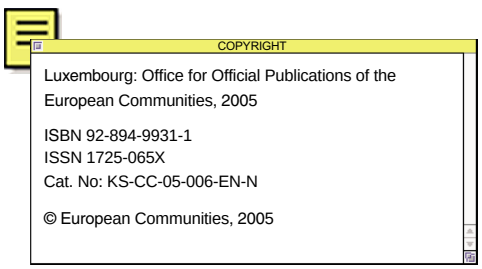
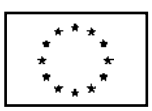
WORKING PAPERS AND STUDIES

2005 EDITION

The continuity of indicators during the transition between ECHP and EU-SILC



ISBN 92-894-9931-1
ISSN 1725-065X
Cat. No: KS-CC-05-006-EN-N



EUROPEAN COMMISSION

THEME
Population and social conditions

The views expressed in this document are the author's and do not necessarily reflect the opinion of the European Commission.

Europe Direct is a service to help you find answers to your questions about the European Union

Freephone number (*)**:**

00 800 6 7 8 9 10 11

(*) Certain mobile telephone operators do not allow access to 00 800 numbers or these calls may be billed.

A great deal of additional information on the European Union is available on the Internet.
It can be accessed through the Europa server (http://europa.eu.int).

Luxembourg: Office for Official Publications of the European Communities, 2005

ISBN 92-894-9931-1
ISSN 1725-065X

Foreword

Executive summary

The European Community Household Panel survey (ECHP) was a pioneering data collection instrument. Launched on a gentleman's agreement basis in 1994 it expired in 2001. However the political scene has changed, notably with the introduction of an open method of coordination in the fields of social inclusion and pensions reform. Other important changes include enlargement of the EU from 15 to 25 member states (and demands for coverage of other neighbouring countries), and the publication by the United Nations expert group on household income statistics of a detailed report and recommendations.

In recognition of these changes, the ECHP is being progressively replaced with data collection under the EU-SILC regulations (no.1177/2003 Community Statistics on Income and Living Conditions). Seven countries launched a preliminary version of EU-SILC in 2003. The project is formally launched in 2004 and EU25 coverage is expected with effect from 2005. The regulations will be fully applicable from 2007. EU-SILC is expected to become the reference source of statistics on income and social exclusion in the European Union. During the transition period until launch of EU-SILC, indicators are being compiled by Eurostat from the best available national sources, harmonised as closely as possible with EU-SILC definitions.

There is an unavoidable disruption in the time series of indicators produced. Data collection under the EU-SILC regulations displays some important differences from its predecessor, the ECHP. Similarly, there are important differences between EU-SILC and transitional national data sources. The impact of these various differences can be significant, depending on the country and the indicators concerned. Although this paper presents various checks and comparisons which have been made - and further information may become available with the receipt of quality reports - it is impossible to isolate individual causes for all such differences and quantify their impact. Nevertheless, notwithstanding the problems of comparability over time and between countries, the validated results for 2002 are considered to give useful information about income poverty and social exclusion, and are suitable for use in the various reporting processes.

Acknowledgements

This report has been developed by the staff of Eurostat Unit D2 'Living conditions and social protection' – Head of Unit: Anne Clemenceau, with contributions from the 'Income poverty and social exclusion statistics' team (Ian Dennis, Head of Project; Anne-Catherine Guio, seconded national expert; Nicolas Marchadour) and the 'EU-SILC' team (Paloma Seoane, Head of Project; Jean-Marc-Museux; Guillaume Osier; Jean Thill; Bruno Bernard; Jonas Johansson). Eurostat would like to thank the national statistics institutes and other national authorities that provided information. A preliminary version of this report was discussed between DG.ESTAT and DG.EMPL in April 2005. The report is presented at the Eurostat Working Group on Living Conditions Statistics in June 2005. The current text was finalised for publication on 1st July 2005.

Contents

1. Background

Statisticians and users alike agree that the European Community Household Panel (ECHP) survey has offered a unique information source with a large range of topics, standardised methodology and procedures and a pure longitudinal panel design.

During the period 1994-2001 the ECHP has traditionally been the primary source of data used by Eurostat for the calculation of many indicators in the field of Income, Poverty & Social Exclusion (such as the Structural Indicators of Social Cohesion; indicators adopted under the Open Method of Coordination such as the 'Laeken' indicators of Social Inclusion and indicators of Pensions Adequacy; Sustainable Development Indicators of poverty and of ageing; and many other indicators published on the Eurostat New Cronos database).

However, and there is no hiding it, the panel has always suffered from operational problems:

- Data timeliness : long delays, despite improvements at the end of the panel (the ECHP 2001 User Database was published Dec.2003)
- Initial response/attrition rates:
  - Initial response rates were very low (eg. LU 41%, DE 48%)
  - Total attrition: 24% over the four initial years, heavy in first year (10%).
- Non participation of Sweden in the project (but inclusion of comparable national source)
- Non-integration of the survey in some National Statistical Systems
- Income definition not fully in accordance with international practice (the UN 'Canberra Guidelines', to which Eurostat contributed, were not published until 2001)

The political scene has also changed, particularly after the Lisbon, Nice, Stockholm, and Laeken summits, where strong support was given to the eradication of poverty and to a better understanding of social exclusion and to the central request for more timeliness. Thus, the timeliness of production needed to be tailored to the evolving political needs and the detailed content (existing set of survey variables) needed to be reviewed.

What was good about the ECHP – its strengths and experience gained from it – are being actively used to develop its successor, data collection under the EU-SILC (Community Statistics on Income and living Conditions) regulations.

For EU-SILC, priority is given to:

- Timeliness

  EU-SILC cross-sectional dimension availability:
  - End of November N+1: Micro-data files for the year N and the social cohesion indicators based on cross-sectional component.
  - End of February N+2: Community micro-data files collected in year N available for scientific purposes.

  EU-SILC longitudinal dimension availability:
  - End of March N+2: micro-data files up to year N.
  - End of July (N+2): Community micro-data files for data collected up to year N available for scientific purposes.

- Flexibility
  - National sample designs
  - Information compiled from interviews or from registers complemented by interviews
  - Encouraged to integrate new source into National Statistical System

- Comparability
  - Common guidelines, definitions and procedures.
  - Maximum possible consistency with Canberra recommendations on the definition of household income

- Full geographical coverage
  - For the first time EU25 member states + candidate countries for accession (BG, HR, RO, TR) + EFTA countries (IS, NO, CH)…
  - By contrast, the ECHP covered 14 countries (+ the inclusion of comparable data from national source for Sweden), although a later pilot project did collect comparable indicators derived from national sources for candidate and accession countries.

2. Main differences between ECHP and EU-SILC

Legal basis

While the ECHP was launched on the basis of a gentlemen's agreement, it was decided to introduce a legal act for EU-SILC[1] and a Framework Regulation of the European Parliament and of the Council was published in July 2003. Technical aspects of the instrument are developed by five Commission Implementation Regulations ('Sampling and tracing rules'; 'Definitions'; 'List of target primary variables'[2] ; 'Fieldwork aspect and imputation procedures'; and 'Intermediate and final quality reports'). Each year, a new Commission Regulation is created on the list of target secondary variables[3] : two have been issued to date ('Intergenerational transmission of poverty'; 'Social participation') and a third is under discussion.

Survey design

The ECHP survey design was based on a pure panel. It means that the sample of people selected for the first year of the survey (sample persons) were followed-up throughout the subsequent duration of the survey (8 years), wherever they may have moved. Children born to sample women were included as sample persons and followed-up.

The cross-sectional and longitudinal data came from the same survey and were collected and processed at the same time.

[1] For a full listing of these regulations, see **appendix VII**.

[2] The core set of subject areas for which data are to be collected on an annual basis.

[3] A module containing subject areas for which data are to be collected at intervals of four years or more.

Since improving timeliness has been one of the new tool's core objectives, and because it is recognised that the longitudinal dimension takes more time in data production, EU-SILC will provide two types of annual data: cross-sectional and longitudinal data that will be treated according to different timetables.

In this way, the cross-sectional and longitudinal data can conceivably come from separate sources, i.e., the longitudinal dataset does not need to be "linkable" with the cross-sectional dataset at the micro-level. Nevertheless, an integrated design 'the rotational design[4]' is recommended by Eurostat for those countries planning to launch a new operation. This design aims to be the most cost effective and efficient for satisfying both the cross-sectional and the longitudinal requirements.

Under this design the minimum panel duration was reduced from 8 years to 4 years (the number of years of observations necessary for building the longitudinal common EU indicators). Consequently, the impact of cumulative attrition will be lower.

The use of a rotational panel allows introduction in the sample of new population sub-groups each year (eg. immigrants), and as a result the cross-sectional data derived from this design will be richer than data derived from a pure panel.

Variables

The number of annual EU-SILC target primary variables is much lower than the number of variables recorded in ECHP (although countries are of course free to include additional variables in their national surveys). Nevertheless, those variables necessary to build the indicators adopted under the open method of coordination have generally been retained (with some exceptions)[5].

Definitions

Although the majority of EU-SILC variables were defined in the same way as the corresponding ECHP variables, the total household gross and disposable income and the different income components were redefined to follow as closely as possible the international recommendations of the UN 'Canberra Manual'.

Income

A key objective of EU-SILC is to deliver robust and comparable data on total disposable household income, total disposable household income before transfers (except old age and survivor's benefits; including old age and survivor's benefits), total gross income and gross income at component level.

This objective will be reached in two steps, insofar as Member States will be allowed to postpone the delivery of gross income at component level and of total household gross income data until after the first year of their operations.

Apart from this, the only data for which delivery will not be compulsory as from the first year of the operation are as follows:

- non-monetary components of employee (with the exception of company cars that is to be calculated as from the first year of the operation) and self-

[4] Rotational design refers to the sample selection based on a number of subsamples or replications, each of them similar in size and design and representative of the whole population. From one year to the next, some replications are retained, while others are dropped and replaced by new replications.

[5] ILO activity status is not an EU-SILC target primary variable. In this way, indicators based on this variable will henceforth be built with the 'current activity status' or 'most frequent activity status'. Furthermore, the distinction between employees and self-employed is no more possible in the calendar of activities, and therefore it may only be possible to compute the working poverty indicator for all people at work, regrouped together.

employed income, imputed rent and interest payments that shall be optional from the first year of the operation and compulsory from 2007;

- gross employers' social insurance contributions shall only be included from 2007 if results of feasibility studies are positive.

As it is mentioned above, in EU-SILC the income at component level is recorded 'gross'. In the ECHP, the income components were recorded net.

New components of disposable income have been introduced in EU-SILC:

- Transfers paid to other households (only transfers received from other households were taken into account in the ECHP);
- Tax adjustment (only taxes paid at source were collected in ECHP)
- Taxes on wealth
- Interest paid on mortgage loans
- Imputed rent
- Non-cash employee income ('income-in-kind')
- Value of goods produced for own consumption ('income-in-kind')
- Employers social insurance contributions[6]
- Furthermore, EU-SILC takes into account negative values of self-employment income, which were previously set to 0 in the ECHP.
- Other variables that can take negative values are variables collected under 'property income'.

The content of some variables has changed:

- The social benefits do not contain the income from 'individual pension plans' (this component was included in theory in the ECHP)
- Survivors' and disability benefits paid after the standard retirement age are included in EU-SILC under 'old-age benefits' (and no more in survivors' and disability functions as in ECHP)
- Early-retirement benefits paid for labour market reasons or in case of reduced capacity to work are included respectively under 'Unemployment benefits' or under 'Disability benefits' (and no more in old age benefits as in the ECHP).
- The total household disposable income[7] includes the new components introduced in the project and tax adjustments are taken into account in the calculation of this variable.

The income reference period is more flexible:

- While in the ECHP the income reference period was the previous year, EU-SILC has fewer constraints. In this way, the income reference period may be a fixed 12-month period (such as the previous calendar year or tax year) or a moving 12-month period (such as the 12 months preceding the interview) or be based on a comparable measure.

[6] As mentioned above, gross employers' social insurance contributions shall only be included from 2007 if results of feasibility studies are positive.

[7] See **Appendix I** 'Computation of household disposable income'

Quality criteria

The following quality criteria are required in EU-SILC:

- To attain *a minimum effective sample size*[8]. This minimum effective sample size is fixed by Regulation for both components. Sample size for the longitudinal component refers, for any pair of consecutive years, to the number of households successfully interviewed in the first year in which all or at least a majority of households members aged 16 or over are successfully interviewed in both years.
- Quality assessment in *a country quality report*. These quality reports will contain information on 'accuracy', 'comparability' and 'data coherence'. On the basis of the individual national reports, a comparative report shall be drawn up by Eurostat.
- Computation of *systematic standard errors* for the income-based indicators as well as from the main survey estimations.
- Methodological studies will be coordinated by Eurostat to identify best practices.

3. EU-SILC implementation

The EU-SILC project was launched in 2003 on the basis of a 'gentleman's agreement', in six Member States (Belgium, Denmark, Greece, Ireland, Luxembourg, and Austria) as well as in Norway.

The starting date for the EU-SILC instrument under the Framework Regulation of the European Parliament and of the Council was 2004 for the EU-15 (with the exception of Germany, Netherlands and the UK who have derogations until 2005) as well as for Estonia, Iceland and Norway.

Germany, Netherlands and UK as well as the new member states (with the exception of Estonia) are allowed to start in 2005 under the condition that they supply comparable data for the year 2004 for the cross-sectional common EU indicators that have been adopted by the Council in the context of the open method of co-ordination before 1 January 2003.

Timetables for implementation in Candidate Countries for EU accession (Bulgaria, Croatia, Romania and Turkey) and in Switzerland are being discussed.

[8] This is the the size required if the survey were based on simple random sampling (design effect in relation to the variable = 1.0). The actual sample sizes will have to be larger to the extent that the design effects exceed 1.0 and to compensate for all kinds of non-response. Furthermore, the sample size refers to the number of valid households which are households for which, and for all members of which, all or nearly all the required information has been obtained

4. Implications of launching EU-SILC at different times

The availability of indicators

The implications of this varying launch timetable include the following:

- Micro data and cross-sectional indicators for 6 EU Member States (all EU15 MS except DE, NL, UK, FR, ES, IT, PT, FI, SV) together with Norway are supposed to be available since December 2004 (gentleman's agreement). Section 6.2.3 details the timetable of the current validation status.
- Micro data and cross-sectional indicators for 12 EU Member States (Estonia together with all EU15 MS except DE, NL, UK) as well as for Norway and Iceland should be available at the latest in December 2005;
- The first set of micro data and cross-sectional indicators from EU-SILC which covers all the EU25 Member States together with Norway and Iceland will only be available in December 2006.

Resolving the problem

Therefore in March 2004 Eurostat launched a collection of indicators derived from national sources **for income reference year 2002** (and if possible 2001 and 2000)[9] for former EU15 countries that did not launch EU-SILC in 2003.

In parallel, the new Member States and the Candidate Countries for EU accession continue to participate in a project coordinated by Eurostat to supply comparable indicators derived from national sources. The third round of figures[10] was published in December 2004.

Table 1 presents the different national data sources used in these countries.

Table 1: source of data to be used during the transition until EU-SILC

Country	Source	Situation at 15/04/2005	
		Survey year	Income year
Czech Republic	Survey on Social Situation of the Household (Sociální Situace Domácností)	2001	2000
	Microcensus	2003	2002
Germany	GSOEP (Sozio-oekonomische Panel)	2001,2002,2003	2000,2001,2002
Estonia	Household Budget Survey (Leibkonna Eelarve Uuring)	2000,2001,2002, 2003	2000,2001,2002, 2003
Spain	Household Budget Survey (Encuesta Continua de Presupuestos Familiares)	2002,2003	2001,2002
France	Tax Survey (Enquête Revenu Fiscaux)	2001,2002	2000,2001
Italy	** N/A **		

[9] The 2000 income reference year data are particularly useful for as these figures will be compared to the last ones available from the ECHP.

[10] The third round sought maximum comparability with EU-SILC methodology. The earlier rounds sought indicators computed as far as possible in accordance with ECHP methodology.

Table 1, continued

Country	Source	Situation at 15/04/2005	
		Survey year	Income year
Cyprus	Household Budget Survey (Family Expenditure Survey)	1997, 2003	1997, ,2003
Latvia	Household Budget Survey (Majsaimniecibu Budzetu Petijums)	2000,2002	2000,2002
Lithuania	Household Budget Survey (Namu ukiu biudzetu tyrimas)	2000,2001,2002	2000,2001,2002
Hungary	Household Budget Survey (Háztartási Költségvetési Felvétel)	2000,2001,2002	2000,2001,2002
Malta	Household Budget Survey (Household Budgetary Survey)	2000	2000
Netherlands	Income Panel Survey (Inkomenspanelonderzoek)	2000,2001,2002	2000,2001,2002
Poland	Household Budget Survey (Badania Budżetów Gospodarstw Domowych)	2000,2001,2002, 2003	2000,2001,2002, 2003
Portugal	Reduced ECHP sample. Only limited indicators are available (at-risk-of-poverty rates before and after transfers at level of total population; s80s20 income quintile share ratio)	2002,2003	2001,2002
Slovenia	Household Budget Survey (Anketa o porabi v gospodinjstvih)	2000,2001,2002	2000,2001,2002
Slovakia	Microcensus	2003	2002
Finland	Income Distribution Survey (Tulonjakotilasto)	2001,2002,2003	2000,2001,2002
Sweden	Survey of Living Conditions (ULF: Undersökning av levnadsförhållanden)	2002	2002
United Kingdom	Household Budget Survey (Family Resources Survey)	2000/01,2001/02 2002/3	2000/01,2001/02 2002/3
Bulgaria	Household Budget Survey	2000,2001,2002	2000,2001,2002
Croatia	Household Budget Survey	2003	2003
Romania	Household Budget Survey	2000,2001,2002	2000,2001,2002
Turkey	Household Income and Consumption Expenditure Survey	2002	2002
Iceland	** N/A **		
Switzerland	** N/A **		

5. Steps to harmonise the different sources as far as possible

The common list of indicators

The indicators to be produced are those which have been adopted under the open method of coordination which applies in the EU in the field of social inclusion; and in the field of pensions adequacy, sustainability and modernity[11].

Steps to harmonise indicators produced using the different national sources

In order to ensure the maximal comparability with indicators produced by countries launching EU-SILC, income definitions applied by countries for the production of indicators using national databases during the transition period should be as close as possible to the ones adopted for EU-SILC.

During the transition period, for EU15 Member-States, the indicators will be based on the definition of disposable income not including the variables which are only mandatory in EU-SILC from 2007[12], namely: Imputed rent; Interest paid on mortgage; Value of goods for own consumption; Non-cash employee income (only the value of company cars for private use is to be included from the beginning); Employer's social insurance contributions;

For countries able to supply such variables already, a second set of indicators including Imputed rent and 'Income-in-kind' (Value of goods for own consumption and Non-cash employee income) were welcomed to perform sensitivity analysis.

Furthermore, for the 10 new Member States and the Candidate Countries for EU accession, income-in-kind is partly included in the total income definition, as it is considered to be a more substantial component of the disposable income for these countries than is the case for EU15 Member States, meaning that its exclusion would have significantly underestimated the actual situation. 'Income-in-kind' covers goods produced directly by the household through either a private or a professional activity (e.g. own production of food by farming households or a household whose leisure activity is connected with agriculture; products from hunting or fishing; withdrawals from stocks by tradespeople, etc.). Services obtained free of charge as part of a professional activity are also classified as 'benefits in kind' (e.g. provision of housing, company vehicle, crèche facilities, free meals at work, etc.).

In order to evaluate the effect of this 'Income-in-kind' component on the value of the indicators, a second set including only income in cash was also supplied. If values for imputed rent were available, a test calculation including this component was also considered.

[11] The current lists are reproduced in **appendix VIII** ("Laeken" indicators of social inclusion) and **appendix IX** (pensions indicators). Detailed methodological guidance is available on the Eurostat website. Details of the political context are available on the DG.Employment website.

[12] In accordance with a bilateral agreement between the Danish authorities and the Commission, Denmark will until 2007 send two series of indicators, with and without non-monetary components (i.e. imputed rent).

For the new Member States and the Candidate Countries for EU accession, the indicators are mainly produced using information from Household budget surveys. In order to approximate as closely as possible to the EU income definition, components such as the following were excluded from the total household income: lottery winnings, insurance claim receipts, non-regular gifts (although regular transfers received from other households were included), all transfers paid to other households, revenue from sales of property (for example houses or cars). The impact of these adjustments on reported values can sometimes be significant by comparison with the national income definitions used in these countries.

It must be kept in mind that the difficulty of capturing income from the hidden economy can introduce significant bias in the income distribution measured through surveys.

Remaining comparability problems

National surveys can have different income reference periods (e.g. monthly vs. yearly, last 12 months vs. previous calendar year, etc.), which may have an impact *inter alia* on the value of the data and their comparability between countries. Furthermore, within a country, the income variable may not be fully comparable between sub-samples if the data collection is conducted at different periods of the year (i.e. in continuous surveys for which the income reference period is less than one year). In this case, the income distribution (and the results in terms of poverty risk) can be biased by the variability of seasonal income components (such as income from agriculture, self-employment, thirteenth and fourteenth month payment).

Another factor that can affect the comparability of the results is the fact that, although a common reference year is specified and applied for most of the countries, there are some exceptions due to the periodic nature of the data source in the countries concerned (see Table 1). For example, data for Czech Republic, Cyprus, Malta and Slovakia is drawn from periodic sources.

The income reference period is the same as the survey year for the national data sources in Estonia, Cyprus, Latvia, Lithuania, Hungary, Malta, Netherlands, Poland, Slovenia and Sweden.

The EU25, EU15 and EU10 averages are calculated as population-weighted averages of available national values, where country coverage represents 75% or more of the total population for the aggregate under consideration.

Other problems for building the indicators

Due to the missing longitudinal dimension in the underlying national data sources (or for the first few year after launch of EU-SILC[13]), "persistent risk-of-poverty" rates could not be calculated.

For some statistical breakdowns (cross-tabulations) in certain countries, the exact definition used by Eurostat could not be reproduced on the basis of national data sources, due to a lack of required information.

For example, for the computation of certain indicators using EU-SILC, information on self-declared activity status is collected monthly during the income reference period (in order inter alia to be able to link adequately poverty risk and activity status). This information is used to construct the variable "most frequent activity status", which is

[13] This indicator will be produced by the first time for EU-25 in 2010

the status that individuals declare to have occupied for more than half the total number of months for which information on any status in the calendar of activities is available. By contrast in most of the national databases used, this 'calendar' information was not available, and the activity status the closest to the income reference period was used (either self-declared 'main activity' or ILO). For similar reasons, the "risk-of-poverty rate by work intensity" could not be calculated by many countries.

The definition of a dependent child is another example of possible differences between the indicators calculated using national databases. According to the definition used at EU level, 'dependent children' comprise two sub-groups: i) all persons below the age of 16; ii) persons aged 16-24 who are living in a household with at least one of their parents, and who are economically inactive. In some countries, not all the activity criteria were available or the age limit was different. Here also, the best proxy was used.

6. Validation of results

6.1. Non EU-SILC Member States

The data presenting the values of the indicators adopted under the open method of coordination which are based on ECHP (for former EU15 Member states) and on national data bases (for others) are available on the Eurostat 'free data' website[14].

For the EU15 countries which did not launch EU-SILC in 2003, **table 2** compares the main common indicators based on ECHP and on transitional data for a common year (or the closest year available). The extent of the break in the series can be important depending on the country and the indicators[15]. Eurostat, in collaboration with the National Statistical Institutes, tried to identify the main methodological differences between ECHP and the national data in order to document this break. The main methodological points are presented in **Box 1** below, even though it must be recognized that fully explaining and documenting the break is far from easy without comparing in details the ECHP and national micro databases, due to the multiple potential methodological reasons that can explain it.

Table 2: Comparison of main indicators, Non EU-SILC Member States

	DE		ES		FR		NL		UK		PT		FI		SE	
	TR	EC	TR	EC	TR	EC	TR	EC	TR	EC	TR	EC	TR	EC	TR	EC
Survey year	2001	2001	2002	2001	2001	2001	2000	2001	2001	2001	2002	2001	2001	2001	2002	2001
Poverty rate	13	11	19	19	13	15	11	11	19	17	20	20	11	11	11	9
S80/S20	4	3.6	5.1	5.5	3.9	4	4.1	3.8	5.4	4.9	7.3	6.5	3.7	3.5	3.3	3.4
before transferts	22	21	22	23	26	24	22	21	26	29	26	24	29	19	29	17
before transferts and pensions	34	39	40	37	44	40	35	36	33	40	:	37	40	30	45	34
Threshold 1 person PPS	8918	9492	6676	6527	7798	8765	8655	8309	8634	8984	:	4967	7301	7680	8314	8502

TR: transition
EC: ECHP, except for SE (HEK)

[14] For access details see **Appendix II** (the Laeken indicators of social inclusion) and **Appendix III** (the Pensions indicators).

[15] Differences are also apparent between particular indicator breakdowns.

BOX 1: Specific national issues

Germany

Source: GSOEP (Sozio-oekonomische Panel)

Sample size: 10000 households

Compared to indicators previously computed on the basis on ECHP (SOEP data source) for years up to and including the 2001 survey year, two main methodological changes occurred during the transition:

- the sample size is doubled compared to the ECHP as all sample supplements were excluded from the ECHP Users' database given the longitudinal nature of this project;

- Imputation and weighting methods used differ.

Spain

Source: Household budget survey, specific income module.

This is a continuous survey about household expenditure using a rotating panel, renewing the sample by 1/8 every quarter. Normally income information is collected in an aggregated amount (only information about the total monthly net household income is asked). In the transition period between ECHP and EU-SILC, an income module has been added in the last quarter of the year.

Sample size: 8,000 households per quarter.

No data is provided for earlier years as the detailed income module was added in the 2002 survey for the first time;

The income definition includes lump sum receipts, and income received from (and contributions paid to) private pension plans;

The most frequent activity status is based on the situation in the last quarter of the reference year, instead of the income reference period (whole year). It is defined as the status that individuals declare themselves to have occupied for more than half the total number of months for which information on any status in the calendar of activities is available. Consequently, where an individual provides information on his activity status over the last 3 months of the calendar, his most frequent activity status will be the status he declares to have occupied for at least 2 months.

No distinction can be made between wage and salary employees and self-employed workers.

Similarly, the definition of “work intensity of the household” has been adapted to the last quarter. The work intensity of the household refers to the number of months that all working age household members have been working during the last quarter as a proportion of the total number of months that could theoretically be worked within the household.

France

Source: Enquête revenus fiscaux (Fiscal Income Survey)

Sample size: 70,000 households

The survey combines labour force survey information with administrative data on fiscal income.

The sample excludes households headed by a person enrolled in the military. Students’ households are dropped from the sample due to technical difficulties to include them. On the basis of a simulation using 2001 figures, the impact on the overall at-risk of poverty rate is

limited (the total rate would be 0.5 higher if students' households were included). The at-risk of poverty rate for people aged 16-24 decreases from 22.4% to 17% due to this exclusion;

Income components not subject to taxation at the personal level (as some capital incomes, for example) are not included in the income definition used. However, social transfers not collected in the fiscal declaration are imputed. Daily sickness benefits and early retirements are included in salaries and wages;

Salaries can include non cash-employee income (company car or dwelling);

Households with negative fiscal income were dropped;

The calendar of activity refers to a twelve month period starting in April of the year preceding the calendar year used as the income reference period.

Italy

No data provided during transition.

Netherlands

Source: Income Panel Survey, based on fiscal administrative data. This is the same source as will be adapted for EU-SILC.

Sample size: 80,000 households.

Income received in (and contributions to) private pension plans are taken into account in an alternative income definition. Similarly, imputed rent is included in a third alternative definition.

Child alimony, parental contributions for students living away from home, regular cash transfers between households are not included in the survey, and it was not possible to impute these variables. In particular, this can bias the situation of students' households. These variables will be included in EU-SILC.

To construct the most frequent activity status or the work intensity of the household, no calendar of activities was available. These variables were therefore defined on the basis of the main source of income.

Portugal

Source: reduced ECHP (in terms of sample and questionnaire)

Sample size: 1,200 households.

Only a partial set of indicators, on a non-strictly comparable methodology, could be computed on this basis (only at-risk-of-poverty rates before and after transfers (total population) and S80/S20 are available).

Finland

Source: Income Distribution Survey. This is the same source as will be used in EU-SILC

Sample size: 10,000+ households.

Persistent risk-of-poverty rates can be estimated from an alternative national source.

Sweden

Source: The Swedish Survey of Living Conditions (ULF). This is the same source as will be used in EU-SILC for the cross-sectional component.

Sample size: 7,900 households.

Compared to previous data used in ECHP, the upper age limit has been dropped with effect from 2002. This increases the risk of poverty for persons aged 65+, as 33 per cent of the added age group 85+ are at risk of poverty (compared to 12% for the age group 65-84).

No data is provided for earlier years as the methodological change was only implemented with effect from 2002. Similarly, the persistent risk-of-poverty rate (60%, 50%) and risk-of-poverty rate anchored at a moment in time indicators could not be calculated.

To construct the most frequent activity status or the work intensity of the household, no calendar of activities was available. The most frequent activity status is based on the current status. No breakdown is provided of the risk of poverty by work intensity of the household or the distribution of the population by work intensity.

United Kingdom

Source: Family Resources Survey

Sample size: approximately 30,000 households

A flagship national annual survey conducted annually by ONS on behalf of DWP involving face-to-face interviews, which collects information about the incomes and living circumstances of private households in the United Kingdom.

Principal differences from the national source previously adapted for ECHP and/or the prescribed methodology:

- The FRS sample size is approximately 3 times larger than that of the BHPS.
- Estimates for UK are for Great Britain only. The data source has recently been extended to cover Northern Ireland.
- In common with similar surveys, the FRS is considered to collect less robust information for households at the lower end of the income distribution. UK statisticians emphasise that reported low incomes do not necessarily mean low living standards.
- No breakdown of indicators by work intensity of the household is provided.
- Negative self-employed income imputed to zero.
- No longitudinal component.

6.2 EU-SILC data

The data comparing the values of the indicators based on SILC data with those from ECHP are available on the Eurostat 'free data' website[16].

6.2.1. Impact of the changes in the income definition

It has to be kept in mind that compared to the ECHP income definition, EU-SILC 2003 takes into account the following income components:

- Transfers paid to other households
- Tax adjustment/tax on wealth
- Furthermore, EU-SILC takes into account negative values of self-employment income, which were previously set to 0 in the ECHP.

The inclusion of the above mentioned variables can also have an impact on the indicators. The impact of each of these changes is measured and can be commented as follows[17]:

Impact of allowing negative self-employment income:

If we constrain negative self-employment income to 0 (as in the ECHP), the impact on the indicators is very limited (if any). It is mainly the S80/S20, the Gini coefficient and the poverty gap that decreases slightly. The mean equivalised income increases.

Impact of transfers paid to other household:

If we put these transfers at 0, the impact is very limited and touches mainly groups that usually make this type of transfers : men, people aged 25-64 years, singles (their poverty rate diminish) and the total inequality decreases very slightly.

Impact of tax adjustment:

If we put the tax adjustment at 0, the impact depends on the relative share of negative versus positive adjustments and the relative impact on the different groups depends on whether there are beneficiary or payer of such taxes adjustment. It is particularly difficult to generalise the impact of such adjustment on all countries as it depends on the redistributive impact of the tax system and also on the reference period of the adjustment (some countries collect tax adjustment paid/received during the income reference year, other countries collect tax adjustment linked to the income collected).

Note also that, as mentioned in **section 2** of this document, after the legal/standard age of retirement, some of the social transfers are considered as old age benefits in EU-SILC; individual pension plans are not included in the EU-SILC disposable income definition; the treatment of early retirement pension is different in both surveys.

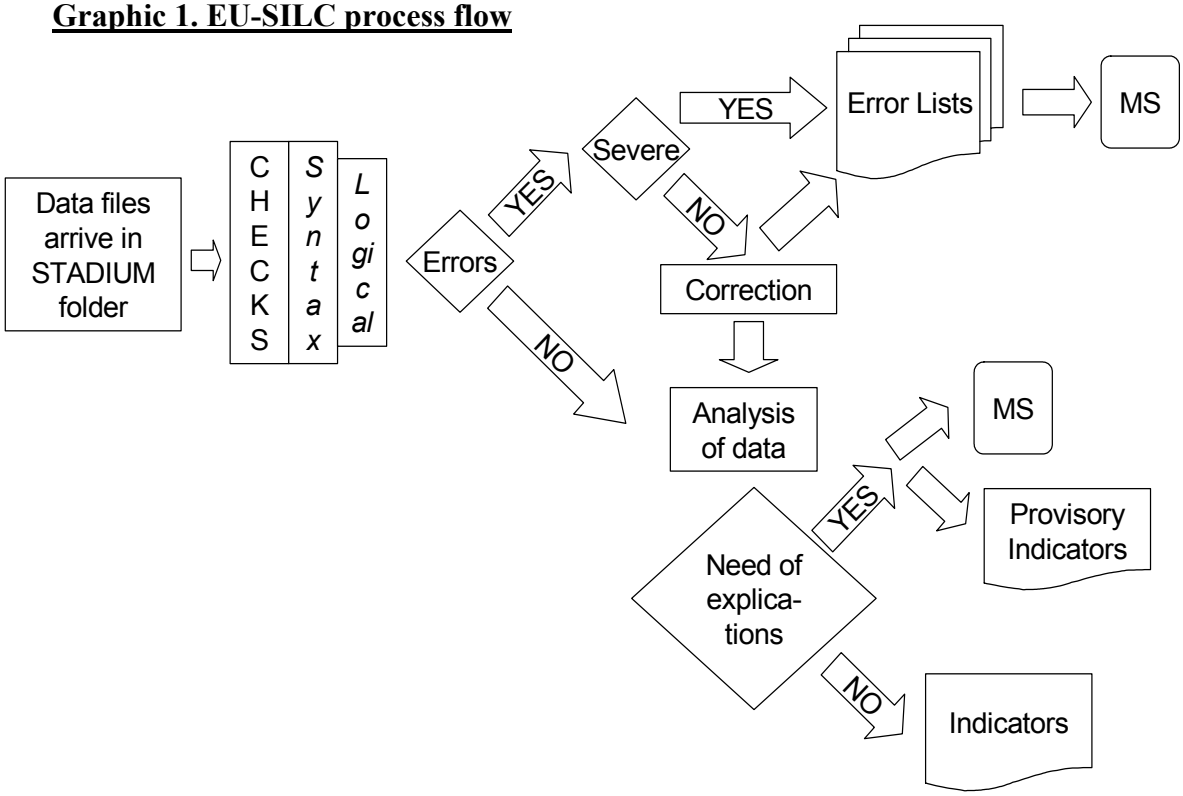
[16] For access details see **Appendix II** (Laeken indicators of social inclusion) and **Appendix III** (Pensions indicators).

[17] Some additional discussion can be found under **section 6.2.2.5**

6.2.2 Validation of EU-SILC data: From the receipt of the data to the production of indicators

Graph 1 shows the EU-SILC process from the receipt of the data to the production of indicators. A two-step process is applied to the EU-SILC data validation: "process checks" and "data analysis".

Graphic 1. EU-SILC process flow



Within the **"process checks**" stage, a "syntax program" checks the values, the flags, the coherence between both and compares the value against a former variable value if a relation exists. The "logical" program checks the structural coherence of the 4 files (Household register, personal register, household data, personal data), the plausibility of the relationship between members of the same household, income (total against components, variables against labour variables) and, in general, coherency between different variables of the same record and other records. In addition, the consistency of the weighting variables and the weighting procedure (design weight, non response correction, calibration) is checked manually on the basis of a standard procedure.

If **severe errors** are detected, error lists (commented where necessary) are sent back to the MS with a request to supply new data files. If only a **few errors** remain in the files, data may be corrected in EUROSTAT or lists sent to the MS.

Within the **"data analysis"** stage, the following issues are studied for each country:

- Response rates/ analysis of frequencies/ fieldwork duration
- An analysis of abnormally low (negative, zero, positive) incomes
- Inconsistencies between gross and net income
- The tracking of anomalies by editing of the "extremes" defined by expert judgments or administrative information
- The influence of imputation on the income distribution
- Detection and editing of outliers obtained from robust regression for major income components

If the analysis reveals anomalies[18], further explanations are requested from the MS, and if necessary new data files will be produced.

Furthermore, a macrodata analysis (to test the coherence with similar information produced from other data sources) is done.

Finally, indicators are calculated and SILC-2003 results are compared with the results based on ECHP (1995-2001) to highlight major differences/similarities. Some additional tests are done to check the impact on the indicators of some methodological changes between the two sources (e.g. introduction of a sample of international civil servants in Luxembourg; evolution of mean income by components; evolution of the distribution of the sample by household and personal characteristics; impact of change of income definition).

Current validation status

Final data are now available for all participant countries. The bilateral validation process typically involved several rounds of data transmission and review before the final acceptance of the microdata. Across the 7 participant countries in the 2003 exercise, the average number of data transmissions required was 3.5, and the average overall delay between first transmission (or deadline date for first transmission, if earlier) and the final acceptance date was 3.5 months. This was a contributory factor to the late delivery of results for the social inclusion political reporting process.

6.2.2.1 Process checks

Syntax/logic

Updated programs related to the checking process are available under CIRCA. Eurostat encourages countries to use these programs and correct errors detected by them before the first transmission of data to Eurostat. Once the errors are corrected by the country it would be desirable that explanations are sent to Eurostat about the edits that failed when the data are right, together with the datasets.

[18] **Appendix V** shows the tables that have been programmed for detecting frequencies, anomalies and average income. These tables also permit to study the response rates, some fieldwork aspects as well as distribution of population and weights.

Special attention should be paid to the construction of the disposable household income, this variable has presented a lot of mistakes in the delivered files for SILC 2003.

Consistency of weighting variables and adequacy of weighting procedures

As mentioned above, design weight, non-response correction, calibration is checked manually on the basis of a standard procedure.

6.2.2.2 Data analysis

Non-response rates and sample sizes

Table 3 shows the household non-response rates, the overall individual non-response rates and individual non-response rates for each country. The response rates are calculated according to common formulae[19].

Table 3. Non-response rates (%)

Country	Household non-response	Overall individual non-response rate	Individual non-response rate
BE	45.54	46.13	1.08
DK	34.09	34.09	0.00
GR	16.92	17.30	0.25
IE	48.02	48.02	0.00
LU	62.42	62.42	0.00
AT	35.63	36.17	0.84
NO	29.20	30.70	1.22

[19] The household non-response rates (Nrh) is computed as follows:

Nrh = (1- (Ra*Rh))*100
where,
Ra= *N° of addresses successfully contacted/N° of valid addresses selected*
Rh = *N° of household interviews completed/ N° of eligible households at contacted addresses*

The Individual non-response rates (NRp) will be computed as follows:

NRp=(1-(Rp))*100
where,
Rp = *N° of personal interviews completed/N° of eligible individuals in completed households*

Overall individual non-response rates (*NRp) will be computed as follows:

*Nrp= (1-(Ra*Rh*Rp))*100

The household non-response rates are ranging from a low bound of 16.92% in Greece, to a high bound of 62.42% in Luxembourg. Belgium presents a rate of 45.54% while for the rest of the countries, the rate ranges from 29 to 36%.

The individual non-response rates are very low for all countries. The rate ranges from 0 to 1.22%.

Table 4 shows the effective sample sizes specified in the Regulation and the sample size achieved in the survey.

The reference in the Regulation is to the *effective sample size* which is the size required if the survey were based on simple random sampling (design effect in relation to the "risk-of-poverty rate" variable = 1.0). The actual sample sizes will have to be larger to the extent that the design effect exceeds 1.0 and to compensate for non-response of all kinds. Furthermore, the sample size refers to the number of valid households which are households for which, and for all members of which, all or nearly all the required information has been obtained.

From the analysis of this table and under the hypothesis that the design effect is equal to 1, the sample size achieved in Denmark and in Ireland are smaller than required. In the case of Ireland, it was agreed in the contract.

Table 4. Sample sizes

Country	Specified in Regulation		Achieved in survey	
	Households	People 16+	Households	People 16+
BE	4 750	8 750	6 199	11 873
DK	5 500	9 500	5 412	10 917
GR	4 750	10 000	6 665	14 923
IE	3 750	8 000	3 112	6 202
LU	3 250	6 500	3 530	7 675
AT	4 500	8 750	4 623	9 543
NO	4 750	8 000	5 852	11 709

6.2.2.3 Frequencies of income components

The frequencies of missing values for the different income variables have been studied.

It has been noted that, as expected[20], the number of missing values after imputation is small for the income variables (with the exception of variable PY200G (Gross monthly income for employees).

[20] "micro-data files transmitted to Eurostat will be weighted, fully checked, edited and imputed in relation to income"

Regarding the number of households receiving income components, we can highlight the case of variable HY090 (interest, dividends, etc.) – see **table 5**. The percentage of households receiving this kind of income is between 2.7% and 26.4% in all countries with the exception of Denmark and Norway.

Table 5. Percentage of households that received interest, dividends,…

	Households with a non-zero value for variable HY090	
	Number of households	Percentage
BE	1 013	16.3
DK	5 342	98.7
GR	183	2.7
IE	715	23.0
LU	607	17.2
AT	1 221	26.4
NO	5 831	99.6

6.2.2.4 Fieldwork

Interview duration

According to According to Article 15 of the Framework Regulation of the European Parliament and of the Council, the total duration of the interview relating to the target primary and target secondary variables of the cross-sectional component, including household and individual interviews, shall not exceed one hour per household on average in each country.

The mean total interview duration per household is calculated as the sum of duration of the household interview plus the sum of duration of all personal interviews, divided by the number of complete households.

Table 6 shows the maximum, minimum and the mean interview duration per type of questionnaire as well as for all questionnaires.

Table 6. Duration of interview

	Household questionnaire			Personal questionnaire			Total interview duration		
	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max
BE	8	20	85	8	20	85	16	58	335
DK	1	1	1	.	.	.	.	.	.
GR	10	20	60	10	20	60	20	65	360
IE	2	14	81	1	14	90	6	42	212
LU	4	19	90	1	16	90	10	54	229
AT	5	10	42	6	11	90	12	33	139
NO	1	8	90	0	6	88	3	19	164

In Belgium and Greece, at household level the interview lasted 20 minutes on average. The average personal interview took 20 minutes. Aggregating the lengths of the personal interviews per household and add them to the length of the household interview, the total interview time (on average) for the household was 58 minutes for Belgium (below the limit value of 60 minutes laid down in the draft EU-SILC Regulation) and it was 65 for Greece. The maximum total interview duration is much higher in these countries compared to other Member States.

It should be kept in mind that both gross and net information on income components were collected in the Belgium survey and approximately 50 non-EUSILC variables were also collected, which has inflated the interview duration. On the other hand, as the EU-SILC secondary target variables are not taken into account, the total interview duration could still exceed the Framework Regulation limit.

The total interview time (in average) for Greece is 65 minutes (only EU-SILC primary target variables). As in Belgium, in Greece both gross and net income components were collected. On the other hand, it should be taken into account that the average number of people aged 16 and over to be interviewed per household is very high compared to other MS (between 2.24 for GR and 1.91 for BE).

The case of Luxembourg, with 54 minutes on average, could be due to introduction, in the questionnaire, of variables not required in EU-SILC at community level despite instructions have been given to the interviewers to take in account only the time they use for the EU-SILC questions. Nevertheless, the introduction of new variables in the questionnaire can increase the household non-response rate (the burden for the household respondent increases) and the item non-response.

No information is available for Denmark. Nevertheless, it can be expected that the duration of the interview be similar to the duration of the interview for Norway as both countries compile income from registers partially and use phone interview.

Type of interview

Table 7.A shows the distribution of people by type of interview. Also of interest is the proportion of interviews which were obtained by proxy. According to Annex I of the Framework Regulation, the mode of collection relating to income components (except for countries relying on income registers), as well as information to be collected for at least one household member aged 16 and over (in all countries)

should be a personal interview, proxy being an exception. (By contrast, proxy interview can be accepted as a normal procedure for other individual information of less detailed and personal kind.)

In the case of Luxembourg and Ireland the number of proxy interviews is very high. This means that of the personal interviews successfully completed, only 71% were obtained through interview with the person concerned.

Denmark and Norway (see **Table 7.B**) used CATI interview for the majority of selected respondents. This kind of interview, together with face to face interview for selected respondents without telephone, is allowed.

Table 7.A Type of interview

	Face to face interview: **PAPI**	Face to face interview: **CAPI**	Telephone interview: **CATI**	**Self-administered** by respondent	**Proxy** interview
BE	----	10 707	----	----	1 166
DK	----	----	4 970	786	5 161
GR	6 169	8 109	144	86	415
IE	----	4 492	----	----	1 710
LU	5 459	----	----	----	2 216
AT	----	8 170	----	----	1 373
NO	----	57	7 154	----	4 350

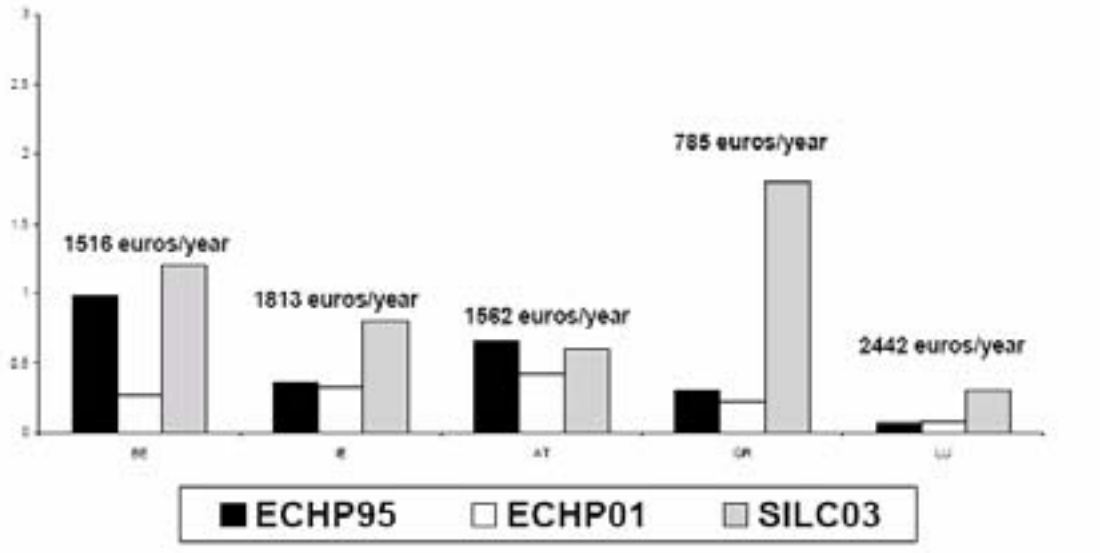
Table 7.B Type of interview (only selected respondent)

	Face to face interview: **PAPI**	Face to face interview: **CAPI**	Telephone interview: **CATI**	**Self-administered** by respondent	**Proxy** interview
DK	----	----	4 626	442	344
NO	----	41	5 805	----	----

6.2.2.5 Extreme values (low, negative, positive and zero income)

Within the "data analysis" stage, an analysis of abnormally low (negative, zero, positive) incomes was made. This section provides an overview of the studies made:

Graph 2: % of people with less than 10% of median income



Note: this graph was prepared using the first data files transmitted to Eurostat. During the data validation process adjustments were made and the values have changed considerably for the final datasets.

Low income is defined as 10% of median equivalised income. The graphic above shows for four countries the percentage of persons with less than 10% of median equivalised income. Greece presents the lower threshold (785 euros/year) and a 2% of people with an equivalised income lower than the threshold.

It is necessary to keep in mind that as the household composition can have changed between the income reference period and the date of the interview, some income components can take negative values and people can work for a salary in kind.

Nevertheless, a careful analysis of the predominant conditions of low income can be very useful to assess the data quality.

The analysis of the total distribution (over the total income) of some income components allows the detection of anomalies (values very far from the set of data) and was applied in EU-SILC data analysis (see outlier detections).

The difference recorded between ECHP and EU-SILC can be the result of the new EU-SILC income definition. The EU-SILC definition on total disposable household income takes into account the following income components:

- Transfers paid to other households
- Tax adjustment/tax on wealth
- Negative values of self-employment income

This may (1°) produce negative household disposable income and (2°) reduce household disposable income. **Table 8** below shows the percentage/number of households for which disposable income is negative due to the tax adjustments, to the self-employed income (PY050) and to the transfers paid to other households (HY130).

While in the case of Greece the household disposable negative income is explained by 'negative self-employed income' (43.5 % of cases), by the 'transfers paid to other households' (26.1 %) and by the 'tax adjustments' (19.6 %), in Luxembourg the negative income is produced by 'self-employed negative income' (11.1%) and by the tax adjustments (88.9%).

In principle, tax adjustments should not produce "negative income" (the amount of taxes paid should not be higher than the amount of income received). Nevertheless, due to the fact that the Commission Regulation on updated definitions permit to collect tax adjustments paid during the income reference period, that normally correspond to taxes paid for the income received in previous years, this taxes can produce negative values in the computation of disposable income. The frequency of occurrence of this pattern is nevertheless an indicator of quality of the data sets. Too frequent negative income may indicate data collection or processing errors.

Table 8. Total household disposable income

	No income		Income >0		Income < 0		Total (1)	Why negative income? (2)			
	No.	%	No.	%	No.	%	%	PY050 (3)	HY130 (4)	Taxes	Total
BE	9	0.1	6 165	99.5	25	0.4	100	33.5	-	66.5	100
DK	1	0.0	5 392	99.6	19	0.4	100	47.4	5.2	21.0	73.6
GR	11	0.2	6 621	99.3	33	0.5	100	60.0	-	40.0	100
IE	17	0.5	3 090	99.3	5	0.2	-	60.0	40.0	-	100
LU	2	0.1	3 519	99.7	9	0.3	100	11.1	-	88.9	100
AT	2	0.0	4 621	100	-	-	100	-	-	-	-
NO	1	0.0	5 847	99.9	4	0.1	100	25.0	25.0	50.0	100

(1) Reference population: Total number of households; (2) reference population: total number of households with negative income; (3) self-employment income; (4) transfers paid to other households.

Table 9 shows that the number of observations with negative values is bigger for the 'total household disposable income before transfers' than for the 'total household disposable income'. This situation is due to the definition of 'total disposable household income before transfers' that removes from the total household disposable income the social transfers received without taking care about the impact of tax adjustments or of transfers to other households.

Table 9. Total household disposable income before transfers

	No income		Income >0		Income < 0		Total (1)	Why negative income? (2)			
	No.	%	No.	%	No.	%	%	PY050 (3)	HY130 (4)	Taxes	Total
BE	729	11.8	4 694	75.7	776	12.5	100	0.7	11.8	85.1	97.6
DK	5	0.1	4 487	82.9	920	17.0	100	0.9	0.2	96.0	97.1
GR	964	14.5	5 349	80.3	352	5.3	100	3	35.7	54.1	100
IE	850	27.3	2 242	72.0	20	0.6	100	6.7	93.3	-	100
LU	13	0.4	3 148	89.2	369	10.5	100	1.4	6.7	78.1	100
AT	824	17.8	3 709	80.2	90	1.9	100	-	45.6	54.4	100
NO	6	0.1	5 708	97.5	130	2.2	100	-	29.1	60.4	89.5

(1) Reference population: Total number of households; (2) reference population: total number of households with negative income; (3) self-employment income; (4) transfers paid to other households.

Although, as mentioned above, a negative disposable income is possible, if a country presented a high percentage of negative values or high amounts for them, the values should be studied.

6.2.2.6 Inconsistencies between gross and net income

Four countries (BE, IE, LU, AT) have provided gross and net income at component level. For these countries an average rate on taxes and social contributions paid at source has been calculated. Eurostat cannot evaluate the quality of data on the basis of these rates. It can just naïvely spot strange values according to the common sense and based on country comparisons. Full assessment can only be made by countries, which have a better knowledge of their fiscal systems.

Table 10. Income as provided and % of taxes and contributions paid at source by income component

	Income as provided			Employee income		Self-employment	
	GROSS	NET	GROSS /NET	Average (gross)	Taxes (%)	Average (gross)	Taxes (%)
BE			x	27 365	34	22 617	45
DK	X			29 701		8 584	
GR		x[21]		----			
IE			x	25 692	22	25 655	20

[21] Although in Greece, according to the questionnaire, both gross and net income components were collected the Greece data files contain only net income.

	Income as provided			Employee income		Self-employment	
	GROSS	NET	GROSS /NET	Average (gross)	Taxes (%)	Average (gross)	Taxes (%)
LU			x	39 106	18	----	
AT			x	23 014	30	21 131	45
NO	X			236 189		205 205	

Within the "data analysis" stage, an analysis of abnormally low (negative, zero, positive) incomes was made. This section provides an overview of the studies made:

Table 11. Percentage of taxes and social contributions paid by low income households, households in first, fifth and ninth deciles

	BELGIUM		IRELAND		LUXEMBOURG		AUSTRIA	
	Amount (upper limit)	Tax (%)	Amount	Tax (%)	Amount	Tax (%)	Amount	Tax (%)
Low income	1 516	232	----	----	----	----	----	----
1st Decile	7 710	34	8 276	3	14 545	17	8 544	15
Median	15 162	22	17 981	10	24 416	18	15 621	21
9th Decile	27 195	28	34 415	21	45 143	23	28 055	28

Note: this analysis was made using the first data files transmitted to Eurostat. During the data validation process adjustments were made and the values have changed considerably for the final datasets. In consequence, final results for Belgium are more coherent.

With the exception of Belgium, it is the households belonging to the highest decile that paid in 2003 a higher percentage of taxes.

In Belgium, households with a very low income paid 232 % as taxes. Manual data editing showed the existence of a mistake in the data available when performing analysis.

This manual editing was done also for other countries and problems with the taxes were found for other MS.

6.2.2.7 The influence of imputation on the income distribution

The percentage of imputed values as well as the impact of imputation on the income distribution was studied.

Table 12. Imputation of PY010 (employee income)

		Completely imputed		Partially imputed		Not imputed	
		Total	%	Total	%	Total	%
BE	PY010G	1 133	19.6	1 357	23.5	3 295	57.0
DK	PY010G	----	----	----	----	8 111	100.0
GR	PY010N	----	----	----	----	4 583	100.0
IE	PY010G	243	8.5	80	2.8	2 533	88.7
LU	PY010G	1 103	29.1	677	17.4	2 007	53.2
AT	PY010G	2 399	50.5	310	6.5	2 043	43.0
NO	PY010G	----	----	----	----	9 173	100.0

Table 12 shows the percentage of imputation for the variable PY010 (employee income). This table should be interpreted with some caution due to the fact that it is not easy to distinguish in some cases the percentage of imputation due to the passage from net to gross, from the percentage of imputation due to the estimation of missing values (the distinction between partial or total imputation is also difficult). However, this rough analysis could be useful for detecting an excessive imputation rate.

With the same constraints mentioned above, a study of the influence of the imputation on the income distribution was made. The study was based on a comparison between the distributions of observed income components (i.e. without imputation) and totally imputed.

6.2.2.8 The tracking of anomalies

Editing "extremes" defined by expert judgments or administrative information

When available, external information from administrative sources was used in the detection of anomalies.

Maximum and minimum thresholds can be established for social benefits (particularly, with retirement, survivors' benefits).

MISSOC (http://europa.eu.int/comm/employment_social/social_protection/missoc_tables_en.htm) is a reference source for establishing these thresholds.

In general, MS are in better position to define threshold because they have a much better knowledge of their social and fiscal system

Editing outliers identified by robust regression

An *outlier* is an observation that lies an abnormal distance from other values. The problem is to take a decision about what should be considered abnormal. Before abnormal observations can be singled out, it is necessary to characterize normal observations.

Two activities are essential for characterizing a set of data:

1. Examination of the overall shape of the graphed data for important features, including symmetry and departures from assumptions.
2. Examination of the data for unusual observations that are far away from the bulk of data (often referred to as outliers).

Graphical representation such as box plots, as well as other techniques (robust regression methods) could be used to identify outliers.

Outliers should be carefully investigated. They often contain valuable information about the process under investigation or the data gathering and recording process. Before considering the possible elimination of these points from the data, one should try to understand why they appeared and whether it is likely similar values will continue to appear.

In order to get insight into the data received, Eurostat has developed programs of detection of outliers for major income components using robust regressions. Eurostat developed regression models to explain the main income variables (Personal Employee Cash Income (Gross), Personal Unemployment Benefits (Gross), Personal Old Age Benefit (Gross), Total disposable household income) on the basis of explanatory variables available in the files. The objective is not to find the best model to explain the dependent variable but mainly to disentangle the cloud of points into various blocks defined by the modalities of the variables of the model very much along the way analysis of covariance is done. Anomalies are then traced at the level of the subgroups. Measure of influence of point such as the Cook distance and standardized residual can be used to rank observations according to their deviation to a predicted value. In practice, residuals have been found to be more useful then Cook distance which remains very small given the size of the sample. Robust method have been found to be more appropriate then traditional methods to detect outliers. In a way, this multivariate approach, allow for detection of of inconsistencies within the explanatory variables themselves.

Appendix VI provides the 4 regressions models used.

The method itself does not define the threshold. It aims to develop selective editing of data. Statistical thresholds of, for example, +/- 3 standardised residuals will provide approximately, under the hypothesis of normality, 0.1% of anomalies. The thresholds can be adjusted in order to select a sufficient number of records to be checked. The manual editing of data can help to detect anomalies in the data sets. If no systematic error is detected at this level, one could thus be pretty confident of the quality of the collection process.

This method has helped Eurostat to spot errors in the data files and to ask MS to carry out further inquiries on these. As long as these errors are not systematic, their impact on Laeken indicators has been found to be limited. However, attention should be paid to their impact on the whole data processing. If mistaken records have been used at the imputation step to define bounds, then the impact on the Laeken

indicators (especially inequality measures) can be significant. If errors have to be corrected, return to data provider, use of administrative sources and expert judgment can be used.

6.2.2.9 External coherence of the data: macro data analysis

The following **tables 13 to 15** show some of the controls and checks that have been done with EU-SILC data, in comparison with external sources[22]

While some of the data are totally comparable between sources (for example: "tenure status"), the degree of comparability of other data are much lower (for example, in the case of social benefits some corrections have been done in order to harmonise sources as much as possible - benefits in kind and lump-sum benefits have been removed from Esspros to make data comparable).

Table 13. Tenure status (%)

COUNTRY	ECHP (%)	EU-SILC (%)	CENSUS (%)
GR			
- Total	100	100	100
- Owners	85	74	74
- Rest	15	26	26
LU			
- Total	100	100	100
- Owners	69	67	67
- Rest	31	33	33
AT			
- Total	100	100	100
- Owners	55	53	---
- Rest	45	47	---

Reference population: Total private households
Data sources:
ECHP : 2001
EU-SILC : 2003
Census : 2001

[22] Weighting factors have been calculated as required to take into account the units' probability of selection, non-response and, as appropriate, to adjust the sample to external data (calibration) relating to the distribution of household and persons in the target population.
When calibration techniques are used, weighted data will be consistent with the external sources used. In this way, the tenure status that has been used in the calibration show consistent data with Census.

Table 14. Self-defined activity status

COUNTRY	Activity status	EU-SILC (%)	LFS (%)
GR	**Total**	100.0	100.0
	At-work	47.5	48.0
	Unemployed	5.4	5.5
	Pupil, student, further training	8.0	7.7
	Others	39.1	38.8
LU	**Total**	100.0	100.0
	At-work	52.9	53.8
	Unemployed	2.3	1.3
	Pupil, student, further training	8.7	8.6
	Others	36.1	36.3
AT	**Total**	100.0	100.0
	At-work	56.9	57.8
	Unemployed	5.3	5.7
	Pupil, student, further training	26.2	26.5
	Others	11.6	10.0

Reference population: Household members aged 16 and over
Data sources:
EU-SILC 2003
LFS 2003 (2nd quarter. For Austria 1st quarter)
ECHP 2001
Comments:
While the sample is not spread over the quarter, the LFS sample is spread over the quarter.

Table 15. Income: Social benefits

	LUXEMBOURG			AUSTRIA		
	Esspros	**SILC**	**ECHP**	**Esspros**	**SILC**	**ECHP**
Total	100.0	100.0	100.0	100.0	100.0	100.0
Unemployed	5.2	4.8	0.9	5.4	4.6	3.5
Old-age/ Survivors/ Invalidity	68.2	75.5	79.7	80.8	80.9	73.1
Family/ Children	22.6	16.4	17.7	12.9	13.9	17.0
Housing	1.5	1.4	0.5	0.3	0.3	6.2
Others	2.5	1.9	1.7	0.3	0.3	0.2

Reference: Total amount of cash-benefits (Sickness benefits have been removed from the total because in the majority of cases they are classified in EU-SILC as salary)
Data sources:

Esspros: 2002
SILC: 2003 (Income reference period 2002)
ECHP: 2001 (Income reference period 2000)

Comments:

- The reference population is different between Esspros and EU-SILC (ECHP) – Esspros collects the pension paid to residents and non-residents, EU-SILC (ECHP) collects information on pensions received by people living in private households in the national territory. When information was available, adjustments were done in order to make data comparable.
- While the distribution of benefits for the Austrian EU-SILC and Esspros data are very close, Luxembourg's data shows big differences. This can be explained for the particular case of Luxembourg where a lot of people (former workers in Luxembourg) are living abroad.
- Analysis based on **table 16** would permit a study of data coherence between waves.
- Data from Greece are not presented in **table 16** due to the fact that the income collected in EU-SILC for Greece is 'net of taxes at source and social insurance contributions'. Information from Esspros is gross.

6.2.2.10 Multilateral comparison

Finally, a multilateral comparison is done to evaluate apparent differences in interpretation/application. In this context it should perhaps be noted that, in accordance with Statistical Programming Committee guidelines, results have anyway to be submitted to member states in advance of publication.

6.2.3 Confidence intervals

For the first time in the history of EU social inclusion indicators, it is foreseen to compute systematic standard errors.

All the computations rely on the linearization technique. The principle is to derive from each EU social inclusion indicator a linear indicator whose variance is asymptotically equal. By computing variance estimates for the so-defined linear indicators, we obtain asymptotic variance estimates for the EU social inclusion indicators. Contrary to the "re-sampling" techniques (Bootstrap, Jackknife….), the linearization method is fully justified on the theoretical point of view. Moreover, once methods are established its implementation is easier.

Variance estimates for linear estimators are then computed using the software POULPE. It is a SAS/AF application which was developed by the INSEE (the French National Institute of Statistics). POULPE can compute variance estimates for linear estimators under quite general sampling designs. Furthermore, in POULPE, the variance component due to non-response, as an additional Poisson phase in the sampling design, and the improvements in the precision due to a calibration of the

sample to external data sources, by applying the residuals technique, are taken into account. However, POULPE cannot take into account the variance component due to imputation.

A typical POULPE process is divided into 4 steps. At each step, a SAS macro carries out preliminary calculations. A SAS/AF environment makes the initialization of the macro parameters interactive and user-friendly. Before running POULPE, it is necessary to create 3 SAS datasets:

- A “design” dataset, where the sampling design is described.

- A “survey” dataset, where are recorded, for each sample unit, the interest variables and some other design variables (identifying variables…).

- A “geographic” dataset where are recorded all the numerical data required for the computation of the inclusion probabilities, and which are not provided by the “survey” dataset.

The creation of these 3 datasets is the most difficult aspect of the POULPE implementation. Once this is fixed, the procedure is systematic and routine (user friendly environment).

For four countries (Belgium, Denmark, Greece and Austria), estimated confidence intervals were calculated for the Laeken indicators. These figures can be found in **appendix IV**. The general level of precision is rather good and in accordance with expectations. For example, the total poverty rate is measured with a precision of +/- 1%, (95% level of confidence).

7. Conclusions and recommendations

General remarks

There is a disruption in the time series of indicators. Data collection under the EU-SILC regulations displays some important differences from its’ predecessor, the ECHP. Similarly, there are important differences between EU-SILC and transitional national data sources. The impact of these various underlying differences on resulting indicators can be important depending on the country and the indicators concerned. Although this paper presents various checks and comparisons which have been made – and further information may become available with the eventual receipt of quality reports – it is impossible to isolate individual causes for all such differences and quantify their impact. Nevertheless, notwithstanding the problems of comparability over time and over space, the validated indicators for 2002 are considered to give important information about income poverty and social exclusion.

Appendix I : Computation of household disposable income

The **Total disposable household income** will be computed as[23]:

The sum for all household members of **gross** personal income components (gross cash or near-cash employee income (PY010G); gross non-cash employee income (PY020G)[24]; gross cash profits or losses from self-employment (including royalties) (PY050G); unemployment benefits (PY090G); old-age benefits (PY100G); survivors' benefits (PY110G); sickness benefits (PY120G); disability benefits (PY130G) and education-related allowances (PY140G)) plus gross income components at household level (income from rental of a property or land (HY040G); family/children-related allowances (HY050G); social exclusion not elsewhere classified (HY060G); housing allowances (HY070G); regular inter-household cash transfers received (HY080G); interests, dividends, profit from capital investments in unincorporated business (HY090G); income received by people aged under 16 (HY110G)) minus (regular taxes on wealth (HY120G); regular inter-household cash transfer paid (HY130G); tax on income and social insurance contributions (HY140G)).
The variable 'tax on income and social insurance contributions' includes tax adjustments-repayment/receipt on income, income tax at source and social insurance contributions (if applicable).

Or as:
The sum for all household members of **net** (of income tax at source and of social contributions) personal income components (cash or near-cash employee income (PY010N); non-cash employee income (PY020N); cash profits or losses from self-employment (PY050N); unemployment benefits (PY090N); old-age benefits (PY100N); survivors' benefits (PY110N); sickness benefits (PY120N); disability benefits (PY130N) and education-related allowances (PY140N)) plus net (of income tax at source and of social contributions) income components at household level (income from rental of a property or land (HY040N); family/children-related allowances (HY050N); social exclusion not elsewhere classified (HY060N); housing allowances (HY070N); regular inter-household cash transfers received (HY080N); interests, dividends, profit from capital investments in unincorporated business (HY090N); income received by people aged under 16 (HY110N)) minus (regular taxes on wealth (HY120G); regular inter-household cash transfer paid (HY130N); repayment/receipt for tax adjustments on income (HY145N)).

Or as:
The sum for all household members of personal income components plus income components at household level, of which some are **net** (net of income tax, net of social contributions or net of both) and others **gross**, or all of them net but some of them net of tax at source, others net of social contributions or net of both, once the tax on income and social insurance contributions (HY140N), the regular taxes on wealth, the regular inter-household cash transfer paid and the employers' social insurance contributions are deducted.
In this case, the variable 'tax on income and social insurance contributions' include repayment/receipt for tax adjustments, income tax at source and social insurance contributions for some income components.

[23] The code in brackets is the EU-SILC reference. See the "Commission Regulation No 1980/2003" for precise definitions.
[24] This variable (with the exception of company car) will be collected only from 2007 onwards.

Appendix II : Time series of the Laeken indicators (website)

Values of the indicators produced from EU-SILC 2003 and national data sources, with comparisons to ECHP for former EU15 countries

See Eurostat website http://epp.eurostat.cec.eu.int/

***Theme*: Population and social conditions**
.*Group*: Living conditions and welfare
..*Domain*: Income and living conditions
…*Collection*: Main Indicators
….*Subset*: Laeken Indicators

Appendix III : Time series of the Pensions indicators (website)

Values of the indicators produced from EU-SILC 2003 and national data sources, with comparisons to ECHP for former EU15 countries

See Eurostat website http://epp.eurostat.cec.eu.int/

***Theme*: Population and social conditions**
.*Group*: Living conditions and welfare
..*Domain*: Income and living conditions
…*Collection*: Main Indicators
….*Subset*: Pensions Indicators

Appendix IV : Confidence intervals

Belgium :

	INDICATOR	VALUE	ST. DEV.	CONFIDENCE INTERVAL AT 95%		CV (%)
				LOWER BOUND	UPPER BOUND	
Belgium	POVERTY THRESHOLD AFTER TRANSFERS					
Total	Total	9 125	98	8 933	9 317	1.1
	POVERTY RATE AFTER TRANSFERS					
Total	Total	15.8	0.52	14.8	16.8	3.3
	Men	14.6	0.56	13.5	15.7	3.8
	Women	16.9	0.6	15.7	18.1	3.6
0-15 years	Total	17.8	0.87	13.1	19.5	4.9
	Men	16.8	1.04	13.8	18.8	6.2
	Women	18.8	1.09	11.3	20.9	5.8
0-64 years	total	14.9	0.51	13.9	15.9	3.4
	men	13.9	0.56	12.8	15.0	4.0
	women	15.9	0.58	14.8	17.0	3.6
16+ years	total	15.1	0.53	14.1	16.1	3.5
	men	13.9	0.58	12.8	15.0	4.2
	women	16.3	0.59	15.1	17.5	3.6
16-64 years	total	13.9	0.49	12.9	14.9	3.5
	men	12.9	0.57	11.8	14.0	4.4
	women	14.9	0.55	13.8	16.0	3.7
16-24 years	total	17.1	0.99	15.2	19.0	5.8
	men	16.7	1.25	14.3	19.2	7.5
	women	17.5	1.15	15.2	19.8	6.6
25-49 years	total	12.8	0.54	11.7	13.9	4.2
	men	11.7	0.63	10.5	12.9	5.4
	women	13.9	0.63	12.7	15.1	4.5
50-64 years	total	14.1	1.01	12.1	16.1	7.2
	men	12.8	1.16	10.5	15.1	9.1
	women	15.3	1.17	13.0	17.6	7.6

	INDICATOR	VALUE	ST. DEV.	CONFIDENCE INTERVAL AT 95%		CV (%)
				LOWER BOUND	UPPER BOUND	
65+ years	total	21.8	1.73	18.4	25.2	7.9
	men	20.1	2.08	16.0	24.2	10.3
	women	22.9	1.79	19.4	26.4	7.8
tenure status	owner	11.9	0.57	10.8	13.0	4.8
	tenant	26.8	1.24	24.4	29.2	4.6
activity status	employed	6.5	0.39	5.7	7.3	6.0
	unemployed	35.1	2.02	31.1	39.1	5.8
	retired	17.6	1.29	15.1	20.1	7.3
	other inactive	25.7	1.16	23.4	28.0	4.5
household type	single , no children (HT=5)	22	1.19	20	24	5.4
	2 adults , no children, both < 65 years (HT=6)	11	0.94	9	13	8.5
	2 adults , no children, one at least >= 65 (HT=7)	20.5	2.21	16	25	10.8
	other households without dependent children (HT=8)	6.8	2.12	3	11	31.2
	single, at least one child (HT=9)	31.3	1.78	28	35	5.7
	2 adults , 1 child (HT=10)	10.7	1.25	8	13	11.7
	2 adults , 2 children (HT=11)	8.2	1.11	6	10	13.5
	2 adults , at least 3 children (HT=12)	18.2	1.85	15	22	10.2
	other households with dependent children (HT=13)	12.8	3.92	5	20	30.6
work intensity	no dependent children, w=0 (WI=1)	26.8	1.61	24	30	6.0
	no dependent children, 0<w<1 (WI=2)	8.2	1.3	6	11	15.9
	no dependent children, w=1 (WI=3)	6	0.71	5	7	11.8
	dependent children, w=0 (WI=4)	57.8	2.75	52	63	4.8
	dependent children, 0<w<0.5 (WI=5)	29.6	1.19	27	32	4.0
	dependent children, 0.5<=w<1 (WI=6)	14.6	1.41	12	17	9.7
	dependent children, w=1 (WI=7)	5.7	0.62	4	7	10.9
	MEDIAN AT-RISK-OF-POVERTY GAP					
Total	total	21.9	1.06	19.8	24.0	4.8
	men	23.7	1.31	21.1	26.3	5.5
	women	21.1	1.04	19.1	23.1	4.9
0-15 years	total	22.4	1.72	19.0	25.8	7.7
16-64 years	total	24.6	1.31	22.0	27.2	5.3
	men	24.7	1.74	21.3	28.1	7.0
	women	24	1.24	21.6	26.4	5.2

	INDICATOR	VALUE	ST. DEV.	CONFIDENCE INTERVAL AT 95%		CV (%)
				LOWER BOUND	UPPER BOUND	
65+ years	total	16.7	1.45	13.9	19.5	8.7
	men	16.7	2.02	12.7	20.7	12.1
	women	16.7	1.43	13.9	19.5	8.6
16+ years	total	21.8	1.07	19.7	23.9	4.9
	men	23.2	1.5	20.3	26.1	6.5
	women	21.1	0.97	19.2	23.0	4.6
	SOME OTHER SOCIAL INDICATORS					
Total	S80/S20	4.5	0.12	4.3	4.7	2.7
	Gini	28.8	0.44	27.9	29.7	1.5

Denmark :

	INDICATOR	VALUE	ST. DEV.	CONFIDENCE INTERVAL AT 95%		CV (%)
				LOWER BOUND	UPPER BOUND	
Denmark	POVERTY THRESHOLD AFTER TRANSFERS					
Total	total	12 553	40	12 475	12 631	0.3
	POVERTY RATE AFTER TRANSFERS					
Total	total	11.7	0.24	11.2	12.2	2.1
	men	11.3	0.26	10.8	11.8	2.3
	women	12.1	0.28	11.6	12.6	2.3
0-15 years	total	9.4	0.38	13.1	10.1	4.0
	men	9.3	0.41	13.8	10.1	4.4
	women	9.5	0.4	11.3	10.3	4.2
0-64 years	total	10.1	0.19	9.7	10.5	1.9
	men	10.1	0.2	9.7	10.5	2.0
	women	10.1	0.2	9.7	10.5	2.0
16+ years	total	12.3	0.25	11.8	12.8	2.0
	men	11.9	0.28	11.4	12.4	2.4
	women	12.7	0.3	12.1	13.3	2.4
16-64 years	total	10.3	0.16	10.0	10.6	1.6
	men	10.4	0.18	10.0	10.8	1.7
	women	10.3	0.18	9.9	10.7	1.7

	INDICATOR	VALUE	ST. DEV.	CONFIDENCE INTERVAL AT 95%		CV (%)
				LOWER BOUND	UPPER BOUND	
16-24 years	total	26.9	0.28	26.4	27.4	1.0
	men	24.4	0.36	23.7	25.1	1.5
	women	29.5	0.31	28.9	30.1	1.1
25-49 years	total	8.8	0.19	8.4	9.2	2.2
	men	9.4	0.22	9.0	9.8	2.3
	women	8.2	0.21	7.8	8.6	2.6
50-64 years	total	4.8	0.27	4.3	5.3	5.6
	men	5	0.3	4.4	5.6	6.0
	women	4.6	0.37	3.9	5.3	8.0
65+ years	total	20.9	1.05	18.8	23.0	5.0
	men	19.6	1.31	17.0	22.2	6.7
	women	21.9	1.24	19.5	24.3	5.7
tenure status	owner	7.6	0.33	7.0	8.2	4.3
	tenant	19.9	0.54	18.8	21.0	2.7
activity status	employed	4.6	0.25	4.1	5.1	5.4
	unemployed	27.3	2.31	22.8	31.8	8.5
	retired	16.9	0.85	15.2	18.6	5.0
	other inactive	36.3	1.1	34.1	38.5	3.0
household type	single , no children (HT=5)	26.4	0.55	25	27	2.1
	2 adults , no children, both < 65 years (HT=6)	4.8	0.24	4	5	5.0
	2 adults , no children, one at least >= 65 (HT=7)	14.5	1.25	12	17	8.6
	other households without dependent children (HT=8)	4.7	1.4	2	7	29.8
	single, at least one child (HT=9)	17.8	1.32	15	20	7.4
	2 adults , 1 child (HT=10)	4.9	0.47	4	6	9.6
	2 adults , 2 children (HT=11)	4.2	0.25	4	5	6.0
	2 adults , at least 3 children (HT=12)	13.4	1.04	11	15	7.8
	other households with dependent children (HT=13)	9.5	2.06	5	14	21.7
work intensity	no dependent children, w=0 (WI=1)	22.6	1.07	21	25	4.7
	no dependent children, 0<w<1 (WI=2)	7.1	0.8	6	9	11.3
	no dependent children, w=1 (WI=3)	4.5	0.37	4	5	8.2
	dependent children, w=0 (WI=4)	50.8	3.59	44	58	7.1
	dependent children, 0<w<0.5 (WI=5)	6.1	2.84	1	12	46.6
	dependent children, 0.5<=w<1 (WI=6)	10.1	1.31	8	13	13.0
	dependent children, w=1 (WI=7)	4	0.42	3	5	10.5

	INDICATOR	VALUE	ST. DEV.	CONFIDENCE INTERVAL AT 95% LOWER BOUND	CONFIDENCE INTERVAL AT 95% UPPER BOUND	CV (%)
	MEDIAN AT-RISK-OF-POVERTY GAP					
Total	total	15	0.66	13.7	16.3	4.4
	men	14	0.77	12.5	15.5	5.5
	women	16.3	0.86	14.6	18.0	5.3
0-15 years	total	19.1	2.58	14.0	24.2	13.5
16-64 years	total	22.9	1.19	20.6	25.2	5.2
	men	21.8	1.55	18.8	24.8	7.1
	women	24.8	1.68	21.5	28.1	6.8
65+ years	total	9.2	0.69	7.8	10.6	7.5
	men	7.7	0.86	6.0	9.4	11.2
	women	10.1	0.83	8.5	11.7	8.2
16+ years	total	14.6	0.61	13.4	15.8	4.2
	men	13.9	0.79	12.4	15.4	5.7
	women	15.4	0.77	13.9	16.9	5.0
	SOME OTHER SOCIAL INDICATORS					
Total	S80/S20	3.6	0.1	3.4	3.8	2.8
	Gini	24.8	0.6	23.6	26.0	2.4

Greece :

	INDICATOR	VALUE	ST. DEV.	CONFIDENCE INTERVAL AT 95% LOWER BOUND	CONFIDENCE INTERVAL AT 95% UPPER BOUND	CV (%)
Greece	POVERTY THRESHOLD AFTER TRANSFERS					
Total	total	4 744	52	4 642	4 846	1.1
	POVERTY RATE AFTER TRANSFERS					
	.					
Total	Total	21	0.52	20.0	22.0	2.5
	Men	20.4	0.56	19.3	21.5	2.7
	Women	21.5	0.55	20.4	22.6	2.6
0-15 years	Total	23.1	1.03	21.0	25.1	4.5
	Men	23.7	1.3	21.2	26.2	5.5
	Women	22.5	1.28	20.0	25.0	5.7

	INDICATOR	VALUE	ST. DEV.	CONFIDENCE INTERVAL AT 95%		CV (%)
				LOWER BOUND	UPPER BOUND	
0-64 years	Total	19.5	0.55	18.4	20.6	2.8
	men	19.4	0.61	18.2	20.6	3.1
	women	19.7	0.59	18.5	20.9	3.0
16+ years	total	20.6	0.5	19.6	21.6	2.4
	men	19.8	0.54	18.7	20.9	2.7
	women	21.4	0.53	20.4	22.4	2.5
16-64 years	total	18.7	0.52	17.7	19.7	2.8
	men	18.4	0.58	17.3	19.5	3.2
	women	19	0.55	17.9	20.1	2.9
16-24 years	total	25.2	1.05	23.1	27.3	4.2
	men	24.9	1.34	22.3	27.5	5.4
	women	25.5	1.31	22.9	28.1	5.1
25-49 years	total	16.7	0.55	15.6	17.8	3.3
	men	16.1	0.6	14.9	17.3	3.7
	women	17.2	0.61	16.0	18.4	3.5
50-64 years	total	18.3	0.82	16.7	19.9	4.5
	men	18.4	0.98	16.5	20.3	5.3
	women	18.2	0.93	16.4	20.0	5.1
65+ years	total	28.2	1.13	26.0	30.4	4.0
	men	25.9	1.31	23.3	28.5	5.1
	women	30.1	1.23	27.7	32.5	4.1
tenure status	owner	21.4	0.58	20.3	22.5	2.7
	tenant	19	1.19	16.7	21.3	6.3
activity status	employed	14.3	0.54	13.2	15.4	3.8
	unemployed	31.7	1.9	28.0	35.4	6.0
	retired	26.8	1.05	24.7	28.9	3.9
	other inactive	25.2	0.86	23.5	26.9	3.4

	INDICATOR	VALUE	ST. DEV.	CONFIDENCE INTERVAL AT 95% LOWER BOUND	CONFIDENCE INTERVAL AT 95% UPPER BOUND	CV (%)
household type	single , no children (HT=5)	27.2	1.38	24	30	5.1
	2 adults , no children, both < 65 years (HT=6)	16	1.3	13	19	8.1
	2 adults , no children, one at least >= 65 (HT=7)	27.9	1.73	25	31	6.2
	other households without dependent children (HT=8)	13.8	0.99	12	16	7.2
	single, at least one child (HT=9)	32.8	5.09	23	43	15.5
	2 adults , 1 child (HT=10)	13.4	1.33	11	16	9.9
	2 adults , 2 children (HT=11)	15.8	1.37	13	18	8.7
	2 adults , at least 3 children (HT=12)	32.5	2.59	27	38	8.0
	other households with dependent children (HT=13)	29.8	1.73	26	33	5.8
work intensity	no dependent children, w=0 (WI=1)	26.4	1.93	23	30	7.3
	no dependent children, 0<w<1 (WI=2)	11.8	0.94	10	14	8.0
	no dependent children, w=1 (WI=3)	13.4	1.19	11	16	8.9
	dependent children, w=0 (WI=4)	53.8	4.56	45	63	8.5
	dependent children, 0<w<0.5 (WI=5)	41.6	3.22	35	48	7.7
	dependent children, 0.5<=w<1 (WI=6)	22.8	1.26	20	25	5.5
	dependent children, w=1 (WI=7)	13.5	1.12	11	16	8.3
	MEDIAN AT-RISK-OF-POVERTY GAP					
Total	total	30.5	1.2	28.1	32.9	3.9
	men	29.7	1.3	27.2	32.2	4.4
	women	30.9	1.1	28.7	33.1	3.6
0-15 years	total	33.2	2.3	28.7	37.7	6.9
16-64 years	total	30.8	1.4	28.1	33.5	4.5
	men	30.9	1.5	28.0	33.8	4.9
	women	30.6	1.5	27.7	33.5	4.9
65+ years	total	27.3	1.5	24.4	30.2	5.5
	men	25.3	1.9	21.6	29.0	7.5
	women	29.3	1.5	26.4	32.2	5.1
16+ years	total	29.7	1.1	27.5	31.9	3.7
	men	29.2	1.3	26.7	31.7	4.5
	women	30.2	1.1	28.0	32.4	3.6
	SOME OTHER SOCIAL INDICATORS					
Total	S80/S20	6.6	0.24	6.1	7.1	3.6
	Gini	35.2	0.49	34.2	36.2	1.4

Ireland :

	INDICATOR	VALUE	ST. DEV.	CONFIDENCE INTERVAL AT 95%		CV (%)
				LOWER BOUND	UPPER BOUND	
Ireland	POVERTY THRESHOLD AFTER TRANSFERS					
Total	total	10 595	118	10 364	10826	1.1
	POVERTY RATE AFTER TRANSFERS					
	.					
Total	total	20.9	0.57	19.8	22.0	2.7
	men	19.7	0.61	18.5	20.9	3.1
	women	22.1	0.67	20.8	23.4	3.0
0-15 years	total	20.5	1.18	13.1	22.8	5.8
	men	20.1	1.33	13.8	22.7	6.6
	women	20.9	1.41	11.3	23.7	6.7
0-64 years	total	18.4	0.58	17.3	19.5	3.2
	men	18.1	0.62	16.9	19.3	3.4
	women	18.7	0.68	17.4	20.0	3.6
16+ years	total	21	0.58	19.9	22.1	2.8
	men	19.5	0.64	18.2	20.8	3.3
	women	22.5	0.68	21.2	23.8	3.0
16-64 years	total	17.7	0.52	16.7	18.7	2.9
	men	17.4	0.58	16.3	18.5	3.3
	women	17.9	0.62	16.7	19.1	3.5
16-24 years	total	18.2	1.07	16.1	20.3	5.9
	men	18.2	1.35	15.6	20.8	7.4
	women	18.2	1.43	15.4	21.0	7.9
25-49 years	total	14	0.56	12.9	15.1	4.0
	men	12.6	0.63	11.4	13.8	5.0
	women	15.4	0.69	14.0	16.8	4.5
50-64 years	total	24.7	1.06	22.6	26.8	4.3
	men	25.9	1.22	23.5	28.3	4.7
	women	23.5	1.32	20.9	26.1	5.6
65+ years	total	41.9	2.34	37.3	46.5	5.6
	men	34.4	2.8	28.9	39.9	8.1
	women	47.9	2.62	42.8	53.0	5.5

	INDICATOR	VALUE	ST. DEV.	CONFIDENCE INTERVAL AT 95% LOWER BOUND	CONFIDENCE INTERVAL AT 95% UPPER BOUND	CV (%)
tenure status	owner	16.2	0.6	15.0	17.4	3.7
	tenant	42.2	1.87	38.5	45.9	4.4
activity status	employed	7.6	0.41	6.8	8.4	5.4
	unemployed	48.9	2.41	44.2	53.6	4.9
	retired	40.3	2.33	35.7	44.9	5.8
	other inactive	37.7	1.05	35.6	39.8	2.8
household type	single , no children (HT=5)	54.8	2.27	50	59	4.1
	2 adults , no children, both < 65 years (HT=6)	16.5	1.15	14	19	7.0
	2 adults , no children, one at least >= 65 (HT=7)	38.5	2.95	33	44	7.7
	other households without dependent children (HT=8)	12.9	1.53	10	16	11.9
	single, at least one child (HT=9)	56.4	4.8	47	66	8.5
	2 adults , 1 child (HT=10)	15.5	1.6	12	19	10.3
	2 adults , 2 children (HT=11)	10.3	1.14	8	13	11.1
	2 adults , at least 3 children (HT=12)	23.7	1.99	20	28	8.4
	other households with dependent children (HT=13)	9.6	1.39	7	12	14.5
work intensity	no dependent children, w=0 (WI=1)	70	2.5	65	75	3.6
	no dependent children, 0<w<1 (WI=2)	13.5	1.28	11	16	9.5
	no dependent children, w=1 (WI=3)	3.7	0.56	3	5	15.1
	dependent children, w=0 (WI=4)	84.9	3.37	78	92	4.0
	dependent children, 0<w<0.5 (WI=5)	45.8	4.12	38	54	9.0
	dependent children, 0.5<=w<1 (WI=6)	13	1.09	11	15	8.4
	dependent children, w=1 (WI=7)	5.6	0.89	4	7	15.9
	MEDIAN AT-RISK-OF-POVERTY GAP					
Total	total	21.9	0.95	20.0	23.8	4.3
	men	23.9	1.06	21.8	26.0	4.4
	women	20.6	1	18.6	22.6	4.9
0-15 years	total	19.1	1.82	15.5	22.7	9.5
16-64 years	total	25.4	1.01	23.4	27.4	4.0
	men	25.7	1.04	23.7	27.7	4.0
	women	25.2	1.29	22.7	27.7	5.1
65+ years	total	14.1	1.13	11.9	16.3	8.0
	men	14.2	1.77	10.7	17.7	12.5
	women	13.9	1.21	11.5	16.3	8.7

	INDICATOR	VALUE	ST. DEV.	CONFIDENCE INTERVAL AT 95%		CV (%)
				LOWER BOUND	UPPER BOUND	
16+ years	total	21.4	0.9	19.6	23.2	4.2
	men	22.9	1.05	20.8	25.0	4.6
	women	20	0.77	18.5	21.5	3.9
	SOME OTHER SOCIAL INDICATORS					
Total	S80/S20	5	0.16	4.7	5.3	3.2
	Gini	30.4	0.39	29.6	31.2	1.3

Austria :

	INDICATOR	VALUE	ST. DEV.	CONFIDENCE INTERVAL AT 95%		CV (%)
				LOWER BOUND	UPPER BOUND	
Austria	POVERTY THRESHOLD AFTER TRANSFERS					
Total	total	9 425	90	9 249	9 601	1.0
	POVERTY RATE AFTER TRANSFERS					
Total	total	13.2	0.57	12.1	14.3	4.3
	men	12.3	0.62	11.1	13.5	5.0
	women	14	0.63	12.8	15.2	4.5
0-15 years	total	15.6	1.2	13.1	18.0	7.7
	men	16.6	1.45	13.8	19.4	8.7
	women	14.6	1.49	11.3	17.5	10.2
0-64 years	total	12.5	0.61	11.3	13.7	4.9
	men	12.1	0.64	10.8	13.4	5.3
	women	12.9	0.69	11.5	14.3	5.3
16+ years	total	12.5	0.51	11.5	13.5	4.1
	men	11.1	0.54	10.0	12.2	4.9
	women	13.8	0.58	12.7	14.9	4.2
16-64 years	total	11.6	0.53	10.6	12.6	4.6
	men	10.8	0.56	9.7	11.9	5.2
	women	12.4	0.62	11.2	13.6	5.0

	INDICATOR	VALUE	ST. DEV.	CONFIDENCE INTERVAL AT 95% LOWER BOUND	CONFIDENCE INTERVAL AT 95% UPPER BOUND	CV (%)
16-24 years	total	11.3	1.02	9.3	13.3	9.0
	men	9.6	1.23	7.2	12.0	12.8
	women	13	1.48	10.1	15.9	11.4
25-49 years	total	11.5	0.63	10.3	12.7	5.5
	men	11.1	0.67	9.8	12.4	6.0
	women	11.8	0.72	10.4	13.2	6.1
50-64 years	total	12.2	0.99	10.3	14.1	8.1
	men	11.1	1.12	8.9	13.3	10.1
	women	13.2	1.18	10.9	15.5	8.9
65+ years	total	16.4	1.12	14.2	18.6	6.8
	men	12.8	1.34	10.2	15.4	10.5
	women	18.8	1.29	16.3	21.3	6.9
tenure status	owner	11.4	0.67	10.1	12.7	5.9
	tenant	16.8	1.1	14.6	19.0	6.5
activity status	employed	8.3	0.49	7.3	9.3	5.9
	unemployed	37.5	3.9	29.9	45.1	10.4
	retired	14.5	0.85	12.8	16.2	5.9
	other inactive	23.8	1.32	21.2	26.4	5.5
household type	single , no children (HT=5)	22.2	1.09	20	24	4.9
	2 adults , no children, both < 65 years (HT=6)	10.7	1.2	8	13	11.2
	2 adults , no children, one at least >= 65 (HT=7)	12.5	1.58	9	16	12.6
	other households without dependent children (HT=8)	6.8	1.24	4	9	18.2
	single, at least one child (HT=9)	30.3	3.75	23	38	12.4
	2 adults , 1 child (HT=10)	9.2	1.49	6	12	16.2
	2 adults , 2 children (HT=11)	12.5	1.56	9	16	12.5
	2 adults , at least 3 children (HT=12)	19.2	3.17	13	25	16.5
	other households with dependent children (HT=13)	8.9	1.82	5	12	20.4
work intensity	no dependent children, w=0 (WI=1)	39.8	3.56	33	47	8.9
	no dependent children, 0<w<1 (WI=2)	12	1.52	9	15	12.7
	no dependent children, w=1 (WI=3)	5.1	0.63	4	6	12.4
	dependent children, w=0 (WI=4)	59.5	9.45	41	78	15.9
	dependent children, 0<w<0.5 (WI=5)	36.7	5.98	25	48	16.3
	dependent children, 0.5<=w<1 (WI=6)	16	1.68	13	19	10.5
	dependent children, w=1 (WI=7)	6.6	0.93	5	8	14.1

	INDICATOR	VALUE	ST. DEV.	CONFIDENCE INTERVAL AT 95%		CV (%)
				LOWER BOUND	UPPER BOUND	
	MEDIAN AT-RISK-OF-POVERTY GAP					
Total	total	19.5	1.2	17.1	21.9	6.2
	men	20.7	1.5	17.8	23.6	7.2
	women	18.9	1.2	16.5	21.3	6.3
0-15 years	total	18.2	2	14.3	22.1	11.0
16-64 years	total	20.9	1.4	18.2	23.6	6.7
	men	21.6	1.7	18.3	24.9	7.9
	women	20.4	1.5	17.5	23.3	7.4
65+ years	total	17.3	1.5	14.4	20.2	8.7
	men	19.5	2.7	14.2	24.8	13.8
	women	16.5	1.3	14.0	19.0	7.9
16+ years	total	20.1	1.2	17.7	22.5	6.0
	men	21.3	1.5	18.4	24.2	7.0
	women	19.1	1.1	16.9	21.3	5.8
	SOME OTHER SOCIAL INDICATORS					
Total	S80/S20	4	0.09	3.8	4.2	2.3
	Gini	27.2	0.47	26.3	28.1	1.7

NB. Confidence intervals not available for Luxembourg or Norway.

Appendix V : EU-SILC data analysis

D01: Number of Records by File

Response Rate
- R01: Non-Response Rate
- R02: DB120 - (Non-)Contact
- R03: DB130 - Non-Response
- R04: DB135 - Acceptance
- R05: RB250 - Individual Response

Fieldwork
- F01: HB100,PB120 - Duration of Interview (Minutes)
- F02: RB260 - Fieldwork Type

Distribution of Population

Population
- P01: Population by Age and Sex
- P02: Population by Age and Sex (weighted: RB050)

Distribution by HH010, HH020, HH080
- P03: Distribution of Population by Dwelling Type (HH010)
- P04: Distribution of Population by Tenure Status (HH020)
- P05: Distribution of Population by Both or Shower in Dwelling (HH080)

Distribution by PE010, PE040
- P10: Distribution of Population by Current Education Activity (PE010)
- P11: Distribution of Population by Highest ISCED Level attained (PE040)

Distribution by PL030, PL040, PL050, PL110, PL140, PL150
- P12: Distribution of Population by Self-defined Current Activity Status (PL030)
- P13: Distribution of Population by Status in Employement (PL040)
- P14: Distribution of Population by Occupation ISCO-88(COM) (PL050)
- P15: Distribution of Population by NACE (PL110)
- P16: Distribution of Population by Type of Contract (PL140)
- P17: Distribution of Population by Managerial Posiotion (PL150)

Income
- I01: Income (Frequencies)
- I02: HY025 - Inflation Factor
- I03: Income Components (Gross/Net)

Imputation of Household Income Components
- I04: Imputation of Household Income Components
- I04a: Imputation of Household Income Components (min mean max)

Imputation of Personal Income Components
- I05: Imputation of Personal Income Components
- I05a: Imputation of Personal Income Components (min mean max)
- I06: Total Income (Total, by Household, by Person)

I07: Household Income (Total, by Household, by Person)
I08: Personal Income (Total, by Person)
I09: Total data collected as ... (%)
I10: Gross/Net data collected as ... (%
I10: Gross data collected as ... (%
I10: Net data collected as ... (%
I11: Net data recorded as ... (%)
I12: Distribution of collected by recording

Main Source of Income

I20: Main Source of Income by group (based on net if available)
I21: Main Source of Income by variable (based on gross if available)

Weights

W01: Weights (Values)
W02: Weights (Flags)

Appendix VI : Detection of outliers for major income components

EMPLOYEE_INCOME = AGE AGE2 SEC_L SEC_U SEC_P UNIV POSTUNIV CDI BOSS PL030 PL060 PL070 PL072

Where,
EMPLOYEE_INCOME: log (PY010)
AGE2: squared age
SEC_L SEC_U SEC_P UNIV POSTUNIV: indicators variables for ISCED level
CDI: unlimited duration contract
BOSS: supervision position
PL030: self-defined activity status for employees
PL060: Total number of hours worked per week
PL070: Total number of months full-time worked during the income reference period
PL072: Total number of months part-time worked during the income reference period

(this regression is restricted on the sub-sample of people who have received a wage (PY010 > 0) and who have been at least at work one month during the income reference period and who are employees (current activity status))

OLD_AGE_BENEFIT=AGE AGE2 CROSSING1 CROSSING2 SEC_L SEC_U SEC_P UNIV POSTUNIV F PL085;

Where,
OLD_AGE_BENEFIT= log (PY100)
F: women indicator
crossing1=indicator of widow women;
crossing2=indicator of married women
PL085: Number of months spent in retirement---

(this regression is restricted on the sub-sample of people having old-age benefits>0 and declaring to be retired)

UNEMPLOYMENT_BENEFIT=SEC_L SEC_U SEC_P UNIV AGE AGE2 PL080 PL070 PL072 PL085;

Where,
UNEMPLOYMENT_BENEFIT = log (PY080)
PL080: Number of months spent in unemployment (numbers of months spent in other activity status are also taken into account in the regression, as they can influence the level of unemployment benefits or early retirement benefits).

(this regression is restricted on the sub-sample of people having unemployment benefits>0 and declaring at least one month of unemployment during the income reference period)

HOUSEHOLD_INCOME = hh_type age assist wi1 jobless secl secp psec univ postuniv nb_children nb_worker

Where,
Household income = log (hy020)
hh_type= indicator of single adult household
assist: percentage of social transfer received
age: median age of the household
wi1 : indicator of full employment
jobless: indicator of jobless
nb_children : number of dependent child
nb_worker : number of workers

(this regression is restricted on the sub-sample of people living in households for which the work intensity is defined, i.e. households with at least one working age person)

Appendix VII : Legal framework for EU-SILC

Framework Regulation:
Regulation of the European Parliament and of the Council (EC) **No.1177/2003**, dated 16 June 2003, concerning Community statistics on income and living conditions (EU-SILC): text with EEA relevance, published in Official Journal L 165, 3/7/2003 P.0001-0009.

Implementation Regulations:
Commission Regulation (EC) **No.1980/2003**, dated 21st October 2003, implementing Regulation (EC) No.1177/2003…(EU-SILC) as regards definitions and updated definitions: text with EEA relevance, published in Official Journal L.298, 17/11/2003 P.0001-0022.

Commission Regulation (EC) **No.1981/2003**, dated 21st October 2003, implementing Regulation (EC) No.1177/2003…(EU-SILC) as regards fieldwork aspects and imputation procedures: text with EEA relevance, published in Official Journal L.298, 17/11/2003 P.0023-0028.

Commission Regulation (EC) **No.1982/2003**, dated 21st October 2003, implementing Regulation (EC) No.1177/2003…(EU-SILC) as regards sampling and tracing rules: text with EEA relevance, published in Official Journal L.298, 17/11/2003 P.0029-0033.

Commission Regulation (EC) **No.1983/2003**, dated 7th November 2003, implementing Regulation (EC) No.1177/2003…(EU-SILC) as regards the list of primary target variables: text with EEA relevance, published in Official Journal L.298, 17/11/2003 P.0034-0085.

Commission Regulation (EC) **No.28/2004**, dated 5th January 2004, implementing Regulation (EC) No.1177/2003…(EU-SILC) as regards the detailed content of intermediate and final quality reports: text with EEA relevance, published in Official Journal L.5, 9/1/2004 P.0042-0056.

Commission Regulation (EC) **No.16/2004**, dated 6th January 2004, implementing Regulation (EC) No.1177/2003…(EU-SILC) as regards the target list of secondary variables relating to “the intergenerational transmission of poverty”: text with EEA relevance, published in Official Journal L.4, 8/1/2004 P.0003-0006.

Commission Regulation (EC) **No.13/2005**, dated 6th January 2005, implementing Regulation (EC) No.1177/2003…(EU-SILC) as regards the target list of secondary variables relating to “social participation”: text with EEA relevance, published in Official Journal L.5, 7/1/2005 P.0005-0009.

Appendix VIII : "Laeken" indicators of social inclusion

(Situation as at May 2005)

Background

In December 2001, the Laeken European Council endorsed a first set of 18 indicators of social exclusion and poverty, organised in a two-level structure of primary indicators – consisting of 10 lead indicators covering the broad fields that have been considered the most important elements in leading to social exclusion – and 8 secondary indicators – intended to support the lead indicators and describe other dimensions of the problem.

After the Laeken European Council, the Indicators Sub-Group has continued working with a view to refining and consolidating the original list of indicators. It highlighted the need to give children a special focus and, to this purpose, to have a standard breakdown by age of all the Laeken indicators, whenever relevant and meaningful (and conditional upon statistical reliability); it redefined the indicator of population living in jobless households and added a new indicator of in-work poverty; pending the adoption of common indicators it recommended the inclusion of tertiary indicators on housing and homelessness. The revised list of commonly agreed indicators as approved by the Social Protection Committee in July 2003, together with their definition, is included in the table below. Those indicators that have been re-defined can be identified thanks to the * sign that has been added in the first column. Similarly, new indicators can be identified thanks to the mention "**new"**.

Breakdowns of the commonly agreed indicators by age, gender and other relevant characteristics

As far as possible, *children* and the *elderly population* must be given a special focus within indicators of social exclusion and poverty. In particular, it is recognised that it is especially important not to base the examination of child poverty and social exclusion on one single at-risk-of-poverty indicator. It is therefore recommended to apply a standard breakdown by broad age groups to all the Laeken indicators, wherever relevant and meaningful. In deciding the degree of disaggregation by age, considerations of statistical robustness must also be taken into account.

Similarly, a *gender breakdown* must be applied to all the indicators, always wherever relevant and meaningful. It should be noted that, in the case of income-based indicators, the gender breakdown is based on the assumption of equal sharing of resources within households. Furthermore, in most instances a gender breakdown is only meaningful when applied to the adult population, as there cannot be any normative interpretation of gender differences in, for example, the poverty risk rate for children.

Two columns in the table below indicate when the age and gender breakdowns have been recommended for the analysis of the situation of poverty and social exclusion. Unless otherwise specified, it is recommended to apply a breakdown by broad age groups, mainly distinguishing between children, the working age population and the elderly population.

As for other relevant dimensions along which the indicators should be examined, they are specified in the definitions below. Clearly, for a good understanding of poverty and social exclusion, these dimensions need to be analysed both in terms of incidence measures – i.e., share of the population in each group who are at risk of poverty – and distribution measures – e.g., distribution of the population at risk of poverty by household type. Both types of measures are closely inter-related and, particularly the latter, require accurate information on the composition of the total population by relevant socio-economic characteristics.

Definitions: the primary indicators

	Indicator	Definition	Age breakdown	Gender breakdown	Specified data source
1	At-risk-of poverty rate	Share of persons with an equivalised disposable income below 60% of the national median equivalised disposable income. Equivalised disposable income is defined as the household's total disposable income divided by its "equivalent size", to take account of the size and composition of the household, and is attributed to each household member.	Yes. Age groups: 0-15; 16 and over; 16-24; 25-49; 50-64; 65+.	Yes (applying to people aged 16 years and over).	ECHP / EU-SILC
1a	At-risk-of-poverty rate by household type	Poverty risk for the total population in the following household types: Households with no dependent children: - Single person, under 65 years old - Single person, 65 years and over - Single women - Single men - Two adults, at least one person 65 years and over - Two adults, both under 65 years - Other households Households with dependent children: - Single parent, 1 or more dependent children - Two adults, one dependent child - Two adults, two dependent children - Two adults, three or more dependent children - Three or more adults with dependent children Dependent children are individuals aged 0 – 15 years and 16 – 24 years if inactive and living with at least one parent.	Already specified in the typology of households.	Already specified in the typology of households.	ECHP / EU-SILC

	Indicator	Definition	Age breakdown	Gender breakdown	Specified data source
1b New	At-risk-of-poverty rate by the work intensity of households	Poverty risk for the total population in different work intensity categories and broad household types. The work intensity of the household refers to the number of months that all working age household members have been working during the income reference year as a proportion of the total number of months that could theoretically be worked within the household. Individuals are classified into work intensity categories that range from WI=0 (jobless household) to WI=1 (full work intensity).	No	No	ECHP / EU-SILC
1c*	At-risk-of-poverty rate by most frequent activity status	Poverty risk for the adult population (aged 16 years and over) in the following most frequent activity status groups: employment (broken down by wage and salary employment and self-employment); unemployment; retirement; other inactivity. The most frequent activity status is defined as the status that individuals declare to have occupied for more than half the number of months in the calendar year.	Yes	Yes	ECHP / EU-SILC
1d	At-risk-of-poverty rate by accommodation tenure status	Poverty risk for the total population in the following accommodation tenure categories: - Owner-occupied or rent free - Rented	Yes	Yes (applying to people aged 16 years and over).	ECHP / EU-SILC
2	At-risk-of-poverty threshold (illustrative values)	The value of the at-risk-of-poverty threshold (60% median national equivalised income) in PPS, Euro and national currency for two illustrative household types: - Single person household - Household with 2 adults, two children	No	No	ECHP / EU-SILC
3	Income quintile share ratio (S80/S20)	Ratio of total income received by the 20% of the country's population with the highest income (top quintile) to that received by the 20% of the country's population with the lowest income (lowest quintile). Income must be understood as equivalised disposable income.	No	No	ECHP / EU-SILC
4	Persistent at-risk-of poverty rate	Share of persons with an equivalised disposable income below the at-risk-of-poverty threshold in the current year and in at least two of the preceding three years.	Yes	Yes (applying to people aged 16 years and over).	ECHP / EU-SILC

	Indicator	Definition	Age breakdown	Gender breakdown	Specified data source
5	Relative median poverty risk gap	Difference between the median equivalised income of persons below the at-risk-of poverty threshold and the threshold itself, expressed as a percentage of the at-risk-of poverty threshold.	Yes	Yes (applying to people aged 16 years and over).	ECHP / EU-SILC
6	Regional cohesion	Coefficient of variation of employment rates at NUTS (Nomenclature of Territorial Units for Statistics) level 2. Employment rates are calculated as the share of the population (aged 15 years or more) who are in employment (ILO definition).	No	Yes	EU-LFS
7	Long term unemploym ent rate	Total long-term unemployed population (≥12 months; ILO definition) as a proportion of total active population aged 15 years or more.	Yes	Yes	EU-LFS
8a*	Population living in jobless households: children	Proportion of children (aged 0-17 years) living in jobless households, expressed as a share of all children.	No	No	EU-LFS
8b*	Population living in jobless households: prime-age adults	Proportion of all people aged 18-59 years who live in a jobless household as a proportion of all people in the same age group. Students aged 18-24 years who live in households composed solely of students are not counted in neither numerator nor denominator.	No	Yes	EU-LFS
9	Early school leavers not in education or training	Share of persons aged 18 to 24 who have only lower secondary education (their highest level of education or training attained is 0, 1 or 2 according to the 1997 International Standard Classification of Education – ISCED 97) and have not received education or training in the four weeks preceding the survey.	No	Yes	EU-LFS
10 New	Low reading literacy performanc e of pupils	Share of 15 years old pupils who are at level 1 or below of the PISA combined reading literacy scole	No	Yes	PISA Survey – OECD
11	Life expectancy	Number of years a person aged 0, 1 and 60 may be expected to live.	No	Yes	Eurostat demographic statistics
12	Self-defined health status by income level.	Proportion of the population aged 16 years and over in the bottom and top quintile of the equivalised income distribution who classify themselves as in a bad or very bad state of health.	Yes	Yes	ECHP / EU-SILC

Definitions: the secondary indicators

	Indicator	Definition	Age breakdown	Gender breakdown	Specified data source
13	Dispersion around the at-risk-of-poverty threshold	Share of persons with an equivalised disposable income below 40%, 50% and 70% of the national equivalised median income.	Yes	Yes (applying to people aged 16 years and over).	ECHP / EU-SILC
14	At-risk-of-poverty rate anchored at a moment in time	In year *t*, share of persons with an equivalised disposable income below the at-risk-of-poverty threshold in year *t-3*, uprated by inflation over the three years.	Yes	Yes (applying to people aged 16 years and over).	ECHP / EU-SILC
15	At-risk-of-poverty rate before social cash transfers	Relative at-risk-of-poverty rate where equivalised income is calculated as follows: - excluding all social cash transfers - including retirement and survivors pensions and excluding all other social cash transfers. - including all social cash transfers (= indicator 1) The same at-risk-of-poverty threshold is used for the three statistics, and is set at 60% of the national median equivalised disposable income (after social cash transfers).	Yes	Yes (applying to people aged 16 years and over).	ECHP / EU-SILC
16	Gini coefficient	Summary measure of the cumulative share of equivalised income accounted for by the cumulative percentages of the number of individuals. Its value ranges from 0% (complete equality) to 100% (complete inequality).	No	No	ECHP / EU-SILC
17	Persistent at-risk-of-poverty rate (50% of median equivalised income)	Share of persons with an equivalised disposable income below 50% of the national median equivalised income in the current year and in at least two of the preceding three years.	Yes	Yes (applying to people aged 16 years and over).	ECHP / EU-SILC
18 New	In-work poverty risk	Individuals who are classified as employed (distinguishing between wage and salary employment and self-employment) according to the definition of most frequent activity status (indicator 1a) and who are at risk of poverty. This indicator needs to be analysed according to personal, job and household characteristics.	Yes	Yes (applying to people aged 16 years and over).	ECHP / EU-SILC
19	Long-term unemployment share	Total long-term unemployed population (≥12 months; ILO definition) as a proportion of the total unemployed population aged 15 years and over.	Yes	Yes	EU-LFS

	Indicator	Definition	Age breakdown	Gender breakdown	Specified data source
20	Very long term unemployment rate	Total very long-term unemployed population (≥24 months; ILO definition) as a proportion of total active population aged 15 years and over.	Yes	Yes	EU-LFS
21	Persons with low educational attainment	Share of the adult population (aged 25 years and over) whose highest level of education or training is ISCED 0, 1 or 2.	Yes. Age groups: 25-34; 35-44; 45-54; 55-64; 25-64; 65 years and over.	Yes	EU-LFS

Data sources for the construction of the common indicators

In order to improve the cross-country comparability of the EU commonly agreed indicators, the Laeken European Council agreed upon common data sources for their calculation as well as common definitions.

The *EU Labour Force Survey (EU LFS)* has been explicitly recognised as the data sources for the construction of all the employment-related commonly agreed indicators. This survey covers all the 25 EU Member States, plus Candidate and EFTA countries. A detailed description of this survey and the definitions used is presented in the publications *Labour Force Survey – Methods and definitions, 2001* and *Labour Force Survey in central and eastern European countries – Methods and definitions, 2000*.

Income-based indicators were initially specified to be calculated on the basis of the *European Community Household Panel (ECHP)*. Information on the characteristics of that survey and availability of data issued from it can be found at the following address: http://forum.europa.eu.int/irc/dsis/echpanel/info/data/information.html. However, this pioneering survey expired in 2001 and is currently being replaced by data collection under the EU-SILC framework regulation and associated implementing regulations[25]. The transitional arrangements are discussed in the current report.

Typically, coverage of household surveys is restricted to private households and excludes certain hard-to-reach groups of the population such as persons who are homeless, nomadic or living in institutions.

Income-based indicators: methodology and limitations

Whilst it is considered to be the best basis for such analyses, income is acknowledged to be an imperfect measure of consumption capabilities and welfare as amongst other things it does not reflect access to credit, access to accumulated savings or ability to liquidate accumulated assets, informal community support arrangements, aspects of non monetary deprivation, differential pricing and other aspects. These factors may be of particular relevance for persons at the lower extreme of the income distribution. The bottom 10 per cent of the income distribution should not, therefore, necessarily be interpreted as having the bottom 10 per cent of living standards.

[25] See **appendix VII**.

Equivalisation

Once total household income is collected, the figures are given per “equivalent adult”, in order to reflect differences in household size and composition. In other words, the total household income is divided by its equivalent size using the so-called “modified OECD” equivalence scale. This scale gives a weight of 1.0 to the first adult, 0.5 to any other household member aged 14 and over and 0.3 to each child. The resulting figure is attributed to each member of the household, whether adult or children. The equivalent size of a household that consists of 2 adults and 2 children below the age of 14 is therefore: 1.0+0.5+(2*0.3) = 2.1.

Detailed methodological guidance

Eurostat has produced computational guidelines describing these indicators (algorithms). The latest version was disseminated in May 2005 (document reference IPSE/73-05/AppendixONE/3.1).

Appendix IX : Indicators of pensions adequacy, sustainability and modernity

(Situation as at May 2005)

Background

Several European Councils, from Lisbon to Barcelona, have highlighted the challenge of an ageing population and its implications for the maintenance of adequate and sustainable pensions. This challenge was underlined in the conclusions of the Stockholm European Council in March 2001, which laid the ground for the open method of coordination on pensions.

This process was finally launched by the Laeken European Council in December 2001 on the basis of eleven common objectives under the following three headings: safeguarding the capacity of systems to meet their social objectives, maintaining their financial sustainability and meeting changing societal needs.

Developing a portfolio of indicators

Building on the achievements under the open method of coordination in the field of social inclusion, the Indicators Sub Group of the Social Protection Committee is again responsible for coordinating the development of pensions indicators.

For the second round of National Strategy Reports to be prepared during 2005, lists of indicators were suggested in a note to Member States circulated by DG.EMPL in October 2004 :

Box A. Main features and major challenges of pension systems

From EU sources (statistics also contained in the data table for country summaries in bold):

Demographic information

- Population breakdown by age groups 0-14, 15-24, 25-44, 45-59, 60-64, 65-74, 75+ (current and projected)
- Life expectancy at birth and at ages 60 and 65, by gender (current and projected for 2010, 2030, 2050)
- Demographic old-age dependency ratio (current and projected for 2010, 2030, 2050): number of persons aged 65+ (60+) in relation to number of working age population (aged 15-64 and 15-59)

Information on household structures

- Housing tenure status: percentage of people aged 65+ (60+) and for complementary age groups (below 65 and below 60) by the housing tenure status of the household they live in (owner-occupied with and without mortgage obligations on the property they live in, rent-free and rented accommodation) (men/women/total).
- Percentage of people aged 65+ (60+) living with their children (men/women/total).
- Percentage of people aged 65+ (60+) living with another adult aged 65+ (60+), men/women/total.
- Percentage of people aged 65+ (60+, 75+) living alone, men/women/total. (Source: LFS)

General socio-economic information

- GDP per capita, recent growth and growth prospects
- Employment and unemployment rates
- Social protection expenditure and **pension expenditure as a % of GDP (Source: ESPROSS)**
- **Public finance situation: debt and deficits**

From national sources:

- Percentage of people aged 65+ (60+, 75+) living in institutions, men/women/total. (Source: administrative data from the Member States)

Box B. Adequacy of pensions

From EU sources (statistics also contained in the data table for country summaries in bold):

- **Risk of poverty for people aged 60+, 65+ and 75+ and below 60, 65, 75 (men/women/total, by household type) (objective 1)**
- Incidence and distribution of risk of poverty for people belonging to the above age groups by the housing tenure status of their households (owner-occupied with and without mortgage obligations on the property they live in, rent-free and rented accommodation).(objective 1)
- Risk of poverty calculated at different income thresholds (40%-50-70% of median national equivalised income) for people aged 60+, 65+ and 75+. (objective 1)
- Relative risk of poverty: risk of poverty for age groups 60+ and 65+ relative to the risk of poverty for complementary age groups (men/women/total). (objective 1)
- Risk of poverty for people whose main activity status is 'retired' (men/women/total). (objective 1)
- **Relative income, i.e. the ratio of median equivalised income of people aged 60+, 65+ and 75+ relative to median equivalised income of people aged <60, <65 and <75 respectively and of people aged 45-54. (objective 2)**
- Composition of income by source, for people aged 60+, 65+, 75+, below 60, below 65, below 75. For each age group: income composition for the group as a whole and for each income quintile. Sources of income: pensions; other social benefits; earnings from work; other sources. (objective 2)
- **Median individual pension income of retirees aged 65-74 in relation to median earnings of employed persons aged 50-59 including and excluding social benefits other than pensions. (objective 2)**
- **Inequality of income distribution (S80/S20),** 60+, **65+** and 75+, men/women/total. **(objective 3)**
- Relative income inequality: income share ratio S80/S20 for age groups 60+, 65+ and 75+ relative to the income share ratio for complementary age groups, men/women/total. (objective 3)

From national sources:

- Any relevant survey data on income and living conditions.
- **Income simulations based on the ISG methodology for theoretical replacement rates (simulations should include interrupted careers due to unemployment, family responsibilities and invalidity; where appropriate, they should be carried out for current scheme rules and for post-reform rules). (objective 2)**
- Current and prospective coverage rates as a percentage of the population aged 15-64 of statutory schemes, occupational schemes and individual schemes; appropriate breakdowns notably by sex, age groups, profession, company size, sector. (objective 2)
- Current and prospective level and share of the income of pensioners provided by statutory schemes, occupational schemes and individual schemes; appropriate breakdowns as above. (objective 2)

Box C. Financial sustainability of pension systems

From EU sources (statistics also contained in the data table for country summaries in bold):

- **Total employment rate: Percentage of people aged** 15-64 **and 30-54 in employment (total/males/females). (objective 4)**
- Current economic or effective old-age dependency ratio: non-active population 65+ (60+) in relation to employed population (aged 15-64; 15-59) (objective 4)
- **Employment rates of older workers: Percentage of people aged 55-59, 60-64, 55-64, and 65-69 in employment (men/women/total). Current and projected scenario (Stockholm target) (objective 5)**
- **Effective age of withdrawal from the labour market (men/women/total) (objective 5)**

From national sources:

- **Projections of public expenditure on pensions (to be validated by the EPC). (objective 6)**
- **Breakdown of expenditure growth by main factors of change (demography, employment, coverage, benefit level)**
- Projected public pensions expenditure per person aged 65+.
- Projected situation of public finances including debt, primary deficits and interest payments. (objective 6)
- Projected evolution of public pension reserve funds. (objective 6)
- Projected budgetary transfers to pension schemes. (objective 6)
- Projected economic or effective old-age dependency ratio: non-active population 65+ (60+) in relation to employed population (aged 15-64; 15-59). (objective 4)
- Stock and flow data (number of beneficiaries) on benefits allowing an early withdrawal from the labour market (see SPC special study on promoting longer working lives for types of benefits to be considered). (objective 5)
- Current and future contribution rates to pension schemes (as far as possible, distinguishing between old age, invalidity and survivors benefits and between contributions to the main public and typical private schemes). (objective 8)
- Current and projected level of reserves of public and private pension schemes in % of GDP. (objective 8)
- Current and projected real rates of return on assets held by pension reserve funds (public and private). (objective 8)
- Current and projected composition of assets held by pension reserve funds (public and private). (objective 8)

Box D. Modernisation of pension systems

From EU sources:

- Gender differences in the risk of poverty by age group (60+, 65+ and 75+ and below 60, 65, 75); calculated for all household types and for women/men living alone. (objective 10)
- Percentage point difference between men and women in the relative income, i.e. the ratio of median equivalised income of people aged 60+, 65+ and 75+ relative to median equivalised income of people aged <60, <65 and <75 respectively and of people aged 45-54; calculated for all household types and for women/men living alone. (objective 10)

From national sources:

- Typical length of vesting/waiting periods. (objective 9)
- Average pension entitlements by sex, individual and derived rights, pensioners aged 65-74 and 75+. (objective 10)

Logically, the same sources should be used to produce the Pensions Indicators as are also used to produce the Laeken indicators of social inclusion[26], wherever possible.

Detailed methodological guidance

Eurostat has produced computational guidelines describing these indicators (algorithms). The latest version was disseminated in May 2005 (document reference IPSE/73-05/AppendixONE/3.2).

[26] See **appendix VIII**.