|_| %	
	Orders received via external computer networks other than Internet (Sales)		
D3.*	**Did your enterprise receive orders via external computer networks other than Internet, during 2006?** (Filter question)	Yes ☐	No ☐ → Go to E1
D4.*	**Please state the value of the turnover resulted from orders received via computer networks other than Internet (in monetary terms, excluding VAT), in 2006.**	*(National Currency)*	
	If you can't provide this value, **Please indicate an estimate of the percentage of the total turnover resulted from orders received via computer networks other than Internet, in 2006.**	|_|_|_| %	

	Module E*: e-Skills – ICT competence in the enterprise unit and the demand for ICT skills		
	(asking enterprises with ICT)		
E1.	**Did your enterprise employ ICT/IT specialists, in January 2007?** (Filter question)	Yes ☐	No ☐ → Go to E3
	Definition ICT/IT specialists : ICT specialists or IT specialists have the capability to specify, design, develop, install, operate, support, maintain, manage, evaluate and research ICT and ICT systems. ICT is the main job.		
E2.	**How many ICT/IT specialists were employed by your enterprise, during January 2007?** If you can't provide this value, **Please indicate an estimate of the percentage of the number of ICT/IT specialists in relation to the total number of persons employed, during January 2007.**	*(Number)* _____ |__|__|__| %	

E3.	**Did your enterprise recruit or try to recruit personnel for jobs requiring ICT specialist skills, during 2006?** (Filter question)	Yes ☐	No ☐ → Go to E6
E4.	**Did your enterprise have hard-to-fill vacancies for jobs requiring ICT specialist skills, during 2006?** (Filter question)	Yes ☐	No ☐ → Go to E6

E5.	**What do you believe were the main reasons of having hard-to-fill vacancies for jobs requiring ICT specialist skills, during 2006?**	Yes	No
	a) Lack or too low number of applicants with ICT specialist skills	☐	☐
	b) Lack of ICT related qualifications from education and/or training	☐	☐
	c) Lack of work experience in the field of ICT	☐	☐
	d) Salary requests too high	☐	☐
	e) Other **- Optional**	☐	☐

E6.	**Did your enterprise recruit or try to recruit personnel for jobs requiring skills in the use of ICT, during 2006?** (Filter question)	Yes ☐	No ☐ → Go to E8
	Definition ICT user skills: Capabilities enabling the effective use of common, generic software tools (basic user skills) or advanced, often sector-specific, software tools (advanced user skills). Jobs requiring ICT user skills: ICT is an important tool for the job and is used to produce work output and/or used intensively at work (in day-to-day activities)		
E7.	**Did your enterprise have hard-to-fill vacancies due to applicants' lack of skills in the use of ICT, during 2006?**	Yes ☐	No ☐

E8.	**Did your enterprise provide training to develop or upgrade ICT related skills of your personnel, during 2006?**	Yes	No
	a) Training for ICT/IT specialists	☐	☐
	b) Training for users of ICT	☐	☐

E9.	**Were any ICT functions requiring ICT/IT specialists performed by external suppliers (fully or partly), during 2006?**	Yes ☐	No ☐
	Definition External Suppliers: Suppliers in a foreign country or foreign suppliers includes 1) foreign affiliates, usually legal entities, established by the enterprise (internal suppliers from abroad), and 2) other foreign enterprises (external suppliers from abroad).		

E10.	**Were any ICT functions requiring ICT/IT specialists performed by suppliers in a foreign country (fully or partly), during 2006?** (Filter question) *Definition Suppliers in a foreign country: Suppliers in a foreign country or foreign suppliers can be 1) foreign affiliates, usually legal entities, established by the enterprise (internal suppliers from abroad) and/or 2) other foreign enterprises (external suppliers from abroad).*	Yes by foreign affiliates established by the enterprise	Yes by other foreign enterprises	No
		☐	☐	☐ → Go to E13/X

E11.	**Which ICT functions were performed by suppliers' ICT/IT specialists <u>in a foreign country</u>, during 2006? - Optional**		
		Yes	No
	a) ICT management (includes e-business and ICT systems management)	☐	☐
	b) ICT development and implementation (includes business software development, programming, web development, database development, communication network development, systems integration and installation)	☐	☐
	c) ICT operations (includes technical support, user help and support, network administration, web administration, database administration)	☐	☐
	d) Other ICT functions	☐	☐
E12.	**From which of the following geographical regions did your enterprise engage suppliers' ICT/IT specialists<u>, during 2006</u>?**		
		Yes	No
	a) other EU Member States	☐	☐
	b) Non-EU Countries	☐	☐

E13.	**Were any business functions requiring users of ICT performed by external suppliers (fully or partly), during 2006? - Optional**	Yes ☐		No ☐

E14.	**Were any business functions requiring users of ICT performed by suppliers <u>in a foreign country</u> (fully or partly), during 2006? - Optional** (Filter question)	Yes by foreign affiliates established by the enterprise	Yes by other foreign enterprises	No
		☐	☐	☐ → Go to X

E15.	**Which business functions (non-ICT) were performed by suppliers' ICT users <u>in a foreign country</u>, during 2006? - Optional**		
		Yes	No
	a) Sales and marketing, customer services	☐	☐
	b) Research and development, product design and engineering	☐	☐
	c) Other (non-ICT) business functions	☐	☐
E16.	**Please indicate the geographical regions from where you engaged business services requiring ICT users<u>, during 2006</u>. - Optional**		
		Yes	No
	a) other EU Member States	☐	☐
	b) Non-EU Countries	☐	☐

Module X: Background information	
(X1-X5) available in some countries from SBS and thus not to be included; latest available information should be provided	
X1.	**Main economic activity of the enterprise, during 2006**	
X2.	**Average number of persons employed, during 2006**	
X3.	**Total purchases of goods and services (in value terms, excluding VAT), for 2006**	
X4.	**Total turnover (in value terms, excluding VAT), for 2006**	
X5.	**Location (Convergence/ non-Convergence region), in 2006**	

I.3.2. Transmission format

General Enterprise Survey

Version 1.1, May 2007.

Note: this transmission format refers to the 2007 survey.

⇨ See document ***MM2007_AnnexI.3.2_TransmissionFormatENT2007.doc***

I.3.3. Reporting template

Note: the reporting template for interim & final methodological report refers to the 2007 surveys.

⇨ See document ***MM2007_AnnexI.3.3_ReportingTemplateENT2007.doc***

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 practise 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 2005 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 1099/2005 (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 then 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, see p. **Error! Bookmark not defined.**).

The operational version 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*.

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 questions) 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 aggregated data to Eurostat, can be found in *II.3.2 Transmission format*.

II.1.6. Explanatory notes

The explanatory notes in this chapter refer to the questions in the 2007 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 2007model questionnaire (but were part of the 2006 and 2005 model questionnaires) have been retained if necessary. More information on former questions can be found in the 2006 manual.

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 a computer at home?

[<u>Scope</u>: all households]

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

For a definition of a household see below, questions F8 and F9 of *Module F: Socio-demographic background variables*.

Access at home refers to the dwelling unit where the household/individual has its <u>address</u>.

Compared to the 2006 survey asking about any devices at home, the question addresses access to a computer of any type, i.e. desktop, laptop or handheld (for further details, see question A3).

"You or *anyone in your household*" may make it clear to the respondent that the questions also concern other members (thus avoiding that the respondent excludes e.g. the case where his/her partner has a computer via work). The wording *have access to* is prefered to *have* because the latter may erroneously be interpreted in the sense of 'ownership' (while the only computer in the household may actually be the property of the employer of one of the members in the household).

A2: Does any member of the household have access to the Internet <u>at home</u>, regardless of whether it is used?

[<u>Scope</u>: all households]

[<u>Type</u>: 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.

The clause 'regardless of whether it is used' 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).

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.

A3: On which of these devices is the Internet accessed at home?

[Scope: all households with Internet access at home, i.e. A2 = Yes]

[Type: multiple answers allowed, i.e. *Tick all that apply*]

The following devices can provide access to the Internet.

a) Desktop computer

A desktop computer can be defined as a computer that usually remains fixed in one place. Normally the user is placed in front of it, behind the keyboard with the monitor placed on top.

b) Portable computer (laptop)

A portable computer (or laptop) can be defined as a computer that is small enough to carry. A portable computer (or laptop) can usually carry out the same tasks as a desktop computer. Includes notebooks but not handheld computers, although portable.

c) TV set with specific Internet device (e.g. digital TV or set top box)

The main issue is if the TV set with the specific device enables Internet access. That means the specific Internet device should allow users to access the WWW or e-mail service.

Digital TV refers to television via digital signals. This is the basic difference compared to ordinary TV transmission which is via analogue signals. Improvement of picture quality and Internet access are some of the features that characterize digital TV. However, there could be new, popular digital TV sets in some countries which more or less only allow watching TV and video on demand.

Today's analogue TV needs e.g. a set top box, which is a device that sits on top of the TV set and is the interface between the home TV and the cable TV company. The set top box changes the digital signal from the cable TV company so the home TV (which is analogue) can read the signal.

d) Games console

Some games consoles have built-in modems and thus give access to the Internet, e.g. Playstation 2 (Playstation 3 is planned to be available in Europe from March 2007). A games console is an electronic device designed to play games. Examples are Sony playstation, Nintendo GameCube and Microsoft Xbox.

e) Other means (i.e. none of the above)

Applies where the household does not have access to the above mentioned devices at home.

Item *e) Other means* in the 2007 model questionnaire lists two optional sub-items, namely:

e1) via Internet enabled mobile phone

e2) via handheld computer (palmtop, PDA)

e1) Internet enabled mobile phone

The household is asked if it is an Internet enabled mobile phone via WAP, GPRS or UMTS (defintions: see below). This should be regardless whether it is used or not. Internet can be defined as a networking infrastructure, where any computer or other device can communicate with each other as long as they are both connected to the Internet. The Internet is used for access to the World Wide Web and is also used for e-mail correspondence, instant messaging etc. During an interview, questions such as 'do you use picture messaging' or 'does your mobile phone have a built-in camera' can help to detect whether a respondent has an Internet enabled mobile phone.

The term GPRS (General Packet Radio Systems) stand for 2nd generation mobile technologies (though GPRS is sometimes referred to as 2,5 generation) while UMTS (Universal Mobile Telephone System) is the term for 3rd generation mobile technology. Both give access to the Internet but 3rd generation at a much higher speed.

Compared to 2006, the term 'WAP' (Wireless Application Protocol) has been deleted in the optional category *e1*, because WAP or i-mode are protocols enabling viewing web pages written by WAP language programs using this standard on a mobile phone screen (see also question A4 and C4).

e2) Handheld computer (palmtop)

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, such as Sony Ericsson P800 and P900 series), there is an overlap with item b) of this question, i.e. these PDAs can be counted for as both a mobile phone (item *b*) and a handheld computer (item *f*).

These items e1) and e2) have been rendered optional as they come back in the individual part of the questionnaire (see Module C, question C4).

A4: What types of Internet connection are used?

[Scope: all households with Internet access at home, i.e. A2 = Yes]

[Type: multiple answers allowed, i.e. *Tick all that apply*]

Below is given an explanation of the most commonly used technologies. 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) Modem (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.

b) DSL (e.g. ADSL, SHDSL etc.)

DSL (Digital Subscriber Line) stands for technology that transports data at high speeds (i.e. equal to or higher than 144 kbit/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.

c) Other broadband connection (e.g. cable, UMTS, etc.)

This question sums up all other types of Internet connection which is not provided via DSL but still can be designated as broadband.

The technologies in question are:

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 bandwith potential.

Fixed wireless technologies (FWA): Technologies in this group are all wireless technologies that provide access between fixed points.

Satellite technologies: Provides wireless access and are predominantly used in remote areas not reached by cable or DSL networks.

UMTS: Relates to mobile phone over broadband. For a definition, see question *A3 e1)*.

d) Mobile phone over narrowband (WAP, GPRS, etc.)

Broadband over the mobile phone is still limited (see UMTS/3G above). Second generation (2G) mobile technologies such as GPRS (General Packet Radio Service, see also question A3 for a definition) 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. WAP is actually a protocol and not a mobile phone system, see question A3 e1 above.

A5: What are the reasons for not having a broadband connection (such as cable or DSL) at home?

[Scope: all households without Broadband access at home, i.e. A4b and A4c = No]

[Type: multiple answers allowed, i.e. *Tick all that apply*]

The question replaces the 2006 question on barriers to Internet access. Both are i2010 core indicators and rotate with each other each year.

The items give an idea of the household reasons or barriers for not having access to broadband at home. Respondents should indicate all reasons that are relevant to them, not only the main reason(s).

a) Too expensive

Too expensive could mean that subscription fees for broadband use, related equipment costs such as DSL Modem/WLAN router etc. are too high.

b) No need

May apply to people that do not find broadband access useful or interesting.

c) Not available in my area

May apply to people living in households in areas where it is not possible to subscribe to broadband because of missing telecommunication and mobile network infrastructures.

d) Can access broadband somewhere else (e.g. at work)

May apply to those who have access to broadband elsewhere (e.g. at work, at other people's homes etc.) and find it not useful or interesting enough to have access at home.

e) None of the above, other

This category is included to close the question. Assuming that respondents that do not have a broadband Internet connection at home must have a reason for not having it, the respondent is supposed to tick at least one of the options *a* to d or option e. In cases where the respondents states not to have any particular reason (not even after having heard the options in the questionnaire) *None of the above, but other* can be ticked.

II.1.6.2 MODULE B: USE OF MOBILE PHONE AND COMPUTERS, LOCATION AND FREQUENCY OF USE

➔ Statistical unit: individuals

B1: When did you last use a computer?

[<u>Scope</u>: all individuals]

[<u>Type</u>: one single answer needed, i.e. *Tick only one* ; filter question]

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). PDA's (handheld computers or palmtops) should also be included. Other equipment with embedded computing abilities, e.g. cell phones, VCR's, TV sets, washing and dishing machines should be excluded.

The questions on computer use are relevant for digital literacy which is the 2007 focus in the i2010 benchmarking framework. Question B1 serves as a filter for questions E1 and E3. Compared to 2006, 'most recently' is rephrased as 'last'.

B2: How often on average have you used a computer in the last 3 months?

[<u>Scope</u>: all individuals having used a computer in the last 3 months, i.e. 1st option in B1]

[<u>Type</u>: one single answer needed, i.e. *Tick only one*]

For definition of a computer see above, B1.

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 *a) Every day or almost every day*.

B3: Where have you used a computer in the last 3 months?

[<u>Scope</u>: all individuals having used a computer in the last 3 months, i.e. 1st option in B1]

[<u>Type</u>: multiple answers allowed, i.e. *Tick all that apply*]

For definition of a computer see above, B1.

Concerning alternative *b) At place of work (other than home)*, this should be interpreted as any place where work is usually conducted excepting at home. The usual place of work can be much wider than just the premises of the employer.

Alternative *c) At place of education*, should be interpreted as the institution where the education takes place ("within the school walls"). Teachers using computers in their school

should tick alternative *b) At place of work* as alternative *c)* rather refers to the point of view of pupils or students.

See also C3 (location of Internet use).

B4: Do you use a mobile phone?

[<u>Scope</u>: all individuals]

[<u>Type</u>: one single answer needed, i.e. *Tick only one* ; binary (*Yes/No*) ; filter question for B4]

This question serves as a stand-alone study variable and refers to any type of mobile phone, not only Internet enabled (includes e.g. GSM).

For definition of a mobile phone see above A3e1, A4c and A4d.

'Use' refers to the actual use of this device. Whether the respondents *owns* the mobile phone is irrelevant to this question. In most cases, it will be a personal mobile phone but in some households a 'collective' mobile phone may be shared by several members – this is especially the case for households that are 'new' to mobile telephony. The mobile phone can be owned by the respondent or can e.g. be provided by the respondent's employer.

II.1.6.3   MODULE C: USE OF THE INTERNET

➔ Statistical unit: individuals

C1: When did you most recently use the Internet?

[<u>Scope</u>: **all individuals** (even those who reported never having used a computer in B1 as they may have accessed the Internet via other means, though this may be very exceptional)]

[<u>Type</u>: one single answer needed, i.e. *Tick only one* ; filter question]

This questions 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*.

C2: On average how often did you use the Internet in the last 3 months?

[<u>Scope</u>: individuals who used the Internet in the last 3 months, i.e. 1st option *(Within the last 3 months)* in C1]

[<u>Type</u>: 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 four 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) At least once a month but not every week

d) Less than once a month

C3: Where have you used the Internet in the last 3 months (using a computer or any other means)?

[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*]

The question is to be answered regardless of the device used or type of connection used. E.g. if a person uses a portable computer with a wireless data card (to be inserted in the card slot) at several locations, he or she should tick all those locations.

a) At home

This should be ticked if the respondent has used the Internet at home for any purpose, private or work-related.

b) At place of work (other than home)

This should be defined as usual place of work. If the respondent usually works at home then *a)* should be coded. If the respondent's work is usually based in a number of places then any of these places count as their place of work. However, use of mobile technology in places such as hotels etc. even for work related purposes, should be coded as *e) At other places*.

c) At place of education

This should refer to the respondent's own place of education. If the respondent works at an educational establishment – e.g. a teacher - then *b)* should be coded.

d) At another person's home

This should not be used if the respondent's place of work is another person's home.

e) At other places

This includes ANY places not already coded above. The further responses are a subset of places which might be included in *At other places* but is NOT a comprehensive list. On the other hand, not all items of the list may be relevant to each country.
Ideally the respondent should be presented with the further list only after they have said "other place"

of which:

e1) Public library: some public libraries offer the possibility to access the Internet, whether it be a free service or if a charge is made.

e2) Post office: see definition below.

e3) Public office, town hall, government agency: The wording to be used in each member state should reflect titles of public offices which respondents would recognise.

e4) Community or voluntary organisation: see definition below.

e5) Internet café: Commercial places that provide Internet access services under payment. There is usually a charge for using the Internet. This charge is normally calculated on the basis of the amount time spent on-line. An Internet Café may also charge for usage not related to the internet (e.g. use of word processing software).

Post offices, *Public offices (town hall, government agencies)*, *Community or voluntary organisations* are all public places capable of providing access to the Internet. The access to the Internet in these public places can be free or under payment of a fee.

e6) Hotspot: added as new optional item for 2007. The Internet can be used at a so-called hotspot with portable or handheld computers via wireless connection at hotels, airports, public places, etc.

C4: Do you use any of the following mobile devices to access the Internet?

[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 question refers to the actual *use* as a means of access not to the 'theoretical' possibility of *access*. Question C4 is not limited to Internet access at home, but can refer to all possible locations (as is in general the case for Module C).

Via *Mobile phone* has been split into two items in order to separate UMTS (see item **b**) from GPRS (see item **a**). First, because of a high interest from users of ICT statistics in the take-up and usage of UMTS. Second, because this also enables to distinguish between mobile phone access over broadband versus mobile phone access over narrowband.

For additional information on GPRS, UMTS or *Handheld computer*, see Module A, question A3 (items e1 and e2).

Item **d)** portable computers (laptop) via wireless connection away from home or work as an important device to access the Internet remotely has been added for 2007. To distinguish real mobile (in the sense of 'on the move') use from situations where the portable computer serves as a desktop at home or at work, the (longer) wording is supposed to exclude such cases.

C5: For which of the following activities did you use the Internet in the last 3 months for private purposes?

[<u>Scope</u>: individuals who used the Internet in the last 3 months (i.e. 1st option in C1)]

[<u>Type</u>: 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.

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 / videoconferencing

Telephoning over the Internet is a relatively inexpensive method to communicate and is often the method used by companies offering reduced cost telephone charges. Users may not be aware that they are communicating using such Internet-based telephony, Voice-over-IP or VoIP.
Next to VoIP, peer-to-peer telephony is becoming more 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.
Videoconferencing includes audio and visual communication between two or more groups or persons. Videoconferencing is more widely used by organisations and replaces the need for face to face meetings. It should be remembered that this question is concerned with private purpose only.

c) Other (use of chat sites, messenger, etc.)

Chat sites can be used by two or more persons for the purpose of communication. This communication is by written word (similar to email). The identity of the person 'chatting' does not have to be divulged to other users (e.g. a username can be created). Users must be 'logged-on' to a chat site to use it. Users that are logged on to a specific chat site can see all correspondence between all users on that site. Correspondence is updated in real time.
This item also covers instant messaging which means real-time communication between people on the basis of typed text.

Information search and on-line services

d) 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).

e) 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 accommodation, for example travel tickets, hotels or any other type of accommodation or web sites containing information for tourists.

f) Listening to web radios / watching web television

This covers both live streaming (real-time) and radio or TV 'on demand' (batch, i.e. the user can listen/watch programs later on). This does not result in any saving of information to, for example, the respondents computer.

g) Playing or downloading games, images, films or music

This refers to actually downloading games, images for use, films for watching, or music for listening.

h) Downloading software

Includes downloading software either free of charge or under payment.

i) Reading or downloading online newspapers / news magazines

This should include all types of online newspapers and magazines either free of charge or under payment.

A sub-item has been introduced for 2007 to distinguish subscription to news services from news consulted on-screen:

i1) This refers to subscribing to news services or news products which are then received regularly.

j) Looking for a job or sending a job application

Includes searching specific web sites for job 'hunting' or for sending an application for a job. Sending a job application should be included in this category only if it was sent on-line.

k) Seeking health-related information (e.g. injury, disease, nutrition, improving health, etc.)

This item refers to Internet use for health related activities. The scope is limited to private purpose, professional use is not to be taken into account. Private should however not be limited to own personal use, but can also include internet use for health related activities on behalf of other family members or friends.
Includes general searches via a search engine (Google, Yahoo!, …) using keywords in one of the mentioned fields. This item also includes more specific searches on specialised websites such as the Ministry of Health, non-governmental bodies or interest groups. The respondent may have obtained the website's address from a folder or an article.
Although not yet available in most countries, this item also includes activities such as making an appointment on-line with a practitioner, requesting a prescription on-line from a practitioner or seeking medical advice on-line from a practitioner.

l) Other information search or on-line service

This item is supposed to close the block on *Information search and on-line services* by ensuring that people who didn't look for anything in particular (i.e. purposes covered by

the above list of items), can nevertheless tick this item *l) Other*. This can be relevant for Internet users who only do general browsing, without looking for a particular good, service or application (e.g. trying someone's name in a search engine like Google).

The item also covers goods/services/applications that are relatively new (and therefore not yet included under *a* to *k*), sensitive or of a lower relevance (and therefore not mentioned as a separate entry).

Banking, (optional: selling of goods or services)

m) 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 *i)* of question D2, see Module D. A simple information search on e.g. shares or financial services is of course included under item *d) Finding information on goods and services*.

n) Selling goods or services (e.g. via auctions)

Selling goods or services on-line 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). In this case, the individual could tick item *l) Other information search or on-line service* as he/she is making use of a website (e.g. a forum) to offer the product or service.

Interaction with public services or administrations

These three items make use of the conceptual framework for the levels of interaction between enterprises or citizens and public authorities, used by DG INFSO. 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 *p* without having answered *Yes* to *o* (or *Yes* to *q* without having answered *Yes* to *o* and/or *p*).

o) Obtaining information from public authorities' websites

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. For example respondents may not realise some public information sites are run by public authorities, e.g. a meteorological office offering weather forecast services.

p) 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.

q) Sending in filled forms

Includes filled in forms sent via internet (public authorities' websites) only. Forms downloaded, printed, filled in and sent by post should not be included in this category.

Training and education

In 2007 the activities related to training and education are more specified regarding on-line activities.

r) Looking for information about education, training or course offers

Includes searching for information relating to course offers at school, college or university whilst part of the education system, but also post educational course offers including leisure courses for gaining and improving knowledge in any subject. Post graduate courses at university should be included.

s) Doing an online course (of any subject)

Doing an online course reflects learning courses distant from the location of education and training organisations or employer where courses can be attended in person (often but not necessarily done at home). Interaction with teachers, trainers and/or learning material is done via the Internet. The use of e-learning software programmes can play a role.

t) Consulting the Internet with the purpose of learning

This new item for 2007 relates to using the Internet for self-directed learning.

Self-directed learning is wider than simply looking for information. It refers to looking for information with the purpose of learning something, of improving knowledge. Looking up a city map or a phone number would not be considered as learning, contrary to e.g. looking up definition of broadband technology or information on an economic concept. To help distinguish from more random finding of any information, the 'stronger' verb *to consult* is used ('consulting the Internet' in the sense of 'seeking advise on the Internet'). The words *with the purpose* is added to stress the *intention* of learning (compared to just looking up some 'daily life' information).

C6: In the past 3 months, have you taken part in a course (any training, not only computer or Internet training, including school or university)?

[<u>Scope</u>: individuals who used the Internet in the last 3 months (i.e. 1st option in C1)]

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

The question is wide in scope regarding course subjects and refers to any type of training, whether formal (e.g. university) or informal (computer class or photography class), therefore the word *course* is used as this is less restrictive than *training* or *education*.

The course aim is learning and gaining knowledge on any subject, not improving the physical condition; physical training courses, e.g. aerobics, football, are excluded.

C7: In the past 3 months, have you used the Internet to...

[<u>Scope</u>: individuals who used the Internet in the last 3 months (i.e. 1st option in C1)]

[<u>Type</u>: multiple answers allowed, i.e. *Tick all that apply*]

Following the filter question C6 meaning that this question is only relevant to persons enrolled in a course, the specific items in C7 look at types of learning via courses and training and thus addresses various activities which are done via the Internet. The items are as follows:

a) Do research as part of a training course or your education

This involves finding and reviewing study content and training material online via websites.

b) Exchange messages relating to the course content with other learners

This addresses all types of messages which can be sent via the Internet, including e-mails, instant messaging, etc.

c) Download learning content which was provided online

Provision of online content which can be downloaded is done via a website.

d) Look for the availability of a book or article for your course in a library

This involves consulting the website of a library to find a book or article and look at its availability with or without reserving it for use online.

C8: Would you like to use the Internet more?

[<u>Scope</u>: individuals who used the Internet in the last 3 months (i.e. 1st option in C1)]

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

This question serves as a filter and is asking whether Internet users would like to make more use of the Internet.

C9: What are your barriers to more intensive use of the Internet?

[<u>Scope</u>: individuals who used the Internet in the last 3 months (i.e. 1st option in C1)]

[<u>Type</u>: multiple answers allowed, i.e. *Tick all that apply*]

Only those who answer positive to C8 are asked for the reasons why they are currently not using the Internet more intensively. The reasons which inhibit a more intense use of the Internet are related to eParticipation aspects and other inhibitors.

a) Foreign language skills are inadequate

The respondent would need foreign language skills to satisfy needs for using the Internet more and especially foreign websites, but these skills are lacking or inadequate.

b) Lack of time

Simply, the respondent does not have the time for more use.

c) Connection is too slow

The Internet connection may be of low speed and too time-consuming. The respondent may live in an area where only low speed/narrowband Internet access is available which is not so suitable for making more intensive use of the Internet. Or the respondent's choice is narrowband Internet access because of cost factors but this inhibits intensive use because of the extended download time involved, etc..

d) Additional connection or per-volume download cost

The respondent may wish to use the Internet more, but the additional cost caused (depending on the payment scheme) through the additional time spent or volume downloaded will inhibit more intensive use.

The item may not be applicable to respondents having subscribed to Internet flat rate payment schemes enabling unlimited monthly Internet usage.

e) Cost of online content

The respondent may be interested to use the Internet more to obtain (more) online content such as news, music, video etc.. However, the content of interest is causing costs e.g. on a per article/music file etc. download basis, which prevents the person to make more use of it.

f) Content (what is there is not interesting enough to make more use of the Internet)

A person may wish to use the Internet more, but the available content (e.g. lists of music files offered by services) is found not interesting enough to make more use.

g) Lack of skills or knowledge (e.g. sites are not user-friendly or too complicated)

The respondent does not have sufficient knowledge to use the Internet, for example also for specific purposes such downloading software, or doing Internet banking. The respondent may find web sites not user-friendly or too complicated to make more use of them. Respondents perceiving technology as too complicated usually have a lower level of skills in dealing with it compared to others.

h) Security or privacy concerns

The respondent may wish to use e-shopping but worries about giving credit card details over the Internet (security concerns), or personal details to be provided over the Internet (privacy concerns), or in general about virus concerns and similar security problems (e.g. spyware).

i) None of the above

This is not an additional alternative. It applies when the respondent did not tick any of the above mentioned items. The category is included to close the question. It is in particular useful in postal surveys where respondents fill in the questionnaires themselves.

C10: Through using the Internet, have you had a computer virus resulting in loss of information or time in the last 12 months?

[**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*) ; filter question]**

The question addresses persons who used the Internet in the last 3 months and then asks this (relatively large) group of recent Internet users about their experiences with computer viruses in the last 12 months. The changed routing in C1 for 2007 has been made deliberately and mainly for reasons of simplification.

The question is about having experienced a computer virus as being one of the most popular security problems perceived in connection with Internet use, posing a threat to the ability to use the computer or files or if recovered, resulting in the loss of time in solving the problem.

C11: How often do you make safety copies or back up files (documents, pictures, etc.) from your computer on e.g. a diskette, a cd or to diskspace on Internet servers?

[<u>Scope</u>: individuals who used the Internet in the last 3 months (i.e. 1st option in C1)]

[<u>Type</u>: one single answer needed, i.e. *Tick only one*]

This question aims to reveal the measures taken to avoid potential security problems in terms of securing **private documents, pictures, and other files,** and their frequency (always or almost always, sometimes, never or hardly ever). Safety copies done by the respondent at workplaces are included if private files are concerned.

In addition to those in the question, USB devices would be another example to store copied files.

It is up to the respondent to judge the frequency. Always/almost always means after working on files; sometimes means less often.

II.1.6.4   Module D: Use of e-Commerce

→ Statistical unit: individuals

D1: When did you most recently buy or order goods or services for private use over the Internet (excluding manually typed e-mails)?

[<u>Scope</u>: individuals who already used Internet, i.e. 1st, 2nd or 3rd option in C1]

[<u>Type</u>: one single answer needed, i.e. *Tick only one* ; filter question]

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 - according to the EUROSTAT definition of e-commerce in enterprises - not be included.

Delivery via electronic means or payment via electronic means is not a requirement for an order via Internet.

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.

D2: What types of goods and services did you order over the Internet for private use in the last 12 months?

[<u>Scope</u>: individuals who already used Internet (1, 2 or 3 in C1) <u>*and*</u> 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 D1]

[<u>Type</u>: multiple answers allowed, i.e. *Tick all that apply* ; filter question for D3]

a) Foods / Groceries

Groceries include necessities of "daily use" such as sanitary products, tobacco products, cosmetics, pharmaceuticals or flowers. This category also includes drinks. Newspapers should be excluded from this category.

b) Household goods

Includes e.g. furniture, toys, washing machines, micro-wave ovens, dish washers, vehicles, gardening equipment, plants, antiques, art, tools, collectibles, etc.

c) 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.

d) Books / Magazines / Newspapers / E-Learning materials

Subscription to newsletters and alike – if not free of charge – also belong to this category. 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.

e) Clothes, sports goods

Includes shoes, textiles, accessories.

f) Computer software and upgrades (incl. computer and video games)

Includes video games and other computer software on any media.

g) Computer hardware

Includes any accessories to computer hardware, such as modem and printer.

h) Electronic equipment (incl. cameras)

Includes e.g. mobile phones, cameras, radios, TVs, stereos, DVD players, video recorders.

i) Share purchases / Financial Services / Insurance

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.

j) Travel and holiday accommodation

Includes e.g. hotel reservation, travel tickets and vehicle hire. Only bookings of accommodation and travel tickets 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) Lotteries or betting

m) 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).

***D3: Were any of the products that you ordered over the Internet downloaded
or accessed from websites rather than delivered by post etc.?***

> [<u>Scope</u>: individuals who already used Internet (1, 2 or 3 in C1) <u>*and*</u> who bought/ordered
> over the Internet in the last 12 months (1 or 2 in D1) <u>*and*</u> who bought films/music,
> books/magazines/newspaper/e-learning material or computer software over the
> Internet (i.e. option *c*, *d* or *f* in D2)]

> [<u>Type</u>: multiple answers allowed, i.e. ***Tick all that apply***]

Digitally delivered products include those products which can be delivered on physical media
(such as CD or tape) but also over the Internet in digitised form as a computer file (and thus
independently of those physical media).

a) Films, music

> Delivery of films and music on-line via files (e.g. a video, an MP3 file or a CD).

b) (Electronic) books , magazines, newspapers, e-learning materials

c) Computer software (incl. computer and video games and software upgrades)

II.1.6.5 MODULE E: E-SKILLS

➔ Statistical unit: individuals

***E1: When did you last take a training course (of at least 3 hours) on any
aspect of computer use?***

> [<u>Scope</u>: individuals who already used a computer, i.e. 1st, 2nd or 3rd option in B1]

> [<u>Type</u>: one single answer needed, i.e. ***Tick only one***]

The emphasis is on formal training via courses (education) and excludes informal training
such as help from colleagues or friends/relatives and learning-by-doing. All kinds of courses
which can relate to computer use have to be included, e.g. word processing or spreadsheet
courses, programming, web-design and so forth.

E2: What are the reasons for not having taken a course on computer use recently?

[Scope: individuals who have taken a computer training course more than 3 years ago or who have never taken a computer training course, i.e. 4th or 5th/last option in E1]

[Type: multiple answers allowed, i.e. *Tick all that apply*]

Formal ICT training is one of the main facilitator for developing necessary skills. The question's emphasis is thus on formal training via courses and excludes informal training. The reasons in the answer categories are expected to fit to varying situations from perceiving no need for a course to specific reasons (cost, time, content etc).

Respondents who have no need to take a computer course, i.e. they answered positive on

a) No need to take one because my computer skills are sufficient

b) No need to take one because I rarely use computers

will not be directed to the remaining set of possible reasons which are:

c) lack of time

d) course costs

e) No suitable offer available

f) Courses are too difficult

g) None of the above

E3: 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* ; indirect filter question for E5]

The purpose of this question is to assess the skills level relating to general computer use of the respondents.

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 ordinality 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.

While the above items existed also in the 2006 questionnaire 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), two new items have been added for the 2007 survey which were meant to add more competence and confidence elements for the more advanced end of basic computer skills in the context of digital literacy. However, they are not necessarily more advanced than existing items at the high end, such as writing a computer programme. The two new items represent additional analytical interest, stand-alone and in relation to high basic computer skills.

new) g) Connecting computers to a local area network

(additional item in the 2007 survey at the more complex end of competencies)
Certain technical aspects need to be known to connect computers and to make a LAN run smoothly.

new) h) Detect and solve computer problems (e.g. computer runs slowly)

(additional item in the 2007 survey at the more complex end of competencies)
Experience in detecting and solving computer problems can address for example detecting and solving slow computer performance. Another example is getting stuck when starting up the operating system. In general problems relate to detecting and solving computer hardware, computer software or file related troubles.

old), 2005 Using a mouse (or other pointing device) to open programs

(item from the 2005 survey, no longer included for the 2006 survey)
It applies when the respondent opens an Internet browser, a word processor or any kind of program. Although this item may sound obvious, it can be considered as a 'first test' of the presence of any e-skills. Moreover, quite some people use e.g. Internet Explorer after having asked another household member to start the program for them (the item is not restricted to simply using the mouse to move the cursor).

E4: Which of the following Internet related activities have you already carried out?

[<u>Scope</u>: individuals who already used Internet, i.e. 1st, 2nd or 3rd option in C1]

[<u>Type</u>: multiple answers allowed, i.e. *Tick all that apply* ; indirect filter question for E5]

The purpose of this question is to assess the skills level of the respondents relating to Internet use. Although the items 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 ordinality in the items presented.

a) Using a search engine to find information

The question refers to search engines such as Yahoo, Google etc.

b) Sending e-mail with attached files

Attached files can be anything from a document, a video or pictures etc.

c) Posting messages to chatrooms, newsgroups or any online discussion forum

A chatroom can be defined as a virtual room, where participants have a chat session. Newsgroups are on-line discussion groups covering all kinds of interest. The member of a newsgroup view and post messages via a news server on the Internet.

d) Using the Internet to make telephone calls

e) Using peer-to-peer file sharing for exchanging movies, music etc.

Peer-to-peer (P2P) refers to a network where the communication/information is distributed to a wide variety of computers rather than residing on one central server. In this network each computer has equal capabilities which differ from a traditional client/server network. P2P is more and more used for sharing of files, movies and music.

f) Creating a web page

While the above items existed also in the 2006 questionnaire and items a) to f) should be considered for classifying and comparing from one year to another the categories low level internet skills (1-2 items ticked), medium (3-4 items ticked) and high level skills (5-6 items ticked), two new items have been added for the 2007 survey which were meant to add more competence and confidence elements for the more advanced end of internet skills in the context of digital literacy. However, they are not necessarily more advanced than existing items at the high end, such as creating a web page. The two new items represent additional analytical interest, stand-alone and in relation to a high level of internet related skills.

new) g) Finding, downloading and installing software

(additional item in the 2007 survey at the more complex end of competencies)
The range of software installed on the computer links provides opportunities to make more and wider use of the computer and the Internet. A large amount of software is available via websites from where software can be downloaded and installed on the computer. To know "which" software may be the greater challenge for the user; the downloading and installing process can be well guided or automated.

new) h) Keeping viruses, spyware and adware off your computer

> *(additional item in the 2007 survey at the more complex end of competencies)*
> This item addresses experience in how to avoid security related problems caused by viruses, spyware and adware. This usually involves the selection and installation of software (ordered/delivered on- or offline) that can detect a virus or unauthorised or violent access to the computer by external parties. Spyware is a software used to collect personal information of Internet users without their consent or that they are informed. Adware is gaining access to the computer via the Internet; a pop-up is shown advertising goods or services. Adware may sometimes force the user to visit certain websites and may interfere with other software applications.

E5: *Where or how did you obtain the skills to carry out these activities?*

[Scope: individuals who reported at least one activity under E3(a-f) or E4(a-f)]

[Type: multiple answers allowed, i.e. *Tick all that apply*]

Items a) through c) refer to skills obtained via formal training courses, while items d) through f) refer to more informal ways of acquiring e-skills.

a) Formalised educational institution (school, college, university)

b) Training courses in adult education center

Refers to courses taken by own initiative, not the employer.

c) Vocational training courses

Apart from *b)*, this refers to skills obtained on demand from the employer.

d) Self-study using books, cd-roms, online courses etc.

e) Self-study in the sense of learning-by-doing

f) Informal assistance from colleagues, relatives, friends

E6: *Do you judge your computer skills to be sufficient if you were to look for a job or change jobs within a year?*

[Scope: all individuals, except retired persons]

[Type: one single answer needed, i.e. *Tick only one*]

Apart from retired persons, all individuals, not only those in the labour market, are requested to reflect on their computer skills adequacy.

For example, young people or students, or housewifes/-men and others might have intentions or obligations to look for/return to employment anyway within a year.

However, if the respondent is not at all, not in the given timeframe or in the short/medium term concerned, the possibility to tick "not applicable" is provided.

II.1.6.6 MODULE F: SOCIO-DEMOGRAPHIC BACKGROUND VARIABLES

[Scope: all households / all individuals, excepting for F5 Occupation]

[Type: F2, F3, F4, F6, F7: one single answer needed, i.e. *Tick only one* ; F1, F8, F9, F10: quantitative ; F5: 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 economical background of the observed statistical units.

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 (F1), gender (F2), educational attainment (F3), employment situation (F4) and occupation (F5) are to be answered by the individual(s) who is (are) selected within the household.

The questions on geographical location (F6) and type of locality (F7) 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 (F8, F9) and household income (F10) of course refer to the household level.

F1: 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 at the during the reference period, is not expected to have a significant impact on the results and comparability.

F2: Gender

Self-explanatory.

F3: Educational level

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 (ISCED 1997).

A detailed description of the classification is available from the UNESCO Institute for Statistics' website (http://www.uis.unesco.org/ev.php?ID=3813_201&ID2=DO_TOPIC), but a short summary is presented below.

Conversion tables matching the national educational levels to the ISCED codes are available from the CIRCA site (information relating to the school year 1999-2000): http://forum.europa.eu.int/Public/irc/dsis/edtcs/library?l=/public/unesco_collection/ programmes_isced97/school_1999_2000&vm=det

For the purpose of this survey, information on the educational level has to be transmitted following a regrouping in three broad categories:
- No formal education completed, primary or lower secondary education (corresponding to ISCED 0, 1 or 2);
- Upper secondary education (corresponding to ISCED 3 or 4);
- Tertiary education (corresponding to ISCED 5 or 6).

Only the *highest* level reached is required (not *all* the levels frequented). Only a *completed* level may be considered; 'completed' in this context is taken to imply successful termination, normally accompanied by an appropriate qualification.

Pre-primary education (ISCED level 0)

Programmes at level 0, (pre-primary) defined as the initial stage of organised instruction are designed primarily to introduce very young children to a school-type environment, i.e. to provide a bridge between the home and a school-based atmosphere. Upon completion of these programmes, children continue their education at level 1 (primary education).

Primary education or first stage of basic education (ISCED level 1)

Programmes at level 1 are normally designed on a unit or project basis to give students a sound basic education in reading, writing and mathematics along with an elementary understanding of other subjects such as history, geography, natural science, social science, art and music. The core at this level consists of education provided for children, the customary or legal age of entrance being not younger than five years or older than seven years. This level covers in principle six years of full-time schooling.

Lower secondary or second stage of basic education (ISCED level 2)

In many, if not most countries, the educational aim is to lay the foundation for lifelong learning and human development on which countries may expand, systematically, further educational opportunities. The programmes at this level are usually on a more subject-oriented pattern using more specialized teachers and more often several teachers conducting classes in their field of specialization. The full implementation of basic skills occurs at this level.

Upper secondary education (ISCED level 3)

The educational programmes included at this level typically require the completion of some 9 years of full-time education (since the beginning of level 1) for admission or a combination of education and vocational or technical experience and with as minimum entrance requirements the completion of level 2 or demonstrable ability to handle programmes at this level. The entrance age to this level is typically 15 or 16 years. More specialization may be observed at this level than at ISCED level 2 and often teachers need to be more qualified or specialized than for ISCED level 2. The programmes can be general, technical or labour-market oriented.

Post-secondary non-tertiary education (ISCED level 4)

> This item captures programmes that straddle the boundary between upper-secondary and post-secondary education from an international point of view, even though they might clearly be considered as upper-secondary or post-secondary programmes in a national context. ISCED level 4 programmes can, considering their content, not be regarded as tertiary programmes. They are often not significantly more advanced than programmes at ISCED level 3 but they serve to broaden the knowledge of participants who have already completed a programme at level 3. Typical examples are programmes designed to prepare students for studies at level 5 who, although having completed ISCED level 3, did not follow a curriculum which would allow entry to level 5.

First stage of tertiary education (ISCED level 5)

> This level consists of tertiary programmes having an educational content more advanced than those offered at levels 3 and 4. Entry to these programmes normally requires the successful completion of ISCED level 3 (esp. general or technical programmes) or a similar qualification at ISCED level 4 (esp. those preparing for entry to ISCED 5). This level includes all the research programmes which are not part of a doctorate, such as any type of Master's degree.

> The programmes can be of a general nature - i.e. largely theoretically based and intended to provide sufficient qualifications for gaining entry into advanced research programmes and professions with high skills requirements - or more technically oriented - i.e. practically oriented/occupationally specific and mainly designed for participants to acquire the practical skills and know-how needed for employment in a particular occupation or trade or class of occupations or trades (the successful completion of which usually provides the participants with a labour-market relevant qualification).

Second stage of tertiary education (ISCED level 6)

> This level is reserved for tertiary programmes which lead to the award of an advanced research qualification, such as a doctorate. The programmes are therefore devoted to advanced study and original research and are not based on course-work only. They typically require the submission of a thesis or dissertation of publishable quality which is the product of original research and represents a significant contribution to knowledge. It prepares graduates for faculty posts in institutions offering ISCED 5 (general) programmes, as well as research posts in government, industry, etc.

F4: Employment situation

Unlike the Labour Force Survey, this survey doesn't contain a number of detailed questions to assess an individual's employment situation or labour status. This variable is more related to the LFS question on the *Main labour status* where the respondent gives a self-perception regarding his/her activity status.

The explanatory notes below should help the respondent and/or interviewer in indicating the appropriate employment situation. 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 an 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 joinly-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 an 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 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).

Only the additional breakdown in *employee* and *self-employed* cannot 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 – would be 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 is 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).

F5: Occupation

[<u>Scope</u>: individuals in employment, i.e. 1st *(employee)* or 2nd option *(self-employed)* in F4]

The individual's occupation should be coded according to the ISCO-88 (COM) classification which is used in most of Eurostat's social statistics and which is based upon the ILO's International Standard Classification of Occupations, ISCO-88

(see http://laborsta.ilo.org/applv8/data/isco88e.html).

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 classification is available in the RAMON-database:

http://europa.eu.int/comm/eurostat/ramon
An introductionary note can be found at:

http://europa.eu.int/comm/eurostat/ramon/documents/isco_88_com/isco_88_com.zip)

The purpose of this variable 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. To be able to make such regroupings, it is necessary to code and store the occupations at a detailed level, at least at 3 digit level but preferably at 4 digit level. The major groups (1 digit) or submajor groups (2 digit) don't allow to distinguish ICT professionals from other workers.

Manual workers

> This category corresponds to major groups 6 to 9 of ISCO:
> - Major group 6: Skilled agricultural 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.
> - Major group 1: Legislators, senior officials and managers;
> - Major group 2: Professionals;
> - Major group 3: Technicians and associate professionals;
> - Major group 4: Clerks;
> - Major group 5: Service workers and shop and market sales workers;
> - Major group 0: Armed forces.

ICT professionals

> The most appropriate recoding will be obtained where ISCO codes are collected at 4-digit level. However, as some countries collect (or enter) the information relating to the respondent's occupation only at 3-digit level, the discussion below also includes an alternative approach covering this case.

> *ISCO available at 4-digit level*

>> The category *ICT professionals* consists of the individuals in one of the following eight ISCO Unit Groups (Unit Groups correspond to the 4-digit level):

>>> **1236**   Computing services managers
>>> **2131**   Computer systems designers, analysts and programmers

2139	Computing professionals not elsewhere classified
2144	Electronics and telecommunications engineers
3114	Electronics and telecommunications engineering technicians
3121	Computer assistants
3122	Computer equipment operators
3132	Broadcasting and telecommunications equipment operators

ISCO available at 3-digit level

The category *ICT professionals* consists of the individuals in one of the following two ISCO Minor Groups (Minor Groups correspond to the 3-digit level):

213	Computing professionals
312	Computer associate professionals

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).

F6: Geographical location

"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.

For 2007, the list of objective 1 regions (valid until end 2006) relates to the new 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 is to be extended: In EU-27, Bulgarian and Romanian NUTS regions are convergence regions.

The list of NUTS regions eligible for support from the Structural Funds under the Convergence Objective (ex-Objective 1), including phasing out Convergence regions, has been decided for the whole period 2007-2013. No update of this list is foreseen in the Structural Funds Regulation.

The new list is available on the next two pages.

Conv = Convergence region

PO = Convergence objectiverelated phasing out region (regions eligible for funding from the Structural Funds on a transitional and specific basis under the Convergence objective)

BE32	Prov. Hainaut	PO
BG00	Balgarija	Conv
CZ02	Střední Čechy	Conv
CZ03	Jihozápad	Conv
CZ04	Severozápad	Conv
CZ05	Severovýchod	Conv
CZ06	Jihovýchod	Conv
CZ07	Střední Morava	Conv
CZ08	Moravskoslezsko	Conv
DE41	Brandenburg - Nordost	Conv
DE42	Brandenburg - Südwest	PO
DE80	Mecklenburg-Vorpommern	Conv
DE93	Lüneburg	PO
DED1	Chemnitz	Conv
DED2	Dresden	Conv
DED3	Leipzig	PO
DEE1	Dessau	Conv
DEE2	Halle	PO
DEE3	Magdeburg	Conv
DEG0	Thüringen	Conv
EE00	Eesti	Conv
GR11	Anatoliki Makedonia, Thraki	Conv
GR12	Kentriki Makedonia	PO
GR13	Dytiki Makedonia	PO
GR14	Thessalia	Conv
GR21	Ipeiros	Conv
GR22	Ionia Nisia	Conv
GR23	Dytiki Ellada	Conv
GR25	Peloponnisos	Conv
GR30	Attiki	PO
GR41	Voreio Aigaio	Conv
GR43	Kriti	Conv
ES11	Galicia	Conv
ES12	Principado de Asturias	PO
ES42	Castilla-La Mancha	Conv
ES43	Extremadura	Conv
ES61	Andalucía	Conv
ES62	Región de Murcia	PO
ES63	Ciudad Autónoma de Ceuta	PO
ES64	Ciudad Autónoma de Melilla	PO
FR91	Guadeloupe	Conv
FR92	Martinique	Conv
FR93	Guyane	Conv
FR94	Réunion	Conv

ITF3	Campania	Conv
ITF4	Puglia	Conv
ITF5	Basilicata	PO
ITF6	Calabria	Conv
ITG1	Sicilia	Conv
LV00	Latvija	Conv
LT00	Lietuva	Conv
HU21	Közép-Dunántúl	Conv
HU22	Nyugat-Dunántúl	Conv
HU23	Dél-Dunántúl	Conv
HU31	Észak-Magyarország	Conv
HU32	Észak-Alföld	Conv
HU33	Dél-Alföld	Conv
MT00	Malta	Conv
AT11	Burgenland	PO
PL11	Łódzkie	Conv
PL12	Mazowieckie	Conv
PL21	Małopolskie	Conv
PL22	Śląskie	Conv
PL31	Lubelskie	Conv
PL32	Podkarpackie	Conv
PL33	Świętokrzyskie	Conv
PL34	Podlaskie	Conv
PL41	Wielkopolskie	Conv
PL42	Zachodniopomorskie	Conv
PL43	Lubuskie	Conv
PL51	Dolnośląskie	Conv
PL52	Opolskie	Conv
PL61	Kujawsko-Pomorskie	Conv
PL62	Warmińsko-Mazurskie	Conv
PL63	Pomorskie	Conv
PT11	Norte	Conv
PT15	Algarve	PO
PT16	Centro (PT)	Conv
PT18	Alentejo	Conv
PT20	Região Autónoma dos Açores	Conv
RO00	Romania	Conv
SI00	Slovenija	Conv
SK02	Západné Slovensko	Conv
SK03	Stredné Slovensko	Conv
SK04	Východné Slovensko	Conv
UKK3	Cornwall and Isles of Scilly	Conv
UKL1	West Wales and The Valleys	Conv
UKM4	Highlands and Islands	PO

The former list of Objective 1 regions in the Member States was valid until 31/12/2006.

F7: Type of locality

The concept of "urbanisation" has been introduced to indicate the character of the area where the respondent lives. The definition of the type of locality is based on the classification of *degrees of urbanisation* which is also used in the Labour Force Survey (LFS, column 176).

The following degrees of urbanisation have been identified:

Densely-populated area

> This is a contiguous set of local areas, each of which has a density superior to 500 inhabitants per square kilometre, where the total population for the set is at least 50,000 inhabitants.

Intermediate area

> This is a contiguous set of local areas, not belonging to a densely-populated area, each of which has a density superior to 100 inhabitants per square kilometre, and either with a total population for the set of at least 50,000 inhabitants or adjacent to a densely-populated area.

Thinly-populated area

> This is a contiguous set of local areas belonging neither to a densely-populated nor to an intermediate area.

(a set of local areas totalling less than 100 km², not reaching the required density, but entirely enclosed within a densely-populated or intermediate area, is to be considered to form part of that area. If it is enclosed within a densely-populated area and an intermediate area it is considered to form part of the intermediate area)

In this definition of *degree of urbanisation* there is a criterion of geographical contiguity together with a population threshold. Harmonised, comparable correspondence between the *degree of urbanisation* and NUTS 5 regions are being (re-)defined on the basis of 2001 census data (for Member States and Candidate countries) in 2005. These correspondence tables enable the NSIs to easily recode the addresses or postal codes into the three categories. In principle, this conversion table should cover all 25 EU countries at NUTS 5 (or LAU 2) level. However, for some countries there may be no results at all or the results will only be at NUTS 4 (or LAU 1) level.

An "area" consists of a group of contiguous "local areas" where a "local area" corresponds to the following entities in the respective Member States :

> **Belgium:** Gemeente / Commune
> **Czech Republic:** Obce (6 251 in year 2000)
> **Denmark:** Kommuner
> **Germany:** Gemeinde
> **Estonia:** Vald+Alev+Linn (254)
> **Greece:** Demotiko diamerisma / Koinotiko diamerisma (after the kapodistria reform, ca. 6000 units)

Spain: Municipio

France: Commune

Ireland: DED / ward

Italy: Commune

Cyprus: Demos/Koinotites

Latvia: Pagast+ Pilsetas (560)

Lithuania: Seniunija

Luxembourg: Commune

Hungary: Telepules (3 135)

Malta: Localities

The Netherlands: Gemeente

Austria: Gemeinde

Poland: Gminy+Miasta (2 486)

Portugal: Freguesias

Slovenia: Obcinah (192 since 1 Jan 1999)

Slovakia: Obce a Mesta (2 920 in year 1999)

Finland: Kunnat

Sweden: Kommune

United Kingdom: Ward

Iceland: Sveitarfélag (165 until 1997, 124 from 1998)

Norway: Kommuner (435)

Bulgaria: Naseleni miasti

Croatia: not yet available

Romania: Comuni, Municipii, Orasi

Turkey: not yet available

F8 and F9: 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 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 ICT's or Internet in a household.

F10: Household income (optional question)

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 ICT's 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 slighly 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 1/12th of the annual bonuses, whether or not including an estimate of the value of payments-in-kind, etc.).

For additional information, see *Annex Transmission format*.

II.2. PRODUCTION METHODOLOGY

This chapter is almost kept unchanged and is based on the 2006 version of the manual since the chapters remain valid for the purposes of illustration and guidance.

II.2.1. Timetable – Survey period

According to Annex II of Commission Regulation (EC) No 1031/2006 of 4 July 2006 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 2007" (see Annex **Error! Reference source not found.**). 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 preceeding 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 recommandation 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.

The table and figure below present an overview of the survey periods in the 2005 survey.

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.

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.

Table: Overview of the survey periods (2005 survey)
Source: Interim reports for the Community survey on ICT usage in households and by individuals (2005)

Belgium	Second quarter
Czech Republic	January - March
Denmark	February, March and April *(2004 survey)*
Germany	End of April - June
Estonia	11 April - 17 July
Greece	1 April - 31 May
Ireland	6 June - …
Italy	14 February - 28 February
Cyprus	April - June

Latvia	23 April - 31 May
Lithuania	1 March - 30 April
Luxembourg	April - May
Hungary	1 April - 21 April
Malta	April - June
Netherlands	1 April - 31 May
Austria	February - May
Poland	1 April - 20 April
Portugal	18 April - 27 May
Slovenia	4 April - 22 April
Slovak Republic	June - July
Finland	29 March - 6 May
Sweden	11 April - 19 June
United Kingdom	21 February - 11 March and a 3 weeks' period in May
Iceland	2 or 3 weeks end February - mid March
Norway	4 April - 8 May

Figure: Overview of the survey periods (2005 survey)
Source: Interim reports for the Community survey on ICT usage in households and by individuals (2005)
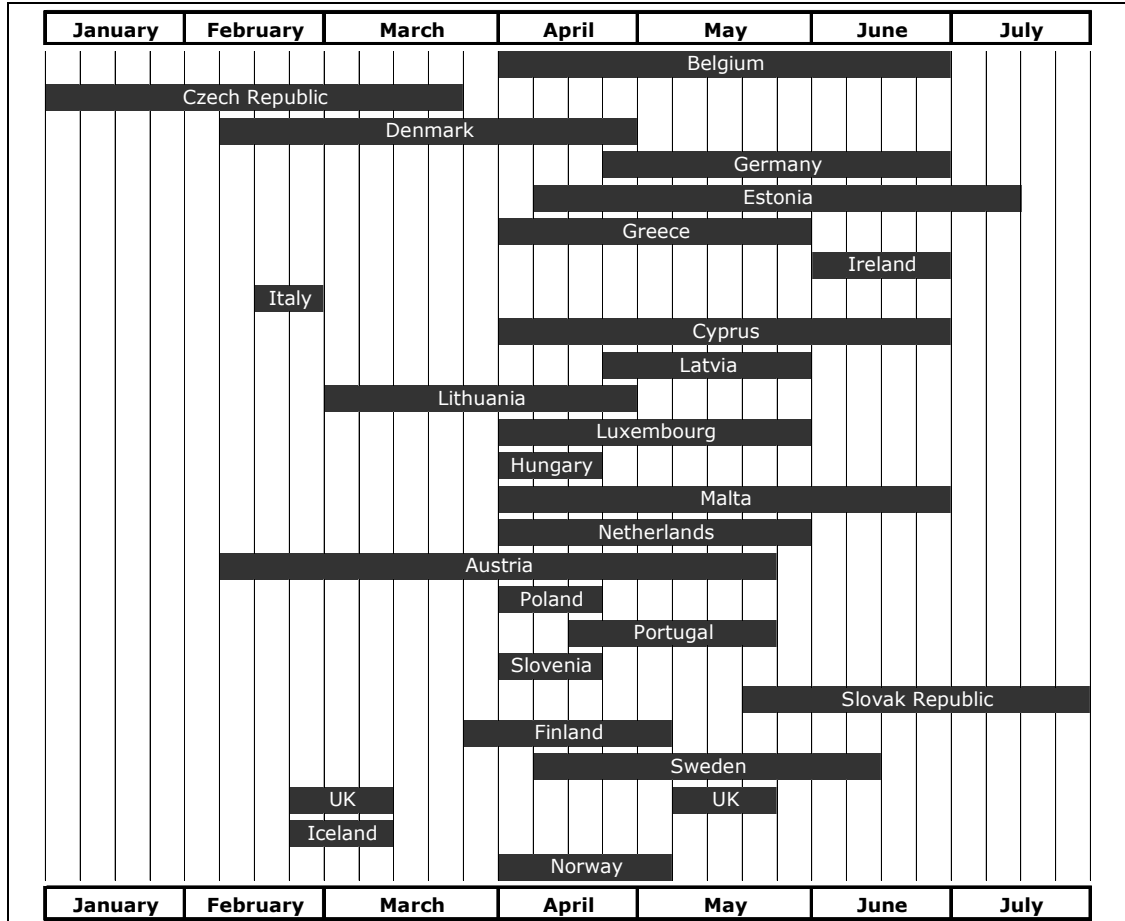
II.2.2. Frame population

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 2005 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 (2005 survey)
Source: Quoted from the Interim reports for the Community survey on ICT usage in households and by individuals (2005)

Belgium	Population register (contains all residents of Belgium, regardless of nationality, together with information on household)
Czech Republic	Register of census areas
Denmark	Central Population Register (contains general personal data on individuals with residence in Denmark; individuals without registered address in the CPR do not enter the survey)
Germany	No sampling frame (quota sampling)
Estonia	Since 2005 the new sampling frame, the Population Register, is used. Therefore in the second quarter of 2005 50% of households are sampled from the old frame, which is 2000 Population and Housing Census database, and 50% of households from the Population Register. 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.

Greece	The sampling frame containing the primary units (cluster of households in one or more unified blocks) for the ICT survey is the same as for the LFS of the year 2004, which is an area frame constructed using the necessary information from the recent Greek General Population Census 2001, and provides complete coverage of the target population of this survey. In each selected primary unit, the sampling frame containing the households is updated on month before the data collection.
Ireland	Census of Population
Cyprus	Population Census 2001 register
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 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.
Lithuania	Population Register
Luxembourg	It will be worked with RDD (random digital dialling) by using computer generated national phone numbers. There will be a number of phone numbers randomly chosen to be called. In order to reach representative results, a quota system for the individuals to be interviewed will be applied.
Hungary	The stock or dataset of the addresses of the 2001 population and housing census was used. This was augmented by the set of addresses of new constructions built between 1st of February 2001 and the middle of 2004.
Malta	Database maintained by the NSO containing all households and individuals in Malta and Gozo.
Netherlands	Municipal Population Register. 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. Every day information about birth, death, move (house), new addresses et cetera is coming in. 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
Austria	The "Zentrales Melderegister (ZMR)" is used as a sampling frame for the Continuous Labour Force Survey. A sub-sample of this sample was drawn for the ICT survey.
Poland	The basis for selection of Primary Sampling Units (PSU) and dwellings is 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.
Portugal	This survey uses a sample selected from Master-sample (Census 2001). It constitutes the sampling basis to all households and individuals related surveys done by INE.
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.

Slovak Republic	Population and housing census is organised by So SR in 10year periodicity. The last one was realised in May 2001. On base of census results the information on whole population is 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 is created on the state of population as of 31 December 2003.
Finland	Updated Central Population Register. It contains the official information of Finnish citizens and aliens residing permanently in Finland, and on buildings and residences and their holders.
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 the Tax Authorities.
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.
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.
Norway	BEBAS - the central demographic-/population database at Statistics Norway. The database is updated from The National Population Register several times a month.

Table: Overview of the target and non-target population (2005 survey)

Source: Interim reports for the Community survey on ICT usage in households and by individuals (2005)

	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				4 000 000	400 000	91%
Denmark (2004)	3 933 851	1 400 000	74%	2 307 329	350 000	87%
Germany	62 000 000	19 000 000	77%	35 000 000	3 000 000	92%
Estonia	1 026 000	330 000	76%			
Greece	8 160 045	2 440 939	77%	3 726 806	415 801	90%
Ireland	3 002 900	1 040 900	74%	1 301 000	104 900	93%
Italy	44 000 000	13 500 000	77%	20 000 000	2 300 000	90%
Cyprus	524 151	206 216	72%	220 695	18 105	92%
Latvia	1 778 195	541 008	77%			
Lithuania	2 575 400	870 500	75%	1 232 700	113 300	92%
Luxembourg	333 866	117 734	74%	156 451	15 484	91%
Hungary	7 720 917	2 218 474	78%			
Malta	296 400	103 400	74%			
Netherlands	12 700 000	3 300 000	79%			
Austria	6 050 000	2 000 000	75%	3 080 000	340 000	90%
Poland	28 503 984	9 726 096	75%	12 592 289	744 751	94%
Portugal	7 873 696	2 511 810	76%			
Slovenia	1 551 637	429 676	78%			
Finland	3 941 000	1 295 000	75%	2 405 000	229 500	91%
Sweden	6 500 000	2 500 000	72%			
United Kingdom	42 873 728	15 699 620	73%	21 491 569	3 211 384	87%
Iceland	207 000	86 000	71%	103 000	8 300	93%
Norway	3 273 926	1 330 818	71%			

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 (e.g. 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).

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 coefficient of variation (or relative standard error) should not exceed 2% for the overall proportions and should not exceed 4% for the proportions relating to the different subgroups of the population where these subgroups constitute at least 5% 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.

Table: Overview of the sampling design and sampling methods (2005 survey)
Source: Quoted from the Interim reports for the Community survey on ICT usage in households and by individuals (2005)

Belgium

    Sample of households (through "Reference persons")
    Stratification by NUTS2 (11 strata)
    Two-stage sampling:
- 598 (each quarter) Primary Units (= parts of municipalities), selected with probability proportional to size, repetitions allowed
- Secondary Units = households (23 for each PSU selection; exception: 26 in the Brussels stratum)

    Initial sample size 14027 households (some 35 000 individuals)
    Individual within household: Interviewer chooses person aged 16-74 with most recent birthday at the date of interview.

Czech Republic

    Stratified random sampling
    All individuals aged 15+ in the household are interviewed.
    Post-stratification by gender, age and region was used.

Denmark

    The sample is formed by a random selection from the CPR. As stated earlier the survey is conducted over 3 months with a new sample being drawn every month. When the sample has been drawn, the individuals taken part 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. In this survey round a qualified sample of 4.006 individuals was formed (February through April). Total interviews for analysis (excluding loses etc.) ended up at 2.899.
    After the data has been collected, the data is grouped by gender, age (4 groups), civil status (5 groups) and gross income (3 groups + unknown/don't know).

Germany

    The ICT survey is based on a representative quota sample. The representativeness was achieved
    by a stratification frame taken from German Microcensus 2003.
    Stratification variables are:

- Federal State (16 Bundesländer);
- Household Type (5 types): 1) single adult; 2) couple without children; 3) single parent with at least one child aged less than 18; 4) couple with at least one child aged less than 18; 5) other household;
- Social status of head of the household (6 types): 1) self employed; 2) civil servant; 3) employee; 4) worker; 5) retired – only for household type 1 and 2 – ; 6) other inactive;
- Household net monthly income (4 classes): 1) < 1300 EUR; 2) 1300 to<2600 EUR; 3) 2600 to < 3600 EUR; 4) > 3600 EUR;

That is 108 strata for each federal state (Bundesland).

Sampling unit is the private household at main residence. Households are selected either from an access panel or from other address databases (i.e. from household budget survey) that give also information on the stratification variables.

In the first step, 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 10 years and over. Then the household questionnaire and the correct number of personal questionnaires is sent to the household. All persons 10 years and over living in the household are selected respondents. There is no second step of sampling within the households. When the participation rate in a certain stratum is low, other households with the same stratification characteristics can be recruited to fulfil the quota restrictions.

To achieve results broken down by "Household type" or "Social status of head of household", the categories of these two stratification variables are taken disproportionally into the sampling frame. In case of "Household type" the categories "couples with children" and "single parents" are overrepresented. In case of " Social status of head of household" the categories "self employed", "civil servant", "worker" and "other inactive" are overrepresented. This disproportional distribution is going to be corrected in the grossing up procedure by adjusting the sample to the microcensus.

Estonia

The sampling design is a stratified systematic two-phase sampling of individuals, whose households are included in the sample.

In the first phase the sample is selected by systematic sampling inside strata and the information on the size of sampled households is collected. In the second phase the sample is grouped by the number of persons aged 15-74 years in the household of the sampled individual. The final sample is then selected by systematic sampling from each household's size group with inclusion probability inverse to the household size. This yields an equal probability sample of households (and its 15-74 years old members) inside strata. All persons aged 15-74 years in the households of the final sample are interviewed.

A stratified systematic sampling is used on the first step of sample design. The 15 counties of Estonia and Tallinn are divided into four strata according to the population numbers (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.

Greece

The two-stage stratified area sampling with self-weighting estimators is applied for the survey. The primary units are the areas (one or more unified blocks) and the final sampling units selected in each sampling area are the households, apart from households not containing at least one member 16 to 74 years old.

Stratification

In each geographical region (NUTS II), the urban agglomerations, the Municipalities and Communes were allocated by the degree of urbanization (urban, semi-urban and rural regions).

The Municipalities of the Greater Athens (the largest urban agglomeration) were allocated to 31 approximately equally sized strata (equal number of households), and the Municipalities belonging to the agglomeration of Thessalonica (the second urban centre of Greece) were allocated to 9 equally sized strata, taking into account social and economic criteria for this further stratification. After the stratification, 79 strata

were produced.

Sampling fraction

The overall sampling fraction is $\dfrac{1}{\lambda} \approx 0,144\%$ (5.500 households).

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 sampling, 15 households are surveyed in each block to give a total quarterly sample 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 (20%) 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. […]

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 for each district (urban areas) respectively.

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 neighbouring 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.

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), six other 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.

All individuals aged 16-74 (15+) will be selected in each sampled household.

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

Stratified two-stage sample design. The biggest towns of Lithuania counties, small and medium towns and rural area were divided into separate strata. Simple random sample of individuals of 16 and older from Population Register was drown in the biggest towns of the counties. A household, which lives in selected persons address, was surveyed. Two–stage sample design was used in small and medium towns. The Pareto sample with probability proportional to the clusters size was used in the first stage. Each town was a cluster. Simple random sample of persons of 16 and older was drowning from Population register in already selected cluster in second stage. In rural area the similar sample design like in medium and small towns was used. Cluster was a rural territorial unit in this case.

Luxembourg

The national sample was constructed in 2 stages:
1rst level: selection of households by RDD.
2nd level: selection of respondent in the household. 1 individual per household was selected.
Quota sampling: age, sex, region, nationality and professional activity
A weighting procedure for these quota variables will be applied (including sex by age). The reference figures will be taken from STATEC.

Hungary

The ICT sample is a sub-sample of the sample of the micro-census whose reference period is 1st April 2005. 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. 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.
When selecting the ICT sample from the micro-census sample, an effort was made to keep is close to a self-weighting sample.
When selecting the sample of localities within the GEDs, we were able to compare population counts estimated from the sample with the corresponding actual values of 1 January 2004. In the case of extremely large sampling errors, the selection procedure was repeated.

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. In the second stage the number of individuals, as determined for every municipality, are selected. Individuals aged 12 to 74 year are selected. Variables used for stratification are NUTS3 and regions used for the fieldwork

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. 8.000 households has been drawn for which the "reference week" of the survey are the calendar weeks 6 to 13 in the year 2005. The sample is stratified by region. All household members between 16 and 74 years in the sampled households are surveyed.

Poland

A two-stage sampling scheme was adopted with stratification on the first stage. At the

first stage of sampling design the same ASPs will be used as in a sample for Household Budget Survey (HBS). The sample already exists. ASPs 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. There are 96 strata (of which 31 covering rural areas).

The number of the ASPs selected to the HBS 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. The procedure of systematic sampling was applied after random ordering of the units (ASPs selected with PPS using Hartley-Rao method). 675 ASPs were selected and that is the first stage sample for HBS and ICT usage surveys.

On the second stage 12 dwellings will be selected in each ASP by means of simple random sampling method, but additional condition will be applied: dwellings selected to the HBS cannot be selected to the ICT usage survey. Samples for the surveys will be separated. Gross sample size for ICT usage surveys: 12 * 675 ASP = 8100 dwellings.

Portugal

The sample used for the household survey on ICT usage is directly selected from Master-sample and for this specific subject, and inquires, in a first approach, 6 026 dwellings by sampling method and around 10 500 individuals (estimate) who are 16-74 years old.

The sample is representative at NUTS I level: Portugal (Mainland), Azores and Madeira.

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 200 sampling units are selected with the probability proportional to size (PPS with replacement), and at the second stage 10 persons aged between 10 and 74 years are selected. Thus 2000 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:

$$\frac{N_a}{\sum_{Slo} N_a} \cdot \frac{10}{N_a},$$

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

The sample is realised on the base of the same methods, which are used in other surveys organised by SO SR (census, inquiry methods face-to-face)

Sample contains 4000 households including households with member's aged 16 to 74. The breakdown of households is following: with 1 member, with 2 members, with 3 members, with 4 members, with 5 and more members.

Two –stage sampling is realised:

- Number of persons in households
- Location

The households are selected inside of each strata random sample.

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 sampling method can be compared to the use of simple random sampling with the addition that the sample is geographically self-weighting. The target area is the whole country, and the respondents represent the 10 to 74-year-old population in Finland, according to age, gender, province, and native language. Two samples are drawn. The other sample consists of 15 to 74 year-old population and the other sample consists of 10 to 14-year-old population.

Sample size is based on the knowledge of response rates in respective surveys earlier. After drawing the sample from the Central Population Register we search with different

methods the telephone number for the target, 93 to 95 per cent of them are within reach of the telephone. Fixed telephone, if any, will be used as the primary way of communication. During the April 2004 interviews half of the interviews were given from fixed telephone numbers and another half from mobile numbers. A couple of weeks before the contact by phone, Statistics Finland sends out a target letter telling him or her briefly the contents and the meaning of the survey. The sample population without telephone number (5 to 7 per cent of the total sample) will receive their own letter by post, indicating that they are asked kindly to contact Statistics Finland. Around 10 per cent will normally make contact and give an interview.

Sweden

The LFS-sample is drawn at the end of the first 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, citizenship (Swedish, foreign), and employment (employed, not employed) as appears in the Register of Regional Employment. In this way 192 strata are constructed. Within each stratum then population is ordered by age, and a systematic probability sample is drawn. The sampling differs from different regions according to NUTS III and between Swedish and non-Swedish citizens. Within regions and each of the two citizen groups the sample size is directly proportional to stratum size.

The LFS sample consists of three separate samples, one for each month in the quarter. Each of the samples, which consists of about 20 000 people, is rotated in such a way that an eighth leaves the sample between two survey cycles. For every sample this occurs every three months. 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 sample for the Survey on ICT use is made by selecting all individuals in the third wave and a few individuals in the fourth to seventh wave. The ICT survey is conducted at the end of the interview, i.e. after the LFS questions.

United Kingdom

In February 2005, a sample of 100 postal sectors will be 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 will be selected with probability proportionate to size and, within each sector, 30 addresses (delivery points) will be selected randomly. If addresses contain more than one household, the interviewer will use 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 will be selected with the use of random number tables. The interviewers endeavour to interview that person – proxy interviews will not be taken.

In May 2005, a sample of 67 postal sectors will be selected and stratified by: region, the proportion of households with no car, the proportion of households where the household reference person is in the National Statistics Socio-economic Classification (NS-SEC) categories 1 to 3, and the proportion of people who are aged over 65. The postal sectors will be selected with probability proportionate to size and, within each sector, 30 addresses (delivery points) will be selected randomly. If addresses contain more than one household, the interviewer will use 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 will be selected with the use of random number tables. The interviewers endeavour to interview that person – proxy interviews will not be taken.

Iceland

Stratified random sampling without replacement is used to select individuals. Population is devided into two strata by residence. Within households with more than one member in the age of 16-74 years, one person is randomly selected, i.e. each individual represents one household.

Norway

The sampling follows the Statistics Norway's standard sample procedure. The country is first divided into a number of sampling areas and these again are divided into 109 strata. The sampling areas are municipalities or a group of municipalities. The municipalities with few inhabitants are put together with other municipalities to ensure

that each sampling area consists of at least seven per cent of the total inhabitants in the stratum the area belongs to. In some cases small municipalities close to highly populated municipalities are put together with the large one in that region. All municipalities with more than 30 000 inhabitants and some with 25 000 to 30 000 inhabitants make separate strata.

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 for the 2004 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 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 coefficient of variation) are discussed in more detail in the next chapters.

Table: Overview of the sample sizes and response rates (2004 survey)

Source: Final reports for the Community survey on ICT usage in households and by individuals (2004)

	Unit	Gross sample size	Ineligible: out-of-scope	Other ineligible	Number of eligible elements	Non-contact	Refusal	Other non-response	Net sample size	Unit response rate
Denmark	IND	4 499		493	4 006	336	445	326	2 899	72.4%
Germany	HH	6 590	171		6 419			1 692	4 727	73.6%
Greece	IND	4 970	427	0	4 543	207	156	0	4 180	92.0%
Spain	HH	19 366		2 034	17 332	1 256	2 065	81	13 930	80.4%
Ireland	HH	7 190	24		7 166		1 019	726	5 421	75.6%
Latvia	IND	7 992	44	298	7 650	683	330	129	6 508	85.1%
Lithuania	HH	4 500	250	100	4 150	210	223	39	3 678	88.6%
Luxembourg	IND	38 767	4 160	8 959	25 648	15 397	8 056	660	1 535	6.0%
Hungary	IND	6 500	0	143	6 357	1 290	1 034	173	3 860	60.7%
Austria	HH	6 294	182	235	5 877	931	1 674	0	3 272	55.7%
Poland	IND	12 202	3 027		9 175	179	36	121	8 839	96.3%
Portugal	IND	14 115	1 099	1 644	11 372	778	251	0	10 343	91.0%
Slovenia	IND	2 000	0	193	1 807	142	184	26	1 455	80.5%
Slovak Republic	IND	5 000	0	0	5 000			259	0	4 741	94.8%
Finland	IND	4 127	0	26	4 101	498	302	167	3 134	76.4%
Sweden	IND	5 429	0	9	5 420	410	589	795	3 626	66.9%
United Kingdom	IND	6 000	411	481	5 108	629	1 427	79	2 973	58.2%
Bulgaria	HH	4 614	32	130	4 452	143	60	0	4 235	95.1%
Romania	HH	10 816		246	10 568	132	167	222	10 047	95.1%
Iceland	IND	2 000	0	50	1 950	116	152	49	1 633	83.7%
Norway	IND	2 000	109	21	1 870	146	458	78	1 188	63.5%

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 are used by only two countries.

Table: Overview of the data collection methods (2005 survey)
Source: Interim reports for the Community survey on ICT usage in households and by individuals (2005)

Belgium	Face-to-face (for persons aged 65+: face-to-face or telephone)
Czech Republic	Computer-assisted Personal interview
Denmark	CATI *(2004 survey)*
Germany	Self-administered mail questionnaire
Estonia	Face-to-face
Greece	Telephone interviews
Ireland	CAPI
Italy	The survey uses paper and pencil interview (PAPI) technique and it is based on a face-to-face interview questionnaire and a self administrate questionnaire. The ICT module is in the self administrate questionnaire.
Cyprus	Face-to-face using CAPI
Latvia	Face-to-face
Lithuania	Face-to-face interview. Household members not at home during the visit of interviewer can be contacted later by telephone. Proxy interviews are allowed only for persons who don't use computer or Internet.
Luxembourg	CATI (with computer generated phone-numbers (mobile phones not included))
Hungary	Face-to-face
Malta	Face to Face interview with initial telephone call to fix appointment
Netherlands	CATI
Austria	CATI

Poland	The main survey technique will be face to face interview. Interviewers are supposed to fill in the household questionnaires together with heads of households or other persons who will be able to give credible answers. Then all members of the household aged 12-74 will be interviewed by means of individual questionnaires. In case of some members' absence, interviewers will visit the respondents once again in order to complete the missing questionnaires. If more convenient, the forms may be left to the respondents for filling in by themselves. Interviewers will gather these questionnaires on prearranged day.
Portugal	The survey method will be face-to-face interviews, with a portable computer, for the household representative and for the individuals who are 10-74 years old. For the cases where the interview to the own person at the dwelling is not possible, we won't admit the adoption of proxy interviews (information conceded by another individual of the household, which is present at the moment of the interview). In such cases we'll use the telephone survey, in order to surpass the proxy interviews and ensure a high response rate.
Slovenia	Face-to-face
Slovak Republic	Face-to-face
Finland	CATI
Sweden	Telephone interviews
United Kingdom	All interviews are carried out face-to-face by members of the general field force of the interviewers trained to carry out National Statistics surveys. Advance letters are sent to all addresses, prior to the interview, giving a brief account of the survey. The interviewing period starts during the last two weeks of the month and continues into the first week of the following month. Interviewers call at the selected addresses unless a refusal has been made beforehand in response to the advanced letter. The interviewer makes at least three calls at an address at different times of the day and week before coding the household as a non-contact.
Iceland	CATI
Norway	CATI

Independent versus embedded survey

For practical reasons, an important number of countries has 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 survey vehicles (2005 survey)

Source: Interim reports for the Community survey on ICT usage in households and by individuals (2005)

Belgium	Labour Force Survey
Czech Republic	Labour Force Survey
Denmark	Annex to the monthly Danish consumer survey *(2004 survey)*
Germany	Stand-alone survey
Estonia	Labour Force Survey
Greece	Stand-alone survey
Ireland	The ICT survey was attached to the Quarterly National Household Survey, which collects information continuously throughout the year, with a total sample of 39,000 households in each quarter. The QNHS is the national title given to the continuous quarterly Labour Force Survey carried out in accordance with Council Regulation 577/98. The ICT questionnaire was asked of direct respondents in wave 3 households in June of 2005, which constituted approximately one-fifth of the total sample of households.
Italy	The module on ICT usage in households and by individuals is embedded in the annual multipurpose survey "Aspects of Daily Life" that is the main survey of Istat's integrated system of multi-purpose social surveys. This annual survey is the supporting and normalising element of the whole social informative framework. It collects a set of data concerning individuals, households and events which affords to construct and analyse the citizen's demand, besides comparing it with services supply, already surveyed by Istat.
Cyprus	Stand-alone survey
Latvia	Stand-alone survey
Lithuania	Stand-alone survey
Luxembourg	The EU-ICT-HH-survey is embedded in another survey called "e-luxembourg", which is about the usage of ICT in the Grand-Duchy of Luxembourg. The questions of "e-luxembourg" follow the "e-Europe"-indicators. "e-luxembourg" is a regular project ordered by the Luxembourg's government since 2001. In order to continue the existing national statistics, it is necessary to conduct those questions of "e-luxembourg" as well, that were originally not included in the EU-ICT-HH-survey. The national authorities decided to merge the two surveys in order to catch synergies and to lower the response burden for the households.
Hungary	Supplementary survey to the micro-census 2005. The socio-demographic background characteristics of the households and the individuals of the sample will originate from the results of the micro-census.
Malta	Stand-alone survey
Netherlands	Stand-alone survey
Austria	Labour Force Survey
Poland	Stand-alone survey
Portugal	Stand-alone survey
Slovenia	Stand-alone survey
Slovak Republic	Stand-alone survey
Finland	Stand-alone survey
Sweden	Labour Force Survey

United Kingdom	'Community Survey on ICT usage in households and by individuals 2005' is a module referred as 'Internet access' within the National Statistics Omnibus Survey. The National Statistics Omnibus Survey is a multi-purpose survey developed by the Office for National Statistics for use by government departments, other public bodies, charities and academics. It provides a fast, cost-effective and reliable way of obtaining information on topics too brief to warrant a survey of their own, and for topics of immediate interest. Its flexibility allows this survey to be used for piloting questions for other surveys.
Iceland	Stand-alone survey
Norway	Statistics Norway's Omnibus surveys embrace questions on different themes, on behalf of various customers. The surveys consist of a permanent set of background variables, in addition to the questions of the customers. In the second Omnibus in 2005 questions on use and access to ICT is included.

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 proces usually makes the collection period longer as one needs 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 (see also below, 0).

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 will be mandatory as of 2006 in all EU countries following the legal acts (Regulation 808/2004 and Regulation 1099/2005).

Table: Overview of the modes of survey participation (2005 survey)
Source: Interim reports for the Community survey on ICT usage in households and by individuals (2005)

Belgium	Voluntary module in mandatory LFS
Czech Republic	Voluntary
Denmark	Voluntary *(2004 survey)*
Germany	Voluntary
Estonia	Voluntary
Greece	Mandatory
Ireland	Voluntary
Italy	Mandatory
Cyprus	Mandatory
Latvia	Mandatory
Lithuania	Voluntary
Luxembourg	Voluntary
Hungary	Voluntary
Malta	Mandatory
Netherlands	Voluntary

Austria	Voluntary
Poland	Voluntary
Portugal	Mandatory
Slovenia	Voluntary
Slovak Republic	Mandatory
Finland	Voluntary
Sweden	Voluntary
United Kingdom	Voluntary
Iceland	Voluntary
Norway	Voluntary

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 preceeding (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. 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[9]

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 (§2.7.2.1): 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 households and/or individuals that responded to the survey to compensate for those that did not respond.

Partial non-response or item non-response (§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.

---

[9] 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.

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.

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 2004 survey (based on the 2004 Final Reports) can be found above (II.2.3.2, Sample size). It is clear that in most 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 socioeconomical (employment, income) variables from administrative data.

More advanced methods for creating weighting classes is 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 (poststratification 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.

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 indicator 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 degressive 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).

I.1.1.1   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. 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. 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 (see also Chapter 9 of the *Reporting template*, Annex II.3.3). The estimation of the sampling variance should ideally take into account the sampling design (e.g. the stratification).

Some guidelines for interpretation and use of the coefficients of variation can be found in section II.2.3 where the sampling design is discussed.

The computated accuracy measures will also determine the reliability of the estimates. If the coefficient of variation for a certain proportion or absolute value is too high, the cell needs to be suppressed.
To avoid that the level of the estimate 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) could be used as a reference measure.
In case the computation of the standard deviation or coefficient of variation 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).

The table on the next page gives an overview of the sampling error for one proportion, namely the proportion of individuals regularly using the Internet (i.e. at least once a week in the three months preceeding the interview). The table clearly indicate the parameters that affect the sampling error: the estimate (i.e. the actual proportion, in this case the level of Internet use) and the sample size (though only the total sample size for the survey is presented in the table – between brackets next to the country name).

Although the overall proportions (i.e. for all individuals) are relatively accurate in all countries, 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.

Table: Sampling error for the proportion of regular Internet users (2004 survey)
Source: Final reports for the Community survey on ICT usage in households and by individuals (2004)

(next page)

Country	Proportion of individuals regularly using the Internet								
(survey net sample size)	All	Men	Women	16-24 yr	25-34 yr	35-44 yr	45-54 yr	55-64 yr	65-74 yr
Denmark (2899)									
Estimated proportion	70.0	73.0	66.0	82.0	79.0	80.0	73.0	59.0	26.0
Coefficient of variation	1.2	1.6	1.9	2.8	2.4	1.9	2.6	3.6	9.0
Germany (8887)									
Estimated proportion	50.0	55.0	44.0	78.0	72.0	60.0	49.0	27.0	13.0
Coefficient of variation	1.3	2.5	2.6	5.0	7.3	4.4	6.2	11.8	35.7
Greece (4180)									
Estimated proportion	17.3	59.5	40.5	33.4	30.4	23.9	9.7	2.0	0.7
Coefficient of variation	3.5	2.9	4.2	5.1	5.3	5.8	10.8	24.4	41.1
Spain (n/a)									
Estimated proportion	40.4	44.9	35.9	75.5	57.6	43.4	29.7	13.7	2.9
Coefficient of variation	1.3	1.7	1.7	1.9	2.1	2.5	4.0	7.1	15.6
Ireland (4638)									
Estimated proportion	37.0	37.0	37.0	49.0	48.0	43.0	32.0	17.0	6.0
Coefficient of variation	3.4	4.4	3.4	4.7	4.3	3.4	4.9	9.5	11.9
Latvia (6508)									
Estimated proportion	27.3	27.6	27.0	53.9	41.2	28.9	19.2	8.0	1.3
Coefficient of variation	4.8	5.0	5.4	3.2	5.5	6.3	11.7	15.4	61.2
Luxembourg (1535)									
Estimated proportion	59.0	34.0	25.0	11.0	13.0	16.0	12.0	6.0	2.0
Coefficient of variation	1.6	1.4	2.0	3.9	3.4	3.0	4.3	8.3	23.2
Hungary (3860)									
Estimated proportion	28.6	31.1	26.6	71.6	45.1	33.4	24.0	11.1	2.9
Coefficient of variation	2.5	3.5	3.6	3.0	4.5	5.9	6.5	11.1	26.3
Austria (5646)									
Estimated proportion	46.1	52.9	39.4	69.3	65.3	53.0	42.1	23.6	6.6
Coefficient of variation	0.9	1.2	1.1	2.0	2.1	1.7	2.0	1.6	1.1
Poland (8839)									
Estimated proportion	22.4	23.6	21.4	49.5	30.7	18.0	14.6	7.3	1.5
Coefficient of variation	2.6	3.1	3.4	2.9	4.6	6.5	6.4	12.5	31.2
Portugal (10343)									
Estimated proportion	25.1	27.1	23.2	54.0	36.9	25.2	17.5	7.0	1.1
Coefficient of variation	1.4	1.8	1.4	5.9	3.9	2.6	2.2	1.2	0.6
Slovenia (1455)									
Estimated proportion	33.4	35.6	31.3	66.4	53.9	31.2	26.3	8.2	1.0
Coefficient of variation	4.5	6.1	6.3	5.5	7.5	10.5	10.8	21.3	71.8
Slovakia (4741)									
Estimated proportion	39.8	21.6	18.2	12.3	9.8	8.4	7.5	1.5	0.4
Coefficient of variation	1.6	2.3	2.7	4.1	3.3	3.2	4.0	7.6	21.5
Finland (3134)									
Estimated proportion	72.4	72.3	72.5	97.5	95.6	86.8	71.3	55.3	13.5
Coefficient of variation	1.1	1.6	1.5	0.7	1.0	1.6	2.6	3.7	1.3
Sweden (3626)									
Estimated proportion	75.0	78.0	73.0	92.0	87.0	82.0	76.0	64.0	41.0
Coefficient of variation	1.0	1.2	1.4	1.2	1.4	1.7	2.1	2.8	19.5
United Kingdom (2973)									
Estimated proportion	77.0	82.0	72.0	78.0	80.0	82.0	72.0	75.0	67.0
Coefficient of variation	2.0	2.0	2.0	4.0	3.0	3.0	4.0	4.0	9.0
Bulgaria (9785)									
Estimated proportion	13.5	7.1	6.4	5.4	3.8	2.2	1.6	0.5	0.1
Coefficient of variation	3.8	5.5	5.8	6.4	7.7	10.1	12.0	22.1	53.3
Romania (3317)									
Estimated proportion	9.9	52.7	47.3	40.8	25.2	17.9	13.3	2.5	0.3
Coefficient of variation	3.6	1.9	2.1	3.2	5.0	5.6	7.1	17.1	45.8
Iceland (1633)									
Estimated proportion	77.0	79.0	75.0	91.0	87.0	82.0	75.0	59.0	29.0
Coefficient of variation	1.2	1.6	1.8	1.6	1.9	2.3	3.0	5.3	12.0
Norway (1188)									
Estimated proportion	77.7	80.8	74.5	98.4	92.1	88.6	80.7	57.9	22.2
Coefficient of variation	1.3	1.8	1.9	0.9	2.0	2.2	3.4	6.0	17.4

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:

l. General methodological information: reference period(s); survey period; survey vehicle, where applicable; survey type; pre-tests; methodological differences compared to previous data collection exercises.

m. 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.

n. Information on the national questionnaire.

o. Sampling frame : name and description of the sampling frame or register used ; known shortcomings of the sampling frame, if any.

p. Sampling design: sampling method; additional measures taken at the time of sampling design to improve representativeness.

q. 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.

r. Item non-response: variables or items with item response rates below 90%; methods used for dealing with item non-response.

s. Grossing-up procedures for households.

t. Grossing-up procedures for individuals.

u. Sampling error: the report shall indicate the coefficient of variation (relative standard error) for a selected group of indicators or subindicators listed in the template to be provided by the Commission.

v. Problems encountered and lessons to be learnt.

w. 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.

The *Reporting template* used for the European ICT usage surveys, including some instructions, can be found in Annex II.3.3
(see document MM2006_AnnexII.3.3_ReportingTemplateHH2005rev1.doc (14 pages))

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 III.1) and in Annex II, chapter 6 of Regulation 1099/2005 (see Annex **Error! Reference source not found.**).

II.3.   ANNEXES

II.3.1.   Model questionnaire

Community survey on
ICT usage in households and by individuals

2007

Eurostat Model Questionnaire (v3.0)

Module A	Access to Information and Communication Technologies

Note : this module is directed to the household

A1   Do you or anyone in your household have access to a computer at home?
(any type: desktop, laptop, palmtop)

Yes ....☐
No ....☐

[-> go to A2]

**A2   Do you or anyone in your household have access to the Internet <u>at home</u>,
regardless of whether it is used?**

Yes ....☐   **[-> go to A3]**
No ....☐   **[-> go to B1]**
Don't know   ☐   **[-> go to B1]**

A3 **On which of the following devices is the Internet accessed at home?**
(tick all that apply)

    a) Desktop computer .... ☐

    b) Portable computer (laptop) .... ☐

    c) TV set with specific Internet device (e.g. digital TV or set top box) .... ☐

    d) Games console .... ☐

    e) Other means (i.e. none of the above) .... ☐

        of which: (optional)   **e1)** via Internet enabled mobile phone (WAP, GPRS, UMTS, etc.)....... ☐

                            e2) via handheld computer (palmtop, PDA) .... ☐

    f) Don't know .... ☐

[-> go to A4]

A4 **What types of Internet connection are used?**
(tick all that apply)

    a) Modem (dial-up access over normal telephone line) or ISDN .... ☐

    b) DSL (e.g. ADSL, SHDSL, etc.) .... ☐

    c) Other broadband connection (e.g. cable, UMTS, etc) .... ☐

    d) Mobile phone over narrowband (WAP, GPRS, etc.) .... ☐

[<u>not</u> ticked *b* or *c* -> go to A5]

[ticked *b* or *c* -> go to B1]

A5 **What are the reasons for not having a broadband connection (such as cable or DSL) at home?**
(tick all that apply)

    a) Too expensive ....…............ ☐

    b) No need ....…...................................…................ ☐

    c) Not available in my area ....…............................ ☐

    d) Can access broadband somewhere else (e.g. at work) .... ☐

    e) None of the above, but other ....…........................ ☐

[-> go to B1]

eurostat

Module B: Use of mobile phone and computers

Note : this and the following modules are directed to the individual selected within the household

---

B1   When did you most recently use a computer? *(filter question)*

Within the last 3 months .... ☐ **[-> go to B2]**

Between 3 months and a year ago .... ☐ **[-> go to B4]**

More than 1 year ago .... ☐ **[-> go to B4]**

Never used one .... ☐ **[-> go to B4]**

B2   How often on average have you used a computer in the last 3 months?
(tick one)

Every day or almost every day .... ☐

At least once a week (but not every day) .... ☐

At least once a month (but not every week) .... ☐

Less than once a month .... ☐

[-> go to B3]

B3   Where have you used a computer in the last 3 months?
(tick all that apply)

a) At home .... ☐

b) At place of work (other than home) .... ☐

c) At place of education .... ☐

d) At another person's home .... ☐

e) Other (e.g. hotel, airport, internet café, etc.) .... ☐

[-> go to B4]

B4   Do you use a mobile phone?
(tick only one)

Yes .... ☐

No .... ☐

[-> go to C1]

Module C: Use of the Internet

C1   When did you most recently use the Internet? *(filter question)*

Within the last 3 months …………………………………	☐	**[-> go to C2]**
Between 3 months and a year ago ……………………	☐	**[-> go to D1]**
More than 1 year ago ……………………………………	☐	**[-> go to D1]**
Never used it ………………………………………..	☐	**[-> go to E1]**

C2   On average how often did you use the Internet in the last 3 months?
(tick one)

Every day or almost every day …………………………… ☐

At least once a week (but not every day) …………… ☐

At least once a month (but not every week) ………… ☐

Less than once a month ………………………………… ☐

[-> go to C3]

C3   Where have you used the Internet in the last 3 months (using a computer or any other means)?
(tick all that apply)

a) At home ……………………………………………… ☐

b) At place of work (other than home) ………………… ☐

c) At place of education ………………………………… ☐

d) At another person's home ………………………… ☐

e) At other places ……………………………………… ☐

of which: (optional)   **e1)** Public Library ……………………………………………… ☐

e2) Post office …………………………………………… ☐

e3) Public office, town hall, government agency …………… ☐

e4) Community or voluntary organisation …………………… ☐

e5) Internet Café ………………………………………… ☐

e6) Hotspot ……………………………………………… ☐

[-> go to C4]

C4   Do you use any of the following mobile devices to access the Internet?
(tick all that apply)

a) Mobile phone via GPRS ……………………………………………… ☐

b) Mobile phone via UMTS (3G) ……………..……………………………… ☐

c) Handheld computer (palmtop, PDA) ………………………………………… ☐

d) Portable computer (laptop) via wireless connection away from home or work …….. ☐

e) None of the above ……….. ……………………………………………… ☐

[-> go to C5]

C5 **For which of the following activities did you use the Internet in the last 3 months for private purpose?**
(tick all that apply)

Communication

    a) Sending / receiving e-mails ………………………………………..………………………… ☐

    b) Telephoning over the Internet / videoconferencing …………………………………….... ☐

    c) Other (use of chat sites, messenger, etc.) ……………………………..………….……… ☐

Information search and on-line services

    d) Finding information about goods or services …………………………………..…..……… ☐

    e) Using services related to travel and accommodation ………………………….………… ☐

    f) Listening to web radios or watching web television ………………………………………… ☐

    g) Playing or downloading games, images, films or music ……………………………..…… ☐

    h) Downloading software ………………………………………………………..……………… ☐

    i) Reading or downloading online newspapers / news magazines ………………………… ☐

       if yes to i)   **i1)** Have you subscribed to news services or products to receive them regularly? Yes ….. ☐
                                                                                                           No …… ☐

    j) Looking for a job or sending a job application …………………………………….…..…… ☐

    k) Seeking health-related information (e.g. injury, disease, nutrition, improving health, etc) …… ☐

    l) Other information search or online service ………………………………………………… ☐

Banking, (optional: selling of goods or services)

    m) Internet Banking ………………………………………..…………………………………… ☐

    n) Selling of goods or services, e.g. via auctions (optional) ………………………………… ☐

Interaction with public services or administrations

    o) Obtaining information from public authorities' web sites ………………………………… ☐

    p) Downloading official forms …………………………………………..……………………… ☐

    q) Sending filled in forms ……………………………………………………………………… ☐

Training and education

    r) Looking for information about education, training or course offers ……………………… ☐

    s) Doing an online course (of any subject) …………………………………..……………… ☐

    t) Consulting the Internet with the purpose of learning …………………………………… ☐

[-> go to C6]

C6 **In the past 3 months, have you taken part in a course (any training, not only computer or Internet training, including school or university)?**

    Yes …………….. ☐   **[-> go to C7]**

    No ……………. ☐   **[-> go to C8]**

C7   In the past 3 months, have you used the Internet to
(tick all that apply)

    a) Do research as part of a training course or your education ....☐

    b) Exchange messages relating to the course content with other learners.....................☐

    c) Download learning content which was provided online ....☐

    d) Look for the availability of a book or article for your course in a library ....☐

[-> go to C8]

C8   Would you like to use the Internet more?

Yes ....☐   **[-> go to C9]**

No ....☐   **[-> go to C10]**

C9   What are your barriers to more intensive use of the Internet?
(tick all that apply)

    a) Foreign language skills are inadequate...☐

    b) Lack of time ....☐

    c) Connection is too slow ....☐

    d) Additional connection or per-volume download cost ....☐

    e) Cost of online content ....☐

    f) Content (what is there is not interesting enough to make more use of the Internet).......☐

    g) Lack of skills or knowledge (e.g. sites are not user-friendly or too complicated)...........☐

    h) Security or privacy concerns ....☐

    i) None of the above ....☐

[-> go to C10]

C10   Through using the Internet, have you had a computer virus resulting in loss of information or time in the last *12* months?

Yes ....☐

No ....☐

[-> go to C11]

C11   How often do you make safety copies or back up files (documents, pictures, etc.) from your computer on e.g. a diskette, a cd or to diskspace on Internet servers?
(tick one)

Always or almost always ....☐

Sometimes ....☐

Never or hardly ever..☐

Not applicable (because I don't keep files on a computer) ....☐

[-> go to D1]

Module D: Use of e-commerce

D1 **When did you last buy or order goods or services for private use over the Internet (excluding manually typed e-mails) ?**

Within the last 3 months .... ☐   **[-> go to D2]**

Between 3 months and a year ago .... ☐   **[-> go to D2]**

More than 1 year ago .... ☐   **[-> go to E1]**

Never bought or ordered .... ☐   **[-> go to E1]**

D2 **What types of goods or services did you buy or order over the Internet for private use in the last 12 months?**
(tick all that apply)

a) Food or groceries .... ☐

b) Household goods (e.g. furniture, toys, etc) .... ☐

c) Films, music .... ☐

d) Books, magazines, newspapers or e-learning material .... ☐

e) Clothes, sports goods .... ☐

f) Computer software and upgrades (incl. computer and video games) .... ☐

g) Computer hardware .... ☐

h) Electronic equipment (incl. cameras) .... ☐

i) Share purchases, financial services or insurances .... ☐

j) Travel or holiday accommodation .... ☐

k) Tickets for events .... ☐

l) Lotteries or betting .... ☐

m) Other .... ☐

*[If 'yes' to categories **c)** , **d)** or **f)** , go to question **D3** ; otherwise go to question **E1**]*

D3 **Were any of the following products that you bought or ordered over the Internet downloaded or accessed from websites rather than delivered by post etc.?**
(if 'yes' to c, d or f in question D2)
(tick all that apply)

Films, music .... ☐

(Electronic) books, magazines, newspapers, *e*-learning material .... ☐

Computer software (incl. computer and video games and software upgrades) .... ☐

[-> go to E1]

Module E: E-skills

E1   When did you last take a training course (of at least 3 hours) on any aspect of computer use?
(for respondents who didn't answer "Never used one" in question B1)

Within the last 3 months .... ☐   **[-> go to E3]**

Between 3 months and a year ago .... ☐   **[-> go to E3]**

Between 1 and 3 years ago .... ☐   **[-> go to E3]**

More than 3 years ago .... ☐   **[-> go to E2]**

Never taken one .... ☐   **[-> go to E2]**

E2   What are the reasons for not having taken a course on computer use recently?
(tick all that apply)

a) No need to take one because my computer skills are sufficient ☐

b) No need to take one because I rarely use computers .... ☐

*[If 'Yes' to category **a)** or **b)** , go to question **E3** ; if 'No', go to category **c)** of question **E2**]*

c) Lack of time .... ☐

d) Course costs .... ☐

e) No suitable offer available .... ☐

f) Courses are too difficult .... ☐

g) None of the above .... ☐

[-> go to E3]

E3   Which of the following computer related activities have you already carried out?
(tick all that apply)
(for respondents who didn't answer "Never used one" in question B1)

a) Copying or moving a file or folder .... ☐

b) Using copy and paste tools to duplicate or move information within a document .... ☐

c) Using basic arithmetic formulas in a spreadsheet .... ☐

d) Compressing (or zipping) files .... ☐

e) Connecting and installing new devices, e.g. a printer or a modem .... ☐

f) Writing a computer program using a specialised programming language .... ☐

g) Connecting computers to a local area network .... ☐

h) Detect and solve computer problems (e.g. computer runs slowly) .... ☐

i) None of the above .... ☐

[-> go to E4]

E4 **Which of the following Internet related activities have you already carried out?**
(tick all that apply)
(for respondents who didn't answer "Never used it" in question C1)

a) Using a search engine to find information .... ☐

b) Sending e-mails with attached files (documents, pictures, etc.) .... ☐

c) Posting messages to chatrooms, newsgroups or an online discussion forum .... ☐

d) Using the Internet to make telephone calls .... ☐

e) Using peer-to-peer file sharing for exchanging movies, music, etc. .... ☐

f) Creating a web page .... ☐

g) Finding, downloading and installing software .... ☐

h) Keeping viruses, spyware and adware off your computer .... ☐

i) None of the above .... ☐

[-> go to E5]

E5 **Where or how did you obtain the skills to carry out these activities?**
(tick all that apply)
(for respondents who ticked at least one option (a to f) in questions E3 or E4)

a) Formal educational institution (school, college, university) .... ☐

b) Training courses in adult education center (but not on the initiative of your employer) .... ☐

c) Vocational training courses (on the demand of the employer) .... ☐

d) Self-study using books, cd-roms, online courses, etc. .... ☐

e) Self-study in the sense of learning-by-doing .... ☐

f) Informal assistance from colleagues, relatives, friends .... ☐

g) Some other way .... ☐

[-> go to E6]

E6 **Do you judge your computer skills to be sufficient if you were
to look for a job or change jobs within a year?**
(tick only one)
(for all respondents, excepting retired persons)

Yes .... ☐

No .... ☐

Not applicable .. ☐

[-> go to F1]

Socio-demographic background characteristics

F1 Age: ☐☐

F2 Sex:
☐ Man
☐ Woman

F3 Educational level:
(tick only one)
☐ Primary or lower secondary education, no formal educat. *[ISCED 0, 1 or 2]*
☐ Upper secondary education *[ISCED 3 or 4]*
☐ Tertiary education *[ISCED 5 or 6]*

F4 Employment situation:
(tick only one)
☐ Employee or self-employed (incl. family workers)
☐ Unemployed
☐ Student (not in the labour force)
☐ Other not in the labour force
(retired, inactive, in compulsory military service, etc.)

F5 Occupation:
< description >
< to be recoded into at least 3-digit ISCO categories >

F6 Geographical location:
☐ "Objective 1" region
☐ Non-"Objective 1" region

F7 Type of locality:
☐ Densely-populated area
☐ Intermediate area
☐ Thinly-populated area

F8 Number of members in the household: ☐

F9 *of which,* number of children under 16: ☐

OPTIONAL QUESTION

F10 Household income: _____ <currency>
(average net monthly income)

II.3.2.   Transmission format

Version 1.1, May 2007.

Note: this transmission format refers to the 2007 survey.

⇨ See document ***MM2007_AnnexII.3.2_TransmissionFormatHH2007.doc***

II.3.3.   Reporting template

Note: the reporting template for interim & final methodological report refers to the 2007 surveys.

⇨ See document ***MM2007_AnnexII.3.3_ReportingTemplateHH2007***

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 **MM2006_AnnexIII.1_Reg808-2004.pdf** (7 pages)

Versions in other languages can be found on the EU's legislation server *europa.eu.int/eur-lex/en*
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 1031/2006 OF 4 JULY 2006 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, 07.07.2006, L186.

⇨ See document ***MM2007_AnnexIII.2_Reg1031-2006.pdf***

Versions in other languages can be found on the EU's legislation server *europa.eu.int/eur-lex/en*
1. click 'Official Journal'
2. select 'Year' (=2006) and 'OJ Number' (=186)
3. select language (upper right, all official languages excepting Maltese)
4. scroll down in the table of contents to the link to page 11.