EUROPEAN COMMISSION
EUROSTAT

Directorate F: Social Statistics and Information Society
Unit F-3: Living conditions and social protection statistics

Comparative Final EU Quality Report 2006

(Version 3 – October 2009)

TABLE OF CONTENTS

0. LEGAL BASIS

Commission Regulation (EC) No 1177/2003 concerning Community statistics on income and living conditions (EU-SILC) in its Article 16 states the following:

> *1. Member States shall produce by the end of the year N+1 an intermediate quality report relating to the common cross-sectional EU indicators based on the cross-sectional component of year N.*
>
> *Member States shall produce by the end of year N+2 final quality reports that cover both cross-sectional and longitudinal components in relation to the year of the survey N, focusing on the internal accuracy.* […]
>
> *2. The Commission (Eurostat) shall produce by the end of June N+2 a comparative intermediate quality report relating to the common cross-sectional EU indicators of year N.*
>
> *The Commission (Eurostat) shall produce by 30 June N+3 a comparative final quality report that covers both cross-sectional and longitudinal components in relation to the year of the survey N.* […]

The comparative final quality report for 2006 aims at gathering and summarizing all the information contained in the 2006 national final quality reports that the Member States have sent to Eurostat. The objective here is to evaluate the quality of the instrument from the European point of view, i.e. by establishing between-country comparisons of some of its key quality dimensions.

The quality aspects described in this document are those specified in the Commission Regulation N° 28/2004 (Annex IV) as regards the detailed content of final quality reports to be produced by Eurostat.

1. RELEVANCE

The relevance of an instrument has to be assessed in the light of the needs of its users. As for EU-SILC the main users are:

- Institutional users like DG EMPL of the Commission and the Social Protection Committee, in charge of the monitoring of social protection and social inclusion, or other Commission services;
- Statistical users in Eurostat or in Member States National Statistical Institutes to feed sectoral or transversal publications such as the Annual Progress Report on the Lisbon Strategy (structural indicators), the Sustainable Development Strategy monitoring report, the Eurostat yearbook and various pocketbooks, among other reports;
- Researchers having access to microdata; and

- End users – including the media - interested in living conditions and social cohesion in the EU.

With the 2006 operation covering the then 25 Member States plus Norway and Iceland, EU-SILC has proved to be the main source for comparable indicators for monitoring and reporting on living conditions and social cohesion at the EU level. The relevance of the instrument is very high among most users although suggestions for improvement have been clearly expressed for instance during the first EU-SILC Users' Conference recently organised in Mannheim (5-6 March 2009)[1] and the Joint OECD/University of Maryland international Conference held in Paris (16-18 March 2009)[2]:

- Institutional users are looking for more timely results that can be better synchronised with the actual policy needs. For instance, in the context of the current financial crisis, policy-makers are looking for indicators to assess its social impact on households which is only possible with a long time lag in the current EU-SILC set-up.

- Statistical users are keen to have stable results without too significant revisions so that reports or publications relying on a long process maintain their relevance.

- Researchers ask for clean and harmonised datasets with a better documentation (more specific metadata, sample structure and weighting procedures, computation of income components, etc.) and information on the production process and revision. Researchers are also looking for a softening of anonymisation rules (migrations, occupation, regional level, etc.) to increase the possibilities to analyse some important issues.

These elements are taken into account to the maximum extent in the process of improvement of the instrument which will continue in the next years.

2. ACCURACY

The concept of accuracy refers to the reliability of estimates computed from a sample rather than the entire population. This section dwells on methodological features of the EU-SILC samples surveyed in each country and intents to draw a picture of their relevance for estimation purposes.

2.1. Sample design

In 2006, the EU-SILC instrument covered 27 countries: twelve carried out the survey for the second time, while eight did it for the third time and seven countries for the fourth time.

---

1 http://www.gesis.org/forschung-lehre/veranstaltungen/konferenzen/european-user-conference/

2 http://www.oecd.org/document/60/0,3343,en_2649_33933_42139644_1_1_1_1,00.html

The Framework Regulation calls for the selection of nationally representative probabilistic samples, with the exception of Germany where quota samples can be used until 2008[3].

The observation units are both households and individuals. Households are clusters of individuals and all the members of a selected household are eligible for inclusion in the sample. The following table summarizes the sampling design by country[4].

Table 1: Sampling design (2006)

Sampling of dwellings/ addresses	Simple random sampling	Malta, Austria
	Stratified simple random sampling	Luxembourg
	Stratified multi-stage sampling	Czech Republic, Spain, France, Hungary, Latvia, The Netherlands, Poland, Portugal, United Kingdom
Sampling of households	Stratified simple random sampling	Cyprus, Slovakia
	Stratified multi-stage sampling	Belgium, Greece, Ireland, Italy
	Quota plus sampling based on an ACCESS panel	Germany
Sampling of individuals	Simple random or systematic sampling	Denmark, Iceland, Sweden, Norway
	Stratified simple random or systematic sampling	Estonia, Lithuania
	Stratified two-phase sampling	Finland
	Stratified two-stage sampling	Slovenia

Source: National Quality Reports 2006.

Countries that carry out a sampling of individuals generally select persons of age 16 and over. They do not include members aged between 14 and 16 in their sample of 'selected respondents' in order to active them in the panel when they become 16, as recommended by Eurostat. Only Estonia follows the guidelines. Denmark deviates from the Eurostat rules as the sampling frame in this country is all persons aged 13 and over but households where the selected person is less that 16 at the beginning of the survey year are not interviewed at all for that wave.

EU-SILC data is collected by an interview with the exception of seven countries where most or part of the information is administrative, gathered from national registers. These so-called 'register countries' are Denmark, The Netherlands, Slovenia, Finland, Sweden, Iceland and Norway.

Most of the countries have adopted the 4-year rotational design recommended by Eurostat[5]. Norway and France have longer panel duration (8 and 9 years respectively) and Luxembourg and Sweden have a pure panel supplemented with a new sample each year.

[3] In Germany for 2006, 50% of the data is based on probability sampling and 50% on quota samples.
[4] A detailed description of the sample design by country can be found in the Annex.
[5] 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.

Nevertheless, some of the countries have deliberately departed a bit from the standard in order to ensure a minimum sample size:

- Czech Republic: due to the relatively small sample size in 2005, all responding households were carried over to the 2006 operation. One new sample replication was added in 2006.

- Estonia: in 2004 households were randomly divided into four rotational groups. According to original rotational scheme, one of these groups was to be dropped in 2005 and another in 2006, but due to lower than expected response rate, it was decided to keep all the rotational groups in the sample. New subsamples were also introduced into the survey in 2005 and 2006 to ensure cross-sectional representativeness. Thus, in 2006 the sample consists of six rotational groups (four started in 2004, one started in 2005 and one started in 2006).

- The Netherlands: in order to ensure minimum longitudinal sample sizes, the sizes of the four rotation groups have been made unequal. Subsamples of respondents that participate longer in the EU-SILC survey are larger (one subsample was purely cross-sectional and was not followed up in 2006; respondents in the second subsample will participate for two years, in the third subsample for three years, and in the fourth subsample for four years).

- Austria: the rotational group 1 of 2004, which under normal circumstances would have dropped out of the sample in 2005, was added to the rotational group 4 in 2005 to secure a sufficient number of households in the longitudinal sample; as a result the longitudinal component consists of the rotational groups 1, 3, 4 of 2004, which have been recoded in rotational groups 3, 4 in the longitudinal 2006 data files.

- Slovenia: in 2006 one rotational group from 2005 should have been dropped, but it was kept, divided into three parts and reallocated to the remaining three rotational groups of 2005. Therefore all households which responded in 2005 were interviewed again in 2006. In addition, one new rotational group was added, so to have four rotational groups in 2006.

- Sweden: the rotational groups from previous waves are complemented with young people and immigrants who have "grown into the population", constructing a special sampling frame with those individuals and making a systematic random sampling.

Some countries are using alternative survey structures, essentially for integrating EU-SILC into existing surveys:

- Finland uses a modified rotational schema in which the basic two-year rotational panel forming the cross-sectional survey is supplemented by the follow-up of two subsamples for two additional years, to be able to have a four-year trajectory for the longitudinal survey.

  The Finnish cross-sectional SILC data collection year 2006 contains two groups based on the Income Distribution Survey (IDS). The first wave of the EU-SILC longitudinal component selected in 2006 is selected randomly within strata from the first wave of the IDS proportionally to the size of the IDS sample.

- Sweden has two separate operations: a cross-sectional survey and a pure long-term panel, i.e. for the 2004-2006 longitudinal survey they did a separate sample starting in 2004 with four panels to rotate according to the Regulations.

- Luxembourg uses a pure panel to carry out the longitudinal survey, which is supplemented every year with a new sample in order to ensure cross-sectional representation which is also followed up in subsequent years.

2.2. **Sampling errors**

Commission Regulation (EC) No 28/2004, Annex III, specifies the set of statistics for which information on sampling errors should be presented in national final quality reports. In specific terms, the Commission Regulation specifies the following requirements: "For the EU-SILC cross-sectional component and for each wave of the EU-SILC longitudinal component, the following information will be provided":

1. The mean, the total number of observations (before and after imputation) and the standard error for the following income components:

   Total household income (4 variables)
   Net income components at household level
   Gross income components at household level
   Net income components at personal level
   Gross income components at personal level

2. The mean number of observations (before and after imputation) and the standard error for the equivalised disposable income breakdown by sex, age group and household size:

   Subclasses by household size (4 classes)
   Population by age group (6 classes)
   Population by sex (2 classes)

This section highlights some main results on sampling errors in EU-SILC surveys from a comparative perspective. A description of the basic methodology and approach followed in the production of these statistics can be found in the annex.

Sampling error results for selected countries

Detailed results, computed using the standard SAS programs developed for the purpose and following uniform specifications, are shown below for a subset of countries: Austria, Denmark, Iceland, Sweden, Estonia, Lithuania.

The following tables show the sample sizes (households or persons, as relevant) and the percentage standard error. More detailed results are shown for the same countries in the annex, where values of the estimate, its standard error, sample size and design effect (deft) are shown for each statistic included in the final quality report. The variable groups (and the relevant analysis units) are as follows:

- Total household income (all households)

- Household level income components (households receiving the component concerned)
- Personal level income components (persons receiving the component concerned)
- Equivalised mean income by household size (all households)
- Equivalised mean income by age class and by gender (all persons)

Sampling errors for each measure are presented for three sample bases:

(1) Full cross-sectional sample. Source: Cross-sectional data 2006.

(2) 2-years longitudinal sample (2005 and 2006). Source: Longitudinal data 2006.

(3) 3-years longitudinal sample (2004, 2005 and 2006). Source: Longitudinal data 2006.

Table 2: Sampling errors (summary)

	Austria			Denmark			Iceland			Sweden			Estonia*			Lithuania**		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
TOTAL SAMPLE SIZE																		
households																		
total	6.028	3.809	2.102	5.711	3.210	1.997	2.838	1.722	1.109	6.803	4.449	2.903		3.894	3.334		2.971	
minimum	53	28	21	134	85	53	65	29	15	114	67	44		27	21		4	
persons																		
total	14.883	9.724	5.513	14.676	8.353	5.216	8.587	5.163	3.463	17.149	11.924	7.644		11.270	9.724		8.205	
minimum	29	19	9	75	39	19	12	10	5	0	65	40		4	3		66	
RELATIVE STANDARD ERROR (%)																		
maximum	32	42	34	20	23	50	22	35	44	16	19	22		96	94		64	
mean	5	7	9	3	4	6	5	7	9	4	5	6		12	12		9	

(1) Full cross-sectional sample
(2) 2-yars longitudinal sample (2005 & 2006)
(3) 3-yars longitudinal sample (2004, 2005 & 2006)

* Estonia (1) not computed
** Lithuania (1) not computed; (3) survey stated only in 2005.

The upper part of the table shows the total sample size and the minimum sample size encountered for any subclass of the sample for which the estimates have been produced. The sample sizes are in terms of households or persons as relevant. Obviously, the sample size declines as we move from the full cross-sectional sample to the 2-year longitudinal and then the 3-year longitudinal sample base for each country. The extent of reduction is quite varied across countries, however. This is because the sample sizes of the new panels introduced each wave are adjusted so as to meet the overall sample size requirements for that year.

Note that the estimates for some income components are based on extremely small sample sizes. This is when the component concerned is received by very few households or persons.

The lower part of the total shows the value of relative standard error (%) encounter: its mean value over all the statistics included in these tables, and the maximum value of the error encountered. The mean value reflects the variation in national sample sizes. The maximum value is generally for the estimates based on the smallest number of sample cases.

Table 3: Sampling errors

SAMPLE BASE: households or persons, as relevant

(1) Full cross-sectional sample. Source: Cross-sectional data 2006.
(2) 2-yars longitudinal sample (2005 & 2006). Source: Longitudinal data 2006.
(3) 3-yars longitudinal sample (2004, 2005 & 2006). Source: Longitudinal data 2006.

Variable		SAMPLE SIZE																	
		Austria			Denmark			Iceland			Sweden			Estonia*			Lithuania**		
		(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Total household income (all households)																			
Total household gross income	**HY010**	6.028	3.809	2.102	5.711	3.210	1.997	2.838	1.722	1.109	6.803	4.449	2.903		3.897	3.336		2.971	
Total disposable household income	**HY020**	6.028	3.809	2.102	5.711	3.210	1.997	2.838	1.722	1.109	6.803	4.449	2.903		3.897	3.336		2.971	
... excluding transfers except pensions	**HY022**	6.028	3.809	2.102	5.711	3.210	1.997	2.838	1.722	1.109	6.803	4.449	2.903		3.897	3.336		2.971	
... excluding all transfers	**HY023**	6.028	3.809	2.102	5.711	3.210	1.997	2.838	1.722	1.109	6.803	4.449	2.903		3.897	3.336		2.971	
Household level income components (households receiving the component concerned)																			
Property income	**HY040n**	226	148	85							114	67	44		63	57		154	
	HY040g	225	147	84	134	85	53	150	87	56	114	67	44		63	57		154	
Family/Children allowances	**HY050n**	2.120	1.381	804							2.256	1.416	927		1.598	1.378		419	
	HY050g	2.120	1.381	804	2.236	1.263	798	1.005	619	386	2.256	1.416	927		1.598	1.378		419	
Other social exclusions	**HY060n**	120	71	34							211	103	56		27	21			
	HY060g	120	71	34				65	29	15	211	103	56		27	21		96	
Housing allowances	**HY070n**	204	123	66							626	372	225		97	86			
	HY070g	204	123	66	693	295	177	1.010	632	413	626	372	225		97	86		130	
Inter-household transfers received	**HY080n**	410	256	139							347	211	147		135	123			
	HY080g	410	256	139	333	160	95	448	255	150	347	211	147		135	123		215	
Capital income	**HY090n**	4.588	2.917	1.619							5.203	3.555	2.363		1.107	955		98	
	HY090g	4.588	2.917	1.619	5.640	3.179	1.980	1.870	1.174	743	5.203	3.555	2.363		1.107	955		98	
Mortgage interest	**HY100n**										3.387	2.358	1.585		321	280			
	HY100g				3.338	1.993	1.243	2.015	1.225	774	3.387	2.358	1.585		321	280		71	
Children's income	**HY110n**	53	28	21							1.194	849	569		63	56		4	
	HY110g	53	28	21	540	318	213	379	238	137	1.194	849	569		63	56		4	
Regular taxes on wealth	**HY120n**										4.581	3.053	1.997		2.522	2.172			
	HY120g				4.100	2.434	1.530	2.473	1.515	981	4.581	3.053	1.997		2.522	2.172		582	
Inter-household transfers paid	**HY130n**	392	242	132							140	86	62		185	158			
	HY130g	392	242	132	263	124	76	419	252	162	140	86	62		185	158		248	
Tax	**HY140n**										6.681	4.410	2.879						
	HY140g	5.923	3.749	2.076	5.678	3.196	1.987	2.836	1.721	1.108	6.681	4.410	2.879		2.875	2.486		1.941	
Tax adjustment	**HY145n**	2.499	1.683	943											1.346	1.152		424	

(cont.)

Personal level income components (persons receiving the component concerned)																
Employee cash or near cash income	**PY010n**	6.254	3.978	2.209							9.640	6.315	4.129	4.932	4.279	3.119
	PY010g	6.254	3.978	2.209	8.443	4.815	2.975	5.497	3.228	2.026	9.640	6.315	4.129	4.932	4.279	3.119
Non-Cash employee income	**PY020n**										2.148	1.490	996	165	142	
	PY020g				680	404	242				2.148	1.490	996	165	142	66
Contributions to private pension	**PY035n**	2.732	1.815	1.029							4.157	2.744	1.833	479	413	
	PY035g	2.732	1.815	1.029							4.157	2.744	1.833	479	413	92
Self-employment income	**PY050n**	1.098	711	405							1.839	1.191	791	713	631	615
	PY050g	1.098	711	405	2.842	1.616	997	678	426	255	1.839	1.191	791	717	634	615
Production for own consumption	**PY070n**	259	164	88												
	PY070g	259	164	88												
Pension from private plans	**PY080n**	29	19	9							974	692	458	4	3	
	PY080g	29	19	9							974	692	458	4	3	
Unemployment benefits	**PY090n**	724	465	237							1.239	785	528	133	110	101
	PY090g	724	465	237	2.061	1.160	723	211	117	76	1.239	785	528	133	110	101
Old-age benefits	**PY100n**	3.045	1.943	1.098							2.728	1.930	1.267	2.334	1.983	
	PY100g	3.045	1.943	1.098	1.574	916	608	740	479	313	2.728	1.930	1.267	2.334	1.983	1.789
Survivor' benefits	**PY110n**	105	72	37							118	65	40	84	76	
	PY110g	105	72	37	75	39	19	313	170	93	118	65	40	84	76	160
Sickness benefits	**PY120n**	181	106	62							2.523	1.704	1.140	594	511	
	PY120g	181	106	62	1.059	558	327	12	10	5	2.523	1.704	1.140	594	511	
Disability benefits	**PY130n**	366	241	131							929	611	406	596	514	
	PY130g	366	241	131	688	369	223	298	176	113	929	611	406	596	514	446
Education-related allowances	**PY140n**	178	102	67							2.053	1.292	806	182	156	
	PY140g	178	102	67	797	407	258	216	126	76	2.053	1.292	806	182	156	225
Employees' gross monthly earnings	**PY200g**	5.682						4.477			0					
Equivalised mean income by household size (all households)																
1 household member	**HX090**	1.755	1.040	544	1.109	563	350	383	238	158	1.681	1.201	773	783	673	630
2 household members	**HX090**	1.823	1.189	653	2.239	1.288	785	829	508	347	2.389	1.521	1.009	1.113	914	917
3 household members	**HX090**	1.053	676	376	895	498	329	573	352	230	1.010	636	412	834	729	635
4 and more	**HX090**	1.397	904	529	1.468	861	533	1.053	624	374	1.723	1.091	709	1.164	1.018	789
all households	**HX090**	6.028	3.809	2.102	5.711	3.210	1.997	2.838	1.722	1.109	6.803	4.449	2.903	3.894	3.334	2.971
Equivalised mean income by age class and by gender (all persons)																
<25	**HX090**	4.513	3.063	1.805	4.874	2.708	1.686	3.489	2.076	1.484	6.162	4.808	2.985	4.089	3.566	2.726
25 to 34	**HX090**	1.677	1.058	572	1.602	828	479	1.051	651	412	2.059	1.285	862	1.074	929	732
35 to 44	**HX090**	2.382	1.520	841	2.286	1.345	848	1.202	727	461	2.388	1.547	991	1.466	1.250	1.143
45 to 54	**HX090**	2.121	1.363	762	2.276	1.307	837	1.258	707	436	2.475	1.550	1.013	1.599	1.382	1.264
55 to 64	**HX090**	1.859	1.249	689	2.004	1.216	736	791	489	329	2.044	1.365	905	1.261	1.086	996
65+	**HX090**	2.331	1.471	844	1.634	949	630	796	513	341	2.021	1.369	888	1.781	1.511	1.344
Male	**HX090**	7.178	4.716	2.649	7.323	4.198	2.632	4.300	2.593	1.747	8.452	6.012	3.851	5.277	4.534	3.819
Female	**HX090**	7.705	5.008	2.864	7.353	4.155	2.584	4.287	2.570	1.716	8.697	5.912	3.793	5.993	5.190	4.386
all persons	**HX090**	14.883	9.724	5.513	14.676	8.353	5.216	8.587	5.163	3.463	17.149	11.924	7.644	11.270	9.724	8.205

* Estonia (1) not computed

** Lithuania (1) not computed; (3) survey stated only in 2005.

Table 4: Sampling errors

Relative standard error (%)

(1) Full cross-sectional sample. Source: Cross-sectional data 2006.
(2) 2-yars longitudinal sample (2005 & 2006). Source: Longitudinal data 2006.
(3) 3-yars longitudinal sample (2004, 2005 & 2006). Source: Longitudinal data 2006.

RELATIVE STANDARD ERROR (%)

Variable		Austria			Denmark			Iceland			Sweden			Estonia*			Lithuania**		
		(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Total household income (unit: household)																			
Total household gross income	**HY010**	1,0	1,2	1,5	1,1	1,4	2,3	1,4	2,4	2,9	0,9	1,1	1,6		2,0	1,9		2,1	
Total disposable household income	**HY020**	0,9	1,2	1,4	1,0	1,4	2,3	1,5	2,5	3,0	0,8	1,0	1,4		1,8	1,7		1,9	
... excluding transfers except pensions	**HY022**	1,0	1,3	1,6	1,3	1,6	2,7	1,6	2,7	3,3	1,0	1,2	1,5		1,9	1,8		2,0	
... excluding all transfers	**HY023**	1,5	2,0	2,3	1,5	2,1	3,2	1,9	3,3	4,1	1,3	1,8	2,1		2,4	2,4		2,6	
Household level income components (unit: household)																			
Property income	**HY040n**	12,2	21,9	24,7							9,5	15,3	21,6		24,7	24,9		16,7	
	HY040g	13,9	23,1	25,3	10,9	15,4	15,5	9,7	18,4	21,9	9,5	15,3	21,6		24,7	24,9		17,1	
Family/Children allowances	**HY050n**	1,5	1,9	2,7							2,2	2,2	2,5		3,6	3,4		6,0	
	HY050g	1,5	1,9	2,7	1,5	2,5	2,4	5,0	7,7	8,0	2,5	2,5	2,8		4,2	4,0		6,6	
Other social exclusions	**HY060n**	19,6	17,4	25,0							7,8	10,0	14,8		30,4	34,8			
	HY060g	19,6	17,4	25,0				17,9	22,5	38,5	7,8	10,0	14,8		30,4	34,8		15,0	
Housing allowances	**HY070n**	4,9	6,2	9,5							2,9	3,9	5,4		8,6	9,9			
	HY070g	4,9	6,2	9,5	3,1	4,9	7,9	1,9	3,0	2,8	2,9	3,9	5,4		8,6	9,9		6,3	
Inter-household transfers received	**HY080n**	6,2	6,7	8,5							4,2	5,6	6,2		15,3	12,5			
	HY080g	6,2	6,7	8,5	4,7	4,5	7,3	4,3	5,8	7,0	4,2	5,6	6,2		15,3	12,5		8,3	
Capital income	**HY090n**	6,3	8,2	11,6							11,1	15,4	18,9		36,1	32,4		31,9	
	HY090g	6,3	8,2	11,6	19,6	23,1	49,9	12,3	21,5	21,5	11,1	15,4	18,9		32,5	34,6		32,8	
Mortgage interest	**HY100n**										2,1	2,2	3,3		7,3	8,0			
	HY100g				1,3	1,7	2,1	2,7	5,0	6,6	2,1	2,2	3,3		7,3	8,0		17,9	
Children's income	**HY110n**	11,5	16,3	19,2							15,9	15,8	20,1		21,1	21,8		64,0	
	HY110g	12,1	16,2	19,2	11,4	13,4	8,8	8,0	11,7	15,8	15,5	15,4	19,8		21,2	21,9		64,0	
Regular taxes on wealth	**HY120n**										3,7	6,2	7,5		4,2	4,5			
	HY120g				1,4	1,7	2,3	2,0	2,6	3,3	3,7	6,2	7,5		4,2	4,5		5,4	
Inter-household transfers paid	**HY130n**	4,7	5,9	7,4							5,8	8,2	10,3		7,0	7,2			
	HY130g	4,7	5,9	7,4	4,8	5,0	8,4	4,4	7,2	8,8	5,8	8,2	10,3		6,9	7,2		7,8	
Tax	**HY140n**										1,2	1,4	1,9						
	HY140g	1,5	1,8	2,2	1,3	1,6	2,5	1,9	2,6	3,7	1,2	1,4	1,9		2,8	2,8		3,0	
Tax adjustment	**HY145n**	-13,3	-15,7	-23,2											-4,1	-4,7		-4,7	

(cont.)

Personal level income components (unit: person)																	
Employee cash or near cash income	**PY010n**	0,9	1,1	1,6							0,7	0,8	1,0		1,8	1,8		1,9	
	PY010g	1,1	1,3	1,8	0,9	1,2	1,6	1,3	1,6	1,9	0,8	0,9	1,2		1,9	1,9		2,0	
Non-Cash employee income	**PY020n**										3,8	14,7	6,3		6,7	7,6			
	PY020g				5,9	7,3	11,0				3,9	18,9	6,6		7,0	7,8		11,1	
Contributions to private pension	**PY035n**	2,6	4,0	6,1							1,9	2,7	3,3		5,3	5,6			
	PY035g	2,6	4,0	6,1							1,9	2,7	3,3		5,3	5,6		14,0	
Self-employment income	**PY050n**	4,0	4,5	6,1							7,2	10,3	11,4		8,7	10,6		5,9	
	PY050g	3,8	4,6	6,2	9,0	8,4	15,5	6,0	6,8	9,3	7,2	9,3	10,1		9,1	11,3		5,9	
Production for own consumption	**PY070n**	7,6	11,0	14,0															
	PY070g	7,6	11,0	14,0															
Pension from private plans	**PY080n**	27,9	36,0	33,9							5,2	5,5	7,4		95,9	94,4			
	PY080g	32,2	41,7	33,8							5,8	5,9	7,7		95,9	94,4			
Unemployment benefits	**PY090n**	3,6	4,2	7,6							2,6	3,4	3,7		13,6	16,2		8,4	
	PY090g	3,7	4,6	8,2	2,3	3,1	4,6	6,6	10,4	12,5	2,7	3,4	3,7		14,4	17,1		8,7	
Old-age benefits	**PY100n**	1,2	1,5	2,3							1,2	1,3	1,5		0,7	0,6			
	PY100g	1,4	1,8	2,7	1,5	1,9	2,5	2,3	2,7	2,7	1,3	1,5	1,8		1,2	0,7		0,9	
Survivor' benefits	**PY110n**	6,3	6,1	7,2							5,1	6,1	8,1		6,7	6,7			
	PY110g	6,8	6,9	8,6	11,2	18,1	25,8	12,7	20,7	26,5	5,6	6,7	8,8		6,7	6,7		6,3	
Sickness benefits	**PY120n**	12,2	16,9	18,9							3,7	4,8	5,7		11,0	12,5			
	PY120g	11,6	15,5	19,4	5,5	7,2	13,0	15,1	21,6	38,3	3,7	4,9	5,7		11,5	13,1			
Disability benefits	**PY130n**	3,2	3,6	4,0							1,8	2,1	2,5		2,3	2,4			
	PY130g	3,7	4,1	4,5	2,6	3,0	4,8	5,2	5,5	7,4	1,9	2,2	2,6		2,3	2,4		2,3	
Education-related allowances	**PY140n**	10,5	23,5	33,2							2,7	3,2	4,8		31,4	37,3			
	PY140g	10,5	23,5	33,2	3,0	4,9	6,0	22,4	35,5	44,2	2,7	3,2	4,8		31,4	37,3		12,0	
Employees' gross monthly earnings	**PY200g**	0,9						1,5											
Equivalised mean income by household size (unit: household)																			
1 household member	**HX090**	1,5	2,0	2,7	1,7	2,4	5,3	3,5	5,6	5,8	1,3	1,6	2,2		4,3	4,6		3,6	
2 household members	**HX090**	1,4	2,0	2,3	1,1	1,5	1,9	3,5	6,0	6,7	1,2	1,4	2,0		3,1	2,7		3,2	
3 household members	**HX090**	1,9	2,2	2,8	1,6	2,2	2,7	2,2	2,7	3,4	1,3	1,5	1,8		3,3	3,8		3,6	
4 and more	**HX090**	1,4	1,6	2,2	1,5	1,7	2,1	1,8	2,2	2,3	1,3	1,5	1,4		2,0	2,2		3,0	
all households	**HX090**	0,8	1,0	1,3	0,9	1,1	2,1	1,4	2,6	3,0	0,7	0,8	1,2		1,9	1,7		1,7	
Equivalised mean income by age class and by gender (unit: person)																			
<25	**HX090**	1,2	1,3	1,8	1,3	1,4	1,8	1,5	1,7	2,1	1,1	1,0	1,2		1,9	2,2		2,2	
25 to 34	**HX090**	1,6	2,0	2,6	1,7	2,1	2,9	2,1	2,7	2,8	1,2	1,4	1,6		3,7	3,9		3,7	
35 to 44	**HX090**	1,6	2,1	2,6	1,6	1,8	5,0	2,0	2,2	2,5	1,3	1,5	1,4		2,5	2,7		2,7	
45 to 54	**HX090**	1,3	1,9	2,5	1,4	1,5	2,3	3,2	2,9	3,2	1,3	1,7	1,3		2,1	2,3		2,7	
55 to 64	**HX090**	1,5	1,5	2,4	1,5	2,3	2,5	4,9	10,5	10,7	1,7	2,2	3,1		3,2	2,6		3,1	
65+	**HX090**	1,3	1,7	2,6	1,4	1,4	1,7	2,5	4,2	4,2	1,3	1,8	2,6		1,5	1,6		2,1	
Male	**HX090**	0,9	1,1	1,5	0,9	1,1	2,0	1,4	2,0	2,3	0,8	0,8	1,0		1,9	1,9		1,7	
Female	**HX090**	0,8	1,0	1,3	0,8	1,0	1,3	1,4	2,4	2,5	0,8	0,8	1,2		1,5	1,5		1,7	
all persons	**HX090**	0,8	1,0	1,3	0,8	0,9	1,5	1,3	2,1	2,3	0,7	0,7	1,0		1,5	1,5		1,6	

* Estonia (1) not computed

** Lithuania (1) not computed; (3) survey stated only in 2005.

*** Tax adjustment: negative sign indicates negative mean value of the variable

2.3. **Non-sampling errors**

Commission Regulation (EC) No 28/2004, Annex III, specifies the information on non-sampling errors which should be presented in national final quality reports. These cover a description and provision of numerical indices where possible on various types of non-sampling errors, including the following.

(1) Sampling frame and coverage errors, including a description of the main coverage problems and procedures for updating the sampling frame.

(2) Measurement errors, including a description of different sources, procedures of questionnaire development and interviewing, and special studies undertaken.

(3) Processing errors, including a description of data entry, coding and editing control, and on the extent of errors found and corrected in particular concerning income variables.

Any methodological studies undertaken in order to assess the magnitude or impact of response and processing errors should be reported.

(4) Unit non-response and achieved sample size, including standardised computation of response and non-response rates at various stages of the data collection process, substitution of sample cases if allowed, and the achieved sample size for household and personal interviews.

Both cross-sectional as well as various longitudinal rates of unit non-response are required.

(5) Item non-response, including for each income component collected or compiled at the household/personal level, the proportions of households/persons receiving and reporting the amount received, reporting it partially, and not reporting the amount; the same for the common cross-sectional EU indicators computed from the cross-sectional data.

In the context of item non-response, information is required on the procedures and extent of imputation, as well as on the net-to-gross conversion of income components.

The objective of this section is to highlight some main results on non-sampling errors in EU-SILC surveys from a comparative perspective. Further information on the basic methodology and approach followed in the production of this information and in its presentation in the national final quality reports can be found in the annex.

<u>Sampling frames used in EU-SILC surveys</u>

The following table shows the type of units used and frame characteristics in EU-SILC surveys. Almost all surveys used a single-stage or a two-stage design.

In multi-stage designs, the whole country is divided into area units such as localities or census enumeration areas (EAs), and a sample of these areas are selected at the first stage. The type of units selected at the first stage is called *primary sampling units* (PSUs). In a two-stage design, in each selected PSU, ultimate sampling units (USUs), which may be dwellings, households or persons, are selected from each sample PSU. In the survey, information may be collected and analysed for the USUs themselves; or for other types of units ('elements') associated with the selected USUs, such as individual persons within sample households, or

conversely, in some EU-SILC surveys household associated with selected individuals. Selecting multiple 'elements' associated with a single USU (taking all households within each selected dwelling, or all persons in a selected household, etc.) is a very common design.

The converse design is much less common: selecting a sample of individual persons, and then taking into the sample the household and all members of the household of the selected person.

In single-stage designs, lists are required for the USUs covering the whole country. The requirement of coverage is more stringent here than in multi-stage designs where the lists of USUs within the selected areas can be updated more readily.

It is common to use both a single-stage and a multistage sampling in different part of the country. For example, two types of designs may be used: while normally the PSUs may be localities and USUs dwellings in a two-stage design, larger localities may be taken into the sample automatically followed by a single-stage selection of dwellings.

As shown in the table "Type of sampling units and the sampling frame (2006)", countries in the 2006 EU-SILC operation have used different sources for lists. Two main groups are: those using population register; and those using census lists and other sources. Then there are a small number of countries which base EU-SILC on successfully interviewed units in another larger survey.

Registers

Generally, where used, the population registers are believed to be up-to-date, assuming that any modification in the population (both people moving in and people moving out) are reported as quickly as possible. Normally in the countries using registers, sample of persons are selected directly. These form the units for the personal interview, while information on income is complied from registers for the whole household of the selected person. There are some variations, however. For example in Finland, "the sample is drawn from the Population Information System maintained by the Population Register Centre of Finland. The register is a continuously updated population register based on domicile. It is updated daily with information on population changes […]". However, unlike many 'register' countries, "the sampling units are dwellings. Persons aged 16+ are selected in the Population Register and then, on the basis of domicile code, their dwellings are eligible for inclusion in the Master Sample. […] Because they have a specific domicile code, homeless people and people living in institutions can be separated before selecting the Master Sample". In Belgium "the sampling frame is the Central Population Register. This Register includes all private households and their current members residing in the territory. Persons living in collective households and in institutions are excluded from the target population". In Sweden "every year a systematic sample is drawn from the register of total population (TPR). This is sorted by age and covers the entire population according to the national registration".

Some countries use multiple frames for different parts of the sample. An elaborate example is provided by Norway.

Census and other sources

When census and other sources are used for lists, it is essential that the databases are updated so as to represent the units which have come into being after the Census and thus ensure that the sample is representative.

In Hungary for example, "the frame is an updated dataset of addresses used in the 2001 population and housing census, thus the under-coverage is due to the new building completed after the last updating". As another example, in Cyprus "the Electricity Authority of Cyprus (EAC) provided a list of domestic electricity consumers, which contained all the new connections of electricity between 2001 and 2005. […] It has been established that each domestic electricity consumer registered by the EAC corresponds to the statistical definition of a housing unit".

In Greece, it is noted that "the dwellings in each newly selected Census area are enumerated just before the fieldwork, so coverage errors ought to be minor". But in some cases, the updated may be limited, for example, for Portugal, the national quality report notes that since its constitution, "the Master Sample was updated two times in small fractions, mainly in exhausted areas".

In France, "in order to represent the dwellings which came into being after the 1999 Census, the so-called new dwellings, the BSLN (Base de Sondage de Logements Neufs) was used together with the 1999 Census". Vacant or secondary dwellings at the 1999 Census had to be included in the sampling frame, due to the important time lag between the Census and the sample selection.

Similarly, in the United Kingdom, households are sampled from the small users Postcode Address File (PAF). This is an up to date list of all addresses maintained by the UK Post Office. The Postcode address file is ordered by postcode sector, which are similar in size to a UK electoral ward.

The following table also shows the last update of the frame as reported in the national quality reports. Unfortunately, information on frame updating has not been provided in some national quality reports, and in some of these the updates may be rather limited.

The last column of the table shows the percentage of listings which were found to be 'blank', i.e. it does not represent any actual household. The amount of blanks can depend on the nature of the list frame, but more likely it reflects the quality (freshness) of the lists: a high proportion of blanks indicates that the frame has not been updated to incorporate changes in the target population which it is supposed to represent. This information is not available for all countries in their quality reports.

Use of respondents to previous (larger) surveys

This can be economical but is likely to increase bias in the sample obtained. Under-coverage comes not only from that which may already exist in the 'parent' sample but also from non-response in the preceding survey. Non-response is usually selective.

Examples include The Netherlands, where the EU-SILC sample has been selected from the subsample of the responding addresses to Labour Force Survey that are willing to participate

to EU-SILC. This is likely to be a serious source of bias. More recently, Statistics Netherlands has focused on an increased use of register data instead of survey data in the production process of statistical information; by making efficient use of register data, it is possible to improve the accuracy of the statistical information, and, at the same time, to decrease the response burden on households.

In Germany EU-SILC survey is designed as a rotational panel (4 subsamples). The sample hitherto has quota and a random part, the latter gradually replacing the former (the sample 2006 contains 2 random samples and 2 quota samples). Sample frame for the yearly random sampling of a new subsample is an access panel (DSP) – containing former participants of the micro census. The 'access panel' refers to the so-called permanent sample of households ready to co-operate with official statistics that was established in German official statistics in 2004. The households in the DSP are 'recruited' on a voluntary basis and hence do not fully meet the requirements of a proper random sample.

Up to 2005 the sample of the Hungarian EU-SILC survey was also a subsample of another survey, the Income Survey sample which was a subsample of the micro census sample. It was noted that from 2006 this basis was to be changed.

Table 5: Type of sampling units and the sampling frame (2006)

	Type of sampling unit		Sampling frame		
	PSU	**USU**	**source of frame**	**Last update**	**% 'blanks'**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
AT	(single-stage sampling)	Dwellings	Central residence register (ZMR)	31-12-2005	0.9%
CY	(single-stage sampling)	Households	2001 census + supplementary list of new houses	not stated	not stated
DE	(single-stage sampling)	Household	DSP (Subsample of the German microcensus)	not stated	0.0%
DK	(single-stage sampling)	Individuals 16+	Central Population Register (CPR)	Continuously	not stated
EE	(single-stage sampling)	Persons 14+	Population register	Continuously	2.9%
FI	(single-stage sampling)	dwellings	Population register	Continuously	not stated
IS	(single-stage sampling)	Persons 16+	Population register	Continuously	5%
LT	(single-stage sampling)	Persons 16+	Residents register (population register)	Regularly	2.6%
LU	(single-stage sampling)	Tax household	Luxembourg Social Security database (IGSS) + Sample of international civil servants	31-12-2005	19.0%
MT	(single-stage sampling)	Households	Census of Population and Housing 2005 database	November 2005	6.5%
SE	(single-stage sampling)	Persons 16+	TRP (Total Population Register)	not stated	not stated
SK	(single-stage sampling)	Households	2001 Population and Housing Census	2005	not stated
BE	Municipalities (or part thereof in larger ones)	Households	Central Population Register	01-08-2006	not stated
CZ	CEUs- Census enumeration units	Dwelling	Geographical register	Continously	4.4%
EL	Census areas	Dwellings	Population Census	Just before the fieldwork	not stated
ES	Census sections	Dwelling	Municipal Register (population register)	01-04-2005	8.2%
FR	Municipality, or group of them	Dwelling	1999 Census + Sampling frame of new dwellings	End 2005	3.3%
HU	Localities	Dwellings	2001 Population and housing census	not stated	0.7%
IE	Block	Household	not stated	not stated	not stated
IT	Municipalities	Household	Registers of the municipalities	Continuously	3.0%
LV	Census area	Addresses	Population Census 2000 + Population register	Beginning of 2005	4.0%
NL	Municipality	Dwellings	Population register	Not reported	not stated
NO	Municipalities (or groups of)	Persons 16+	1990 Census (FoB90) + Population register	Annualy (1); Monthly (2).	not stated
PL	Enumeration areas	Dwellings	Domestic Territorial Division Register (TERYT)	01/01/2005	6.2%
PT	Area of the 2001 Master Sample	Dwelling	Census of Population and Housing 2001	not stated	4.0%
SI	Clusters of enumeration areas	Persons 16+	Central Register of Population (CRP)	Just before the fieldwork	not stated
UK	postcode sector	Addresses	PAF (Postcode Address File)	not stated	3.6%

Tracing rules for follow-up of the longitudinal sample

Practically all countries have followed the standard tracing rules defined in the Commission Regulation on the subject[6].

The EU-SILC "longitudinal data sets" as distributed to researchers has been structured as follows.

1. Rotation groups which are not in the survey for both of the two most recent years (say, Y and (Y-1)) are excluded. This covers rotation groups dropped prior to year Y, and rotation group introduced for the first time at Y. This constitutes the sample base of the data for each of the two years Y and (Y-1).
2. The data for year (Y-2) are confined to the (normally two of the four) rotation groups which are present in all the three years Y to (Y-2). Similarly, the data for year (Y-3) are confined to rotation groups which are present in all the three years Y to (Y-2). (Normally there is only one such group.)
3. Apart from the above exclusion according to rotation group, the 'longitudinal data set' as distributed to researchers covers the most recent 'n' (up to 4) years, and is composed of the full cross-sectional sample for each of the 'n' years. No additional criteria are used to exclude any units which are not 'longitudinal' in the proper sense of the term as defined above.
4. The standard files D and H for households, and R and P for persons are included for each of the 'n' years. The variables included are of course those in the EU-SILC longitudinal data set, which are not identical to those in the cross-sectional data set.

Obviously, not all units included in this data set are 'longitudinal' in the sense defined. A *longitudinal data set* (consisting of longitudinal units) may be constructed as follows.

Constructing a truly longitudinal data sets

Properly longitudinal samples are identified on the basis of *continuous presence of individual persons in the survey* for the specified number of most recent years, for two most recent years for the 2-year longitudinal sample; three most recent years for the 3-year longitudinal sample, etc.

An "expansion" of the longitudinal sample base is required to ensure inclusion of *whole* households with all their members, as required for computation and analysis of income variables. The final set of units for inclusion in the computation can be identified in terms of the above as follows.

Household level variables (H) for the set of households corresponding to each of the longitudinal individual sets as defined above, i.e. for households containing at least one longitudinal person.

Person-level variables (R) for the set of all persons in each of the above-defined sets of households.

---

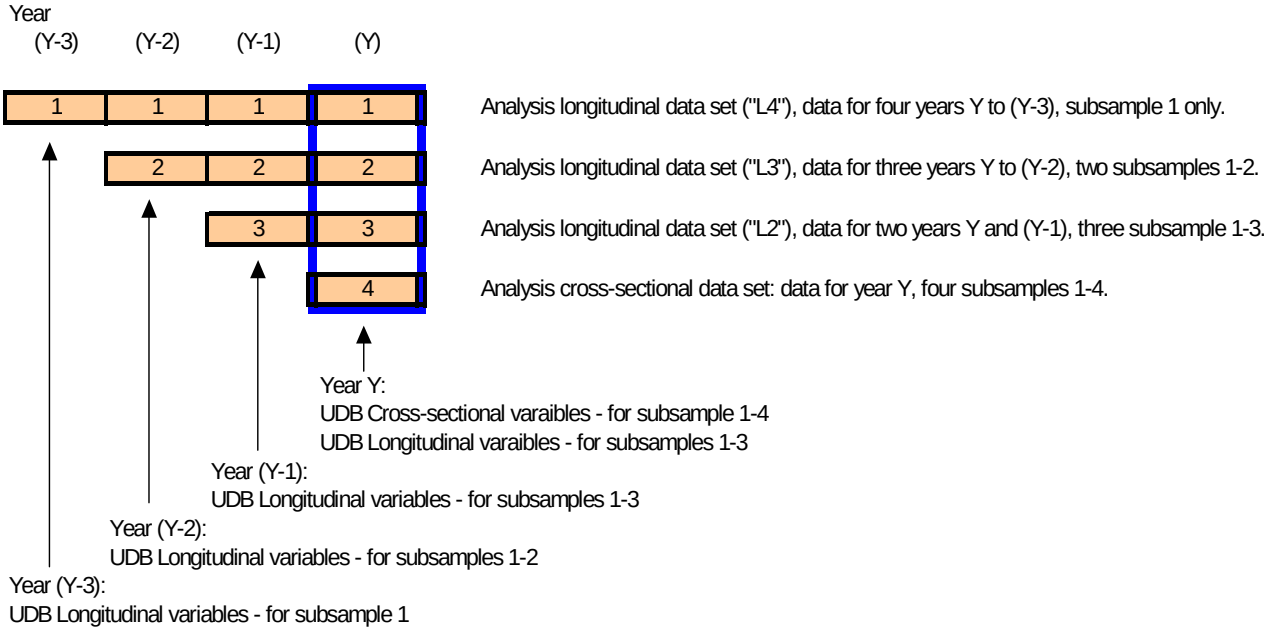
[6] Additional information on tracing rules can be found in section 5.3 of this document.

Adult-level variables (P) for the set of adults in those households; similarly for the subset of variables for selected respondents in those household, if required.

The figure "The UDB and the 'properly longitudinal' data sets" shows the data sets for years Y (current year) to year (Y-3). At the bottom there are shown: (a) one data set of cross-sectional variables for year Y covering subsamples 1-4; and (b) four data sets of longitudinal variables one for each year Y to (Y-3), with the respective sample basis. The sample basis for the last-mentioned four longitudinal data sets are, respectively, subsamples 1-3, subsamples 1-3, subsamples 1-2, and subsample 1. These are the data sets distributed in the UDB. Many of the figures presented below from national quality reports are most likely for this sample base with reference to the current year Y.

On the right hand side of the diagram are listed the four longitudinal data sets for which the figures constructed from the UDB are presented in this report. These data sets are confined to longitudinal units, properly so defined as explained above, the actual data again mostly with reference to the current year Y only. Measures such as longitudinal response rates (see below), for instance, can be meaningfully constructed only for such properly longitudinal data sets.

Figure 1: The UDB and the 'properly longitudinal' data sets



The actual datasets available depend on when the EU-SILC operation was started in the country. The starting dates for different groups of countries are as follows:

2003-2006	EL	LU	NO								
2004-2006	BE	DK	EE	IE	ES	FR	IT	AT	PT	FI	SE	IS
2005-2006	CZ	DE	CY	LV	LT	HU	NL	PL	SI	SK	UK	

The figure below shows for each of the three possible starting year the information included in the 2006 longitudinal data file. The columns represent the concerned subsample (rotational group) and the row the wave.

Figure 2: Data situation for countries stating EU-SILC in different years

	R1	R2	R3	R4	R1	R2	R3
2003				X			
2004				X	X		
2005				X	X	X	
2006				X	X	X	

	R1	R2	R3	R4	R1	R2
2004			X	X		
2005			X	X	X	
2006			X	X	X	

	R1	R2	R3	R4	R1
2005		X	X	X	
2006		X	X	X	

Unit response and non-response rates

Commission Regulation 28/2004 has defined indicators aimed at measuring unit non-response in EU-SILC final quality report as follows. (In addition, breakdown of these rates according to sample status of the person and certain other indicators such as by causes of non-response is also required).

For the first wave of the EU-SILC longitudinal component, the information to be provided includes the following indicators.

- Address contact rate (Ra): the ratio of the number of addresses successfully contacted, to the number of valid addresses selected.

- Household response rate (Rh): the ratio of the number of household interviews completed (and accepted in the data base), to the number of eligible households at the contacted addresses.

- Individual response rate (Rp): the ratio of the number of personal interviews completed (and accepted in the data base), to the number of eligible individuals in completed households.

- Individual non-response rate: NRp= (1-(Rp)

Non-response at the three stages – address contact, household interview and personal interview – is cumulative, so that the overall non-response rates for households and individual interviews are defined, respectively, as follows.

- Overall household interview non-response rate: NRh = 1 – (Ra*Rh)

- Overall personal interview non-response rate: *NRp = 1 – (Ra*Rh*Rp)

For the second and following waves of the EU-SILC longitudinal component, the information provided includes the following indicators:

- Response rate for households = (Ra*Rh)

- Wave response rate: percentage of households successfully interviewed which were passed on to wave t (from wave t-1) or newly created or added during wave t, excluding those out of scope (under the tracing rules) or non-existent.

- Longitudinal follow-up rate: percentage of households which are passed on to wave t+1 for follow-up within the households received into wave t from wave t-1, excluding those out of scope (under the tracing rules) or non-existent.

- Follow-up ratio: number of households passed on from wave t to wave t+1 in comparison to the number of households received for follow-up at wave t from wave t-1.

- Achieved sample size ratio: ratio of the number of households accepted for the database in wave t to the number of households accepted for the database in wave t-1.

- Response rate for persons = (Ra*Rh*Rp)

Cross-sectional response rates (Ra, Rh, Rp) and the corresponding non-response rates

These rates are shown in the next table, distinguishing the new part from the total cross-sectional sample for 2006.

It is clear that the main non-response takes place at the household interview stage. On the average, 97% of selected addresses are successfully contacted; and once a household interview has been completed, 99% of the personal interviews in these households are also successfully completed. But only around 80% of the interviews with contacted households are completed on the average. For the new part of the sample (i.e. the rotation group introduced for the first time), the household interview success rate is considerably lower (73%).

Overall non-response rate for the personal interview, averaged over countries, is as high as 30% for the new sample, and somewhat above 20% for the 'whole' sample (those including the units already in the survey in previous waves and the new units). There is a large variation around this average among the countries, with the non-response rate varying from 5% in Cyprus to 40% in Denmark and Belgium.

The above are the figures for the total cross-sectional sample for 2006. The picture is worse when we consider the new part of the sample, introduced in 2006 for the first time. The overall personal interview non-response rate exceeds 40% in Denmark, Austria and Spain, reaching 53% in Belgium.

The last column of the table shows the percentage points by which the new sample non-response is higher than the total cross-sectional sample. The overall average is 8 percentage points. Big differences are found in Spain (16%), Austria, France, Belgium, Estonia, Lithuania, and Hungary (13%). It is necessary to find out more about the reasons for this, and possible steps which may be taken in order to improve the situation.

No information has been provided in the national quality reports of the United Kingdom and Luxembourg. It is possible that non-response rates are rather high in these countries.

Table 6: Cross-sectional unit non-response rates: comparison of the new sample with the whole sample (Cross-sectional sample 2006)

	Ra		Rh		Rp		Nrh		*NRp	
	W	N	W	N	W	N	W	N	W	N
Belgium	100	100	61	47	99	99	39	53	40	53
Czech Republic	97	96	76	65	100	100	26	38	26	38
Denmark	84	84	72	72	100	100	40	40	40	40
Germany	99	97	78	68	99	99	23	34	24	35
Estonia	92	85	89	81	99	98	18	31	19	33
Ireland	100	100	72	61	100	100	28	39	28	39
Greece	100	100	88	79	99	98	12	21	13	23
Spain	98	97	73	58	98	97	28	44	30	45
France	100	99	84	71	99	98	16	30	17	31
Italy	99	98	86	82	100	100	15	20	15	20
Cyprus	100	99	95	92	100	100	5	9	5	9
Latvia	99	97	79	71	99	99	22	31	23	32
Lithuania	100	99	80	68	100	100	20	33	20	33
Luxembourg	94		75		100		30		30	
Hungary	99	99	83	70	100	100	18	31	18	31
Malta	96	92	90	86	100	100	14	21	14	21
The Netherlands	97	99	83	78	100	100	19	23	19	23
Austria	100	99	72	59	100	100	28	42	28	42
Poland	100	99	87	76	95	95	13	25	17	29
Portugal	98	98	88	82	100	100	14	20	14	20
Slovenia	99	98	79	74	100	100	22	27	22	27
Slovakia	91	100	94	86	99	100	14	14	15	14
Finland	100	100	83	75	100	100	17	25	17	25
Sweden	91	91	81	81	100	100	26	26	26	26
United Kingdom	100		77		100		23		23	
Iceland	100	100	73	71	100	100	27	29	27	29
Norway	99	99	69	64	100	100	32	36	32	36

Respondent status and data status

The following tables show the distribution of all persons by 'respondent status' (variable RB245), and of persons aged 16+ by 'data status' (variable RB250).

The first table gives the results for the 3-year panel, and the second table for the 2-year panel. The figures are computed from the micro databases longitudinal 2006.

The respondent status distinguishes between children (aged below 16) and adults; and among the latter in 'register countries', between the selected respondent (one per household) and other adults. The data status refers to whether the information on income has been obtained for the person concerned. Note that this index does not take into account the overall level of non-

response in the individual interview, but only of individual interview non-response within *interviewed* households. Consequently, the overall mean percentage in the category ('information completed only from the interview (RB250=11) is very high (98%), limiting the usefulness of this indicator for identifying the level of unit non-response[7]. The only disturbing case is that of the UK, where a very low rate (80%) is reported for the 2-year panel.

In principle, however, it should be useful to have this indicator, given that these variables are involved in the computation of the unit non response indicator.

Note also that for register countries, this indicator always has all the cases in category RB250='13' since no non-response in the collection or compiling of income and other personal variables is possible within completed households – by definition, a household is considered 'completed' only if the selected respondent within it has been successfully interviewed.

Table 7: Distribution according to respondent status and data status (2006 Longitudinal Sample, 3-year panel)

RB245: Respondent status (all persons). Percent distribution

	1	2	3	4	total	Number
BE	79.6	.	.	20.4	100.0	4,055
DK	.	41.8	37.3	20.9	100.0	4,780
EE	84.7	.	.	15.3	100.0	9,028
IE	77.1	.	.	22.9	100.0	3,286
EL	85.0	.	.	15.0	100.0	6,447
ES	84.4	.	.	15.6	100.0	15,167
FR	80.2	.	.	19.8	100.0	15,429
IT	85.6	.	.	14.4	100.0	23,414
LU	79.5	.	.	20.5	100.0	7,962
AT	81.2	.	.	18.8	100.0	5,102
PT	85.6	.	.	14.4	100.0	5,824
FI	.	40.9	38.9	20.2	100.0	8,265
SE	.	41.8	38.2	20.0	100.0	6,914
IS	.	36.3	39.6	24.1	100.0	3,099
NO	.	41.7	36.4	21.9	100.0	6,607

RB250: Data Status (persons aged 16+). Percent distribution

	11	12	13	14	21	22	23	31	32	33	total	Number
BE	99.4	.	.	0.5	.	.	.	.	.	0.0	100.0	3,228
DK	.	.	100.0	.	.	.	.	.	.	.	100.0	3,779
EE	98.9	.	.	.	0.1	.	0.7	0.2	0.1	.	100.0	7,650
IE	16.3	.	83.7	.	.	.	.	.	.	.	100.0	2,532
EL	99.6	.	.	.	0.0	.	0.1	0.3	.	.	100.0	5,480
ES	98.0	.	.	.	0.0	.	0.7	0.3	.	0.9	100.0	12,795
FR	99.0	.	.	.	.	.	0.5	.	0.5	0.0	100.0	12,374
IT	100.0	.	.	.	.	.	.	.	.	.	100.0	20,035
LU	100.0	.	.	.	.	.	.	.	.	.	100.0	6,327
AT	99.7	.	.	0.3	.	.	.	.	.	.	100.0	4,141
PT	99.5	.	.	.	0.1	0.0	0.1	0.2	0.1	.	100.0	4,985
FI	.	.	100.0	.	.	.	.	.	.	.	100.0	6,593
SE	.	.	100.0	.	.	.	.	.	.	.	100.0	5,529
IS	.	0.2	99.8	.	.	.	.	.	.	.	100.0	2,352
NO	.	1.2	98.8	.	.	.	.	.	.	.	100.0	5,159

RB245: Respondent status

All household members aged 16 and over are interviewed

1 current household member aged 16 and over

Only selected household member aged 16 and over is interviewed

2 selected respondent

3 not selected respondent

Households members aged less than 16 at the time of interview

4 not eligible person

RB250: Data Status

Information or interview completed

11 information completed only from interview

12 information completed only from registers

13 information completed from both: interview and registers

14 information completed from full-record imputation

Interview not completed, though contact made

21 individual unable to respond (illness, incapacity, etc) and no proxy possible

22 failed to return self-completed questionnaire

23 refusal to co-operate

Individual not contacted because

31 person temporarily away and no proxy possible

32 no contact for other reasons

Information or interview not completed

33 information not completed: reason unknown

Source: UDB Longitudinal sample. 3 waves.

[7] The mean over countries excluded the register countries, as well as Ireland and The Netherlands, where all or

Table 8: Distribution according to respondent status and data status (2006 Longitudinal Sample, 2-year panel)

RB245: Respondent status (all persons). Percent distribution

	1	2	3	4	total	Number
BE	80.0	.	.	20.0	100.0	8,077
CZ	84.7	.	.	15.3	100.0	9,012
DK	.	41.3	37.6	21.2	100.0	7,782
DE	81.8	.	.	18.2	100.0	22,392
EE	84.6	.	.	15.4	100.0	10,674
IE	77.9	.	.	22.1	100.0	5,708
EL	84.6	.	.	15.4	100.0	10,016
ES	84.3	.	.	15.7	100.0	23,339
FR	79.3	.	.	20.7	100.0	19,140
IT	85.3	.	.	14.7	100.0	36,962
CY	79.4	.	.	20.6	100.0	7,960
LV	85.0	.	.	15.0	100.0	7,071
LT	83.7	.	.	16.3	100.0	7,785
LU	78.6	.	.	21.4	100.0	8,068
HU	83.3	.	.	16.7	100.0	12,901
NL	.	38.7	36.3	25.0	100.0	17,164
AT	81.4	.	.	18.6	100.0	9,183
PL	81.9	.	.	18.1	100.0	32,088
PT	85.4	.	.	14.6	100.0	8,701
SI	.	30.9	56.4	12.7	100.0	21,299
SK	84.9	.	.	15.1	100.0	10,863
FI	.	40.3	39.4	20.2	100.0	12,697
SE	.	41.8	38.5	19.7	100.0	10,597
UK	79.9	.	.	20.1	100.0	15,064
IS	.	35.9	40.3	23.7	100.0	4,842
NO	.	40.3	36.8	22.8	100.0	7,102

RB250: Data Status (persons aged 16+). Percent distribution

	11	12	13	14	21	22	23	31	32	33	total	Number
BE	99.5	.	.	0.5	.	.	.	.	.	0.0	100.0	6,460
CZ	100.0	.	.	.	.	.	.	.	.	.	100.0	7,637
DK	.	.	100.0	.	.	.	.	.	.	.	100.0	6,132
DE	99.8	.	.	.	0.0	.	0.2	0.0	0.0	.	100.0	18,324
EE	98.9	.	.	.	0.1	.	0.7	0.2	0.1	.	100.0	9,030
IE	17.6	.	82.4	.	.	.	.	.	.	.	100.0	4,445
EL	99.6	.	.	.	0.0	.	0.1	0.4	.	.	100.0	8,476
ES	97.9	.	.	.	0.1	.	0.8	0.3	.	0.9	100.0	19,671
FR	99.1	.	.	.	.	.	0.4	.	0.5	0.0	100.0	15,170
IT	100.0	.	.	.	.	.	.	.	.	.	100.0	31,521
CY	99.8	.	.	0.2	.	.	.	.	.	.	100.0	6,320
LV	98.6	.	.	.	0.0	.	0.5	0.8	0.1	0.0	100.0	6,009
LT	99.8	.	.	0.1	.	.	0.1	0.0	.	.	100.0	6,518
LU	100.0	.	.	.	.	.	.	.	.	.	100.0	6,340
HU	100.0	.	.	.	.	.	.	.	.	.	100.0	10,751
NL	0.1	.	99.9	.	.	.	.	.	.	.	100.0	12,869
AT	99.7	.	.	0.4	.	.	.	.	.	.	100.0	7,473
PL	95.6	.	.	.	0.3	.	2.2	1.7	0.2	0.0	100.0	26,290
PT	99.5	.	.	.	0.2	0.0	0.1	0.2	0.1	.	100.0	7,427
SI	.	64.6	35.4	.	.	.	.	.	.	.	100.0	18,599
SK	99.8	.	.	.	.	.	0.1	0.1	0.0	.	100.0	9,225
FI	.	.	100.0	.	.	.	.	.	.	.	100.0	10,128
SE	.	.	100.0	.	.	.	.	.	.	.	100.0	8,513
UK	100.0	.	.	.	.	.	.	.	.	.	100.0	12,036
IS	.	1.1	98.9	.	.	.	.	.	.	.	100.0	3,693
NO	.	1.4	98.6	.	.	.	.	.	.	0.0	100.0	5,480

RB245: Respondent status

All household members aged 16 and over are interviewed

1 current household member aged 16 and over

Only selected household member aged 16 and over is interviewed

2 selected respondent

3 not selected respondent

Households members aged less than 16 at the time of interview

4 not eligible person

RB250: Data Status

Information or interview completed

11 information completed only from interview

12 information completed only from registers

13 information completed from both: interview and registers

14 information completed from full-record imputation

Interview not completed, though contact made

21 individual unable to respond (illness, incapacity, etc) and no proxy possible

22 failed to return self-completed questionnaire

23 refusal to co-operate

Individual not contacted because

31 person temporarily away and no proxy possible

32 no contact for other reasons

Information or interview not completed

33 information not completed: reason unknown

Source: UDB Longitudinal sample. 2 waves.

Longitudinal response and follow-up rates

A selection of the longitudinal response and follow-up rates defined before are presented in the following table.

most of the information was collected from registers and interviews together (RB250='13').

Table 9: Longitudinal response rates (waves 2004-2006)

	AT	BE	EE	FR	GR	PT	CY	ES	NO	LU	CZ	LV	NL	PL	SI	UK
Cross-sectional 2006 (1)																
*NRp	28,0	39,6	18,9	16,8	12,9	13,8	5,0	29,9	31,7	29,5	26,3	22,6	19,5	17,4	21,8	23,0
Ra	100,0	100,0	92,0	100,0	100,0	98,0	100,0	98,0	99,0	94,0	97,0	99,0	97,0	100,0	99,0	100,0
Rh	72,0	61,0	89,0	84,0	88,0	88,0	95,0	73,0	69,0	75,0	76,0	79,0	83,0	87,0	79,0	77,0
Rp	100,0	99,0	99,0	99,0	99,0	100,0	100,0	98,0	100,0	100,0	100,0	99,0	100,0	95,0	100,0	100,0
Longitudinal 2005-2006 (2)																
Response rate for households																
Wave response rate	66,8	57,0	85,3	93,1		93,0	94,0	78,3	91,5	68,3	87,4	79,3	77,6	89,0	79,0	73,5
Longitudinal follow-up rate	78,7	87,1	89,6	96,1	93,2		95,3	84,9	92,6		90,0	83,4	87,0		81,0	78,5
Achieved sample size ratio	68,8	158,0	88,1	92,9	93,0	93,0	96,9	81,0	91,5	101,0	88,5	79,9	107,0	89,0	79,0	75,0
Response rate for persons																
Longitudinal follow-up rate	67,1	65,8	86,7	100,0	99,6	97,0	99,8	97,5			99,0	98,6		95,6	100,0	71,3
Achieved sample size ratio	67,5	98,3	90,4	93,3	93,7	92,0	95,0	91,5		98,9				90,7	79,0	72,1
Longitudinal 2004-2005 (2)																
Response rate for households																
Wave response rate	79,4	60,0	90,5	92,3		93,0										
Longitudinal follow-up rate	83,6	83,0	92,7	95,5	93,8	87,0										
Achieved sample size ratio	95,6	111,0	114,0	92,1	86,0	95,0										
Response rate for persons																
Longitudinal follow-up rate	86,3	98,8	89,7	100,0	99,3	98,0										
Achieved sample size ratio	87,4	112,0	112,0	90,9	85,7	95,0										

Wave response rate

Percentage of sample persons successfully interviewed among those passed on to wave t (from wave t-1) or newly created or added during wave t, excluding those out of scope.

Longitudinal follow-up rate

Percentage of sample persons successfully interviewed in wave t out of all of sample persons selected, excluding those who have died or found out of scope.

Achieved sample size ratio

Ratio of the number of completed personal interviews in wave t to the number of completed personal interviews in wave t-1.

Sources:

(1) Comparative Intermediate EU Quality Report 2006 (Version 3 – July 2008)

(2) National Quality Reports.

The calculation of the response/non response rates is more complex for the longitudinal sample because it is necessary to deal with a dynamic picture as survey units change over time; some units cease to exist while new ones are created. This is particularly true for households which are much more transient units than individual persons.

The ratios of achieved sample size can exceed 100% because of the size of the new panel introduced each wave. See the cases of Belgium, Estonia, Luxembourg and The Netherlands, for instance.

Figures in the table have been compiled from national quality reports and earlier comparative quality reports issued by Eurostat. For a number of countries (e.g., DK, FI, IS, IT, HU, DE, SK, LT) the national quality reports do not include the required information, and these countries are not shown in the table. See annex for an outline of the computational procedure and the variables required.

Substitution

It is not the normal practice in EU-SILC to permit substitution for sample cases which cannot be enumerated successfully. According to EU-SILC Regulations, substitutions may be permissible in wave 1 when the sample is originally selected, but not in subsequent waves.

However, three countries have reported the use of substitution in the quality reports: Ireland, Spain and Portugal. Information has not been provided on the percentage of cases substituted, except in the case of Spain (35%). Two basic items of information about the substitution procedure are as follows:

	Ireland	Spain	Portugal
Source of substitute units (substitute chosen from the same …	Block	PSU	Master Sample area
Characteristics controlled in substitutions	NUTS2	PSU	n/a

Some further details provided in the national quality reports are noted below.

Ireland

> In Ireland, lack of information on the substitution operation and on its probable impact is an important shortcoming of the quality reporting. It is mentioned, however, that the country intended "to undertake an exercise to compare the main characteristics of 100 substitute households with 100 original units using census data".
> The following procedure is noted in a later report. "The second sampling stage involved the random selection of four independent samples of one original and three substitute households for each survey area. […] The original sample household constituted the quota of co-operating households to be realised in each survey area and the interviewers systematically approached as many substitute households as was necessary to realise their quotas. In this fashion, variations in response by region and town size were controlled".

Spain

"The new sample is made of 4004 households. 1752 of them were failed to contact and 1385 of these 1752 were substituted. Finally, the percentage of substituted households in the sample is 1385/4004 = 35%. […] In each section, besides the eight addresses selected originally, a further eight were selected as substitutes in case any problem arose with the addresses chosen originally". The following procedure is noted in a later report. "The common variable of an address selected originally and its prospective substitute is the census section. There is not other common variable. There have been multiple substitutions in the sense that further substitutions (until the list of eight substitutes is completely used) have been made for failed substitutions".

Concerning main characteristics of substituted units compared to the original units, only limited information is available. There are some variables that have been collected using a short questionnaire in field when an original unit has not been accepted, but the non-response rate (among such units) has been very high.

Portugal

No information has been provided in the 2006 final quality report. However, the following procedure is noted in a later report. "In each area of the new panel a set of 3 dwellings were selected to substitute the original ones whenever the interviewer was not able to get a response after implementing every perseverance procedure. Dwellings corresponding to secondary residences, vacant, demolished or used for non residential purposes are not substituted. The substitutes are shown in a sequential way per area. The interviewer selects substitutes using this order or sequence".

Achieved sample size

The first impact of unit non-response is on the achieved sample size. (In addition, of course, unit non-response can introduce bias in the results obtained from the survey.)

The following table[8], left panel, shows the achieved sample size (number of households completed) for the cross-sectional component of 2006 EU-SILC, as required by Commission Regulation 28/2004, and also for the 2-year and 3-year longitudinal components.

Column (1) shows the numbers of household interviews completed in the full cross-sectional sample. Column (2) shows the same number for the 2-year longitudinal sample, and column (3) for the 3-year longitudinal sample where available. Ration (2)/(1) is the ratio of the household sample size for the 2-year panel to that for the total cross-sectional sample. In the standard design with 4 rotational panels, this ratio is expected to be 0.75, or somewhat lower due to panel attrition. For around half of the countries, the ratio is in the range 0.67-0.82. For

[8] In order to calculate the number of households in full cross-sectional sample (A1), the records in H files are counted. The numbers of households in the 2-year and 3-year longitudinal samples (A2 and A3) are computed by counting distinct variable HB030 in H file where households completed interview for last 2 and 3 years respectively. The numbers of persons 16+ in the 2-year or 3-year longitudinal sample who completed a personal interview (B2 and B3) are calculated by counting the records in R-file where RB062 or RB063 is greater than zero and RB250 is equal 11 or 12 or 13.

the remaining countries, it is lower, and for some (Denmark and Norway) it falls below 0.50. This can happen if the size of the new panel introduced in 2006 (which contributes to the cross-sectional sample but not to the longitudinal sample) has been increased to meet the overall cross-sectional minimum sample size requirements. Such adjustment does not restore the longitudinal sample size for the current year, but of course can do so for the following years.

The right panel of the table shows the achieved sample size in terms of personal interviews for the 2-year and 3-year longitudinal components for 2006, and also the ratio (3)/(2). The ratio is particularly low for Belgium, Ireland and Austria. It is high (1.0) for Luxembourg because of the special rotational pattern used in the country. A high ratio means that the size of longitudinal panel is well maintained as the duration of the panel increases.

Table 10: Achieved sample size[9]

	(A) Household interviews					(B) Personal interviews		
	(1)	(2)	(3)	ratio (2)/(1)		(2)	(3)	ratio (3)/(2)
BE	5860	3339	1617	0.57	**BE**	6425	3210	0.50
CZ	7483	3812		0.51	**CZ**	7637		
DK	5711	2091	1042	0.37	**DK**	6132	3779	0.62
DE	13799	9639		0.70	**DE**	18282		
EE	5631	3807	3197	0.68	**EE**	8932	7567	0.85
IE	5836	3145	1355	0.54	**IE**	4445	2532	0.57
EL	5700	3836	2436	0.67	**EL**	8440	5460	0.65
ES	12205	8093	5028	0.66	**ES**	19261	12540	0.65
FR	10036	7827	6104	0.78	**FR**	15036	12252	0.81
IT	21499	14708	9209	0.68	**IT**	31521	20035	0.64
CY	3621	2605		0.72	**CY**	6309		
LV	4315	2848		0.66	**LV**	5924		
LT	4660	2932		0.63	**LT**	6503		
LU	3836	3141	2684	0.82	**LU**	6340	6327	1.00
HU	7722	5026		0.65	**HU**	10751		
MT	3494	2342		0.67	**MT**	5533		
NL	8986	6647		0.74	**NL**	12869		
AT	6028	3761	2023	0.62	**AT**	7447	4130	0.55
PL	14914	10714		0.72	**PL**	25128		
PT	4367	3111	2035	0.71	**PT**	7388	4959	0.67
SI	9478	6581		0.69	**SI**	18599		
SK	5105	3757		0.74	**SK**	9207		
FI	10868	5120	3382		**FI**	10128	6593	0.65
SE	6803	4433	2887	0.65	**SE**	8513	5529	0.65
UK	9902	6522		0.66	**UK**	12036		
IS	2845	1749	1110	0.61	**IS**	3693	2352	0.64
NO	5768	2755	2739	0.48	**NO**	5479	5159	0.94

Source: Computed from the cross-sectional and longitudinal data bases, 2006

(A) Number of household for which an interview is accepted for the database
and (if applicable)
Number of selected respondents who are members of the households who completed a personal interview

(B) Number of persons of 16+ who are members of interviewed households who completed a personal interview.
or (if applicable)
Number of persons of 16+ for whom income data were compiled from registers

(1) full cross-sectional sample

(2) longitudinal sample 2-year duration

(3) longitudinal sample 3-year duration

Information on item-non-response in national final quality reports

The following table summarises the availability of information in national final quality reports.

[9] The longitudinal sample in Finland is a subsample of the cross-sectional sample; not all persons from the cross-sectional sample are selected for the longitudinal interviews.

Up to 2006, countries could report income components in either gross or net form. Some countries report only gross components, others report net and gross at least for some components[10].

In some 'register' countries, all income information is obtained from registers, and there is no item non-response by definition, for example in Sweden and Denmark.

In some other register countries, some small components may come from other sources, and hence subject to item non-response. Total disposable income variables are usually constructed from collected net components or constructed from gross amounts using micro-simulation. This for example is the case in Norway, Slovenia and the Netherlands. In Finland, total disposable income variables HY010 and HY020 are constructed from collected gross income components; HY022 and HY023 are constructed by gross/net conversion of gross income components on the basis of taxation register data (imputing).

In a number of countries - such as Italy, France, Latvia - all income variables at component level are collected and recorded net of taxes and social security contribution at source. Therefore the issue of item non-response does not arise in relation to total gross and gross components at household and personal levels (items 3.5.1.1, 3.1.5.5 and 3.5.1.7 in the table).

In some other countries – such as Cyprus, Lithuania, Malta, Slovakia – all components are reported only gross, so that the issue of item non-response does not arise in relation to net components at household and personal levels (items 3.1.5.6 and 3.5.1.8 in the table).

In some countries, some income components are collected gross, while others as net or as both net and gross. Item non-response can occur in any of the items in the table.

[10] Additional information on the form in which income variables have been obtained and the method used for obtaining them in the required form can be found in section 5.2 and in the annex of this document.

Table 11: Information on item-non-response in national final quality reports

Item non-response. Is the breakdown into full, partial and missing provided?

	HY010	HY020	HY022	HY023	Gross H	Net H	Gross P	Net P	Rot group
Belgium	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	no
Czech Republic	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Denmark	Na	Na	Na	Na	Na	Na	Na	Na	Na
Germany	Yes	Yes	Yes	Yes	Yes	No	Yes	No	No
Estonia	No	No	No	No	No	No	No	No	No
Ireland	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Greece	No	Yes	Yes	Yes	No	Yes	No	Yes	No
Spain	No	Yes	Yes	Yes	No	Yes	No	Yes	No
France	Na	Yes	Yes	Yes	Na	Yes	Na	Yes	No
Italy	Na	Yes	Yes	Yes	Na	Yes	Na	Yes	No
Cyprus	Yes	Yes	Yes	Yes	Yes	Na	Yes	Na	No
Latvia	Na	Yes	Yes	Yes	Na	Yes	Na	Yes	No
Lithuania	Yes	Yes	Yes	Yes	Yes	Na	Yes	Na	No
Luxembourg	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Hungary	Yes	Yes	Yes	Yes	Yes	Na	Yes	Na	No
Malta	Yes	Yes	Yes	Yes	Yes	Na	Yes	Na	No
The Netherlands	Yes	Yes	Yes	Yes	Yes	No	Yes	No	No
Austria	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	**No**
Poland	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Portugal	Na	Yes	Yes	Yes	No	Yes	Yes	Yes	:
Slovenia	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Slovakia	Yes	Yes	Yes	Yes	Yes	Na	Yes	Na	No
Finland	Yes	Yes	Yes	Yes	Yes	Na	Yes	Na	Yes
Sweden	Na	Na	Na	Na	Na	Na	Na	Na	Na
United Kingdom	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Iceland	Yes	Yes	Yes	Yes	Na	Na	Na	Na	No
Norway	Yes	Yes	Yes	No	Yes	Na	Yes	Na	No

Source: National Quality Reports 2006.

	Is the breakdown into full, partial and missing provided for:
HY010	Total household gross income
HY020	Total disposable household income
HY022	Abaove, before social transfers other than old-age and survivors' benefits
HY023	Above, before all social transfers including old-age and survivors' benefits
Gross H	Gross income components at household level
Net H	Net income components at household level
Gross P	Gross income components at personal level
Net P	Net income components at personal level
Rot group	Whether the figures are given by rotation group
Na	Not applicable (i.e. only gross or only net component collected)

Some further information is summarised below from comments in the national quality reports.

Denmark

> "Item non-response is generally very low between 0 and 2 pct. The most striking exception is HS130: Lowest monthly income to make ends meet, where it is 10.7%. Information about income is taken from a register. Against this background, Denmark has no item non-response for income variables."

Greece

> The rates of item non-response are extremely low. The quality report notes the following on why it has not been necessary to report item non-response rates. "For the income variables the initial item non-response was approximately 0.2%. Mostly item

non-response was observed in the self-employment income, however due to the limited percentage of non-response we decided to call back the households and their members in order to get the missing information. Hence, in our final data no items missing are included. Also, no imputation was made in the data as partial information didn't exist". The quality report presents a table where "only the percentages of households (per income components collected or compiled at household level) / persons (per income components collected or compiled at personal level) having received an amount for each income component are presented".

Italy

In Italy, all income variables at component level are net of taxes and social security contribution at source. No gross amount has been constructed before the 2007 survey.

Latvia

In Latvia, no data on gross incomes have been collected in 2005 or in 2006.

Lithuania

Employee cash and near-cash income (PY010), self-employment income (PY050), unemployment benefits (PY090), family/children related allowances (HY050), interest, dividends, profit from capital investments (HY090), income received by people aged under 16 (HY110) were collected in net and/or gross. The remaining variables were collected only in gross.

Malta

Income components were collected gross. Net income at the household level had to be constructed, but not at the level of individual components.

Portugal

"Item non-response is not available for HY020, HY022, HY023 because it corresponds to the sum of various components independently of item non-response pattern". Only PY020 (non-cash employee income) is given gross.

Slovakia

Income components are recorded gross.

Sweden

"Calculations of income variables are based on administrative register data. Imputation procedures are consequently not necessary".

Iceland

Item non-response is not to be found for income variables as they come from registers. All income data are recorded gross at component level.

Norway

Presumably, item non-response is not an issue with income coming from registers.

Data collection errors

Now we consider the specific category 'measurement errors'[11] that affect the process of data collection. Such errors occur when the response provided differs from the real (unknown) value. These errors originate from various sources:

- the questionnaire (effects of the design, content and wording)
- the data collection method (effects of the modes of interviewing)
- the interviewer (effects of the interviewer on the response to a question, including errors of the interviewer)
- the respondents (effects of the respondent on the interpretation of items).

Such errors may be random or they may result in a systematic bias if they are not random. The occurrence of these errors and their effects is almost unavoidable; however, each country can implement various methods and procedures to reduce such errors.

As regard to the original questionnaire, the basis is provided by the EU-SILC regulations and the EU-SILC doc 65 *Description of Target Variables: Cross-sectional and Longitudinal*. Experience from pilot surveys and/or former EU-SILC waves have been used to identify potential sources of problems, such as concerning questionnaire content and wording. In so far as these procedures have now become established, less emphasis is given to the detailed reporting of these aspects in the national quality reports of subsequent years.

Concerning the data collection method, it is expected that computer-assisted interviewing (CAPI or CATI) is useful for reducing measurement problems and facilitating data collection. Another advantage of computer-assisted interviewing is that most of the processing errors (inconsistencies and incompatibilities within a household or within an interview) can be identified and corrected during the interview.

To reduce interviewer effects it remains necessary to provide the interviewers with sufficient training and support measures. These training measures help to ensure that all respondents are interviewed under similar conditions as far as possible.

The respondent error tends to increase by proxy responses. This kind of interviewing can result in biased responses, because the proxy generally takes place in the case of selective categories of persons, for example people in employment or self-employment which are less accessible than retired or unemployed persons. That problem can become much more serious in a complex survey like EU-SILC, with complex content. For instance, EU-SILC collects non-monetary income components (e.g., income from private use of company car…) which are difficult to report by proxy. The same applies of course to subjective and personal questions.

[11] Additional information on "errors in measurement", "measurement biases" and "measurement variance" can be found in the annex.

Processing errors

For countries adopting the CAPI/CATI methods of interviewing, the processing errors due to data entry (from a written to an electronic format) are expected to be minimised.

Checking of data quality is an important part of the post-data-collection editing process. Basic principles of this process are standardisation and transparency, in which all relevant tasks are included in a predefined process and data editing rules are generalized for subgroups to avoid single-case solutions. Transparency of changes made to data has to be ensured by documentation such as program code, copies of data files at various stages, flag variables for the identification of the form of information recorded in the substantive variables, and written documentations and descriptions of all the operations.

The information available on records of processing procedures and errors in national quality reports is limited.

Quality control studies (re-interview, record check studies…)

Special quality-control or evaluation studies were undertaken in a few countries. Here are some examples as reported in national final quality reports for 2006.

Czech Republic

> "The questionnaires were first tested in pilot survey of 600 randomly sampled households (Spring 2004). The pilot project involved 14 future regional co-ordinators of the survey and small group of experienced interviewers (2-3 per region). After this fieldwork test, questionnaire was updated and partly re-designed, with active involvement of the regional staff and the participating interviewers. Together with the questionnaires, detailed interviewers guidelines were developed with binding instructions to all questions".

Hungary

> "After the fieldwork the inspectors called 5% of the households asked about the interviewer (whether the interviewer visited the households, was he/she polite, etc.)".

Poland

> "After the household and individual interview completion the respondents were obliged to answer a few questions concerning interview performance. [...] about three quarters of respondents [...] showed a favourable attitude towards the survey, while about 3% [...] were unwilling towards it. In the interviewers' opinion, in about 88% of questionnaires [...] the quality of non-income data collected could be recognised as good or very good and in 1% - as doubtful. The quality of income data was evaluated as slightly worse, mainly because of item non-response. It should also be pointed out that, in our opinion, the quality of data concerning net income categories is much higher than in the case of gross income. The reason is that non-response to the highest degree affected the information on taxes and social and health insurance contributions".

Portugal

> "An additional questionnaire to evaluate the interviewer's performance was applied by telephone to a sample of 10% respondent households (528 households). 319 accepted to cooperate".

Sweden

> "The EU-SILC data are from 2004 to 2006 through face-to-face interviews. The interview form has been specially designed for this type of survey. Telephone interviews whit computer aid CATI is now currently use as the main way to make interviews and half of the interviews during 2006 was CATI. Experiments with split samples have been carried out. The results indicate very little difference between the two interview methods. Indirect interviews can be a source of errors. Applied on appropriate questions experience says that indirect interviews can be an efficient method to collect information."

2.4. **Mode of data collection**

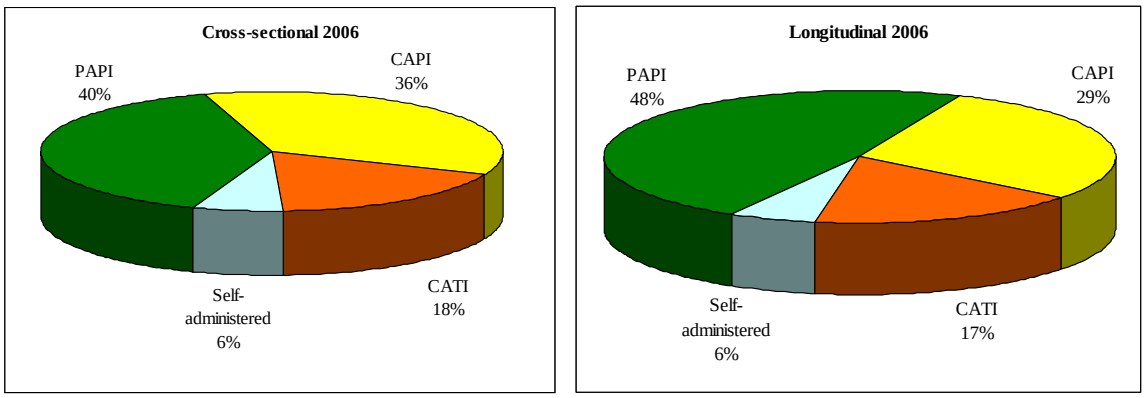
The EU-SILC Regulation allows some degree of flexibility to countries regarding the mode of data collection. The information can be either extracted from registers or collected from interviews. For the interview, four different ways to collect the data are possible:

- Paper-Assisted Personal Interview (PAPI)
- Computer-Assisted Personal Interview (CAPI)
- Computer-Assisted Telephone Interview (CATI)
- Self-administered questionnaire

Countries may use only one method or a combination of various methods. In the EU-SILC legal basis, priority is given to face-to-face personal interviews (PAPI or CAPI) over the other modes of data collection. The following graph represents the different modes of data collection used by the countries for the 2006 operation[12].

[12] Figures are obtained adding up the number of interviews carried out by each mode of data collection by all countries and dividing it by the total of interviews carried out in all countries. The countries are the EU-27 countries except Bulgaria and Romania plus Iceland and Norway. Detailed percentages for each mode of data collection by country for the 2006 operation can be found in the annex.

Figure 3: Mode of data collection (EU27 minus BG, RO plus IS, NO; cross-sectional and longitudinal 2006)



Source: Micro-database (March 2009)

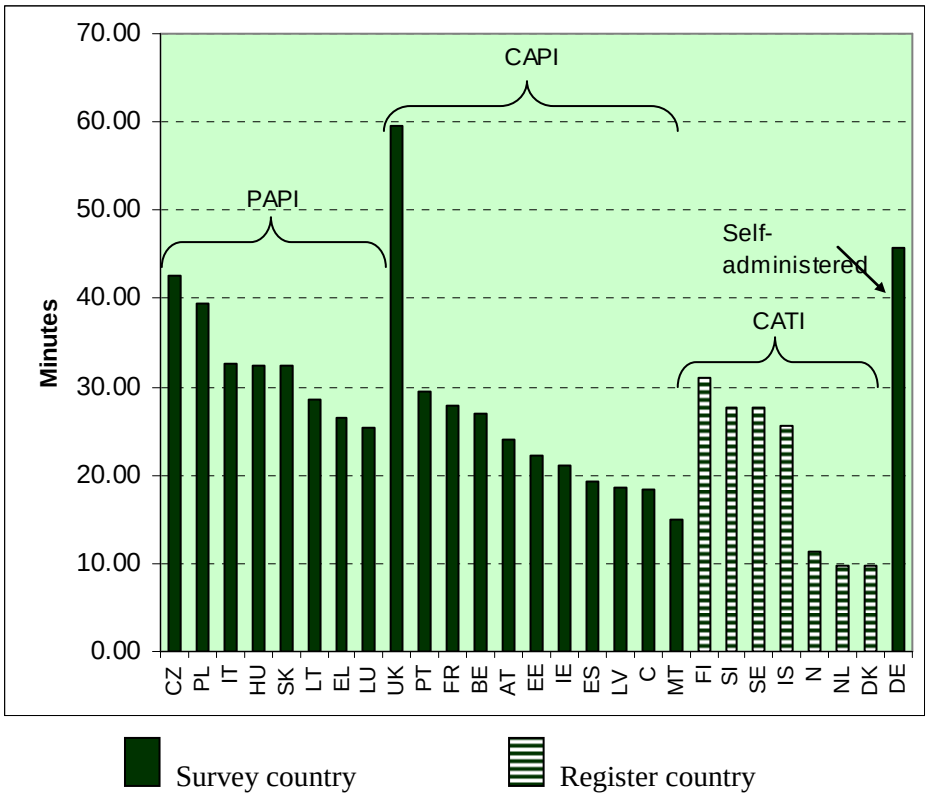
Three main conclusions can be extracted from these graphs: 1. Face to face interviews are used for over 75% of the interviews (either in paper or with a computer); 2. Computer assisted interviews in general, by telephone or in paper, are used in over 45% of the cases. 3. For the cross-sectional component there was an increase in the use of computer assisted interviews compared to the longitudinal component.

The mode of data collection might affect the duration of the interview. Indeed from the following figure we can extract some interesting conclusions combining the information on the mode of data collection by country with the duration of the interview[13]. This graph represents the mean interview duration in minutes calculated as the sum of the duration of all household interviews (HB100) plus the sum of the duration of all personal interviews (PB120), divided by the number of household members aged 16 and over whose household questionnaire is completed and accepted for the database (PB030)[14]. Countries are grouped by the most used mode of data collection.

[13] It should be kept in mind two key inaccuracies in data provided by countries: 1. some countries include national questions in addition to the EU-SILC ones and countries cannot determine the length of the "actual" SILC questionnaire. 2. in some countries the recorded interview duration relates to the minutes elapsed from the first question to the last one, whereas the actual time can be considerably longer.

[14] If the household interview duration (HB100) or one personal interview duration (PB120) is missing for one member of the household, then the household is excluded from the calculation.

Figure 4: Average interview duration per individual (2006 cross-sectional)



Source: Micro-database (August 2009) except for Denmark, Finland and Iceland (figure from the national quality reports).

From this graph we can conclude that survey countries use face to face interviews (PAPI or CAPI) and register countries use telephone interviews. Among the face to face interviews those done with the assistance of a computer are shorter on average.

In the case of United Kingdom, EU-SILC questions are included as part of the General Household Survey questionnaire and there is no information on the interview duration of EU-SILC alone. GHS mean total interview time is 60 minutes.

Proxy interviewing is permitted if the proxy rate is kept as limited as possible. Some countries that encountered rather high non-response rates chose to use proxies to ensure a certain degree of accuracy in their data. In addition, in countries that use the selected respondent type of survey, the household respondent (in most cases selected respondent) is asked for information about all household members, therefore, these countries have a high percentage of proxy interviews concerning personal interviews. The table below presents the percentage of proxies in 2006 (cross-sectional and longitudinal).

Table 12: Percentage of proxy interviews (2006)

	Cross-sectional	Longitudinal
Belgium	13.8	13.9
Czech Republic	8.3	9.1
Denmark	48.7	48.7
Germany	20.6	16.4

	Cross-sectional	Longitudinal
Estonia	5.8	4.0
Ireland	33.3	31.7
Greece	4.0	4.2
Spain	41.2	38.7
France	26.8	26.4
Italy	15.5	16.0
Cyprus	12.7	13.0
Latvia	6.6	6.4
Lithuania	16.5	15.4
Luxembourg	25.3	24.5
Hungary	13.2	11.5
Malta	32.6	31.7
The Netherlands	43.4	42.0
Austria	19.6	19.9
Poland	18.6	19.1
Portugal	13.5	12.9
Slovenia	26.9	26.5
Slovakia	5.9	5.6
Finland	50.8	50.9
Sweden	3.7	4.5
United Kingdom	10.0	9.9
Iceland	0.0	0.0
Norway	29.2	32.7

Source: Micro-database (March 2009).

The percentage of proxy interviews varies greatly among countries[15]. In the register countries, the level of proxies diverge from below 5% in Sweden and Iceland, to around 50% in Denmark, The Netherlands and Finland, and with a level of about 30% in Slovenia and Norway. In survey countries the range of proxies is the following: six countries with a percentage below 10% (Czech Republic, Estonia, Greece, Latvia, Slovakia and United Kingdom), nine countries between 10 and 20% (Belgium, Germany, Italy, Cyprus, Lithuania, Hungary, Austria, Poland and Portugal) and five countries above 20% but below 42% (Ireland, Spain, France, Luxembourg and Malta). There is a need of more detailed information on the reason for the high percentage of proxies in the national quality reports. In the case of Malta the explanation is the following: Since it is very difficult to ensure that all household members will be present for the interview at the same time, proxy interviews are the only alternative to higher non-response rates. In an attempt to minimise the errors that may result from proxy interviewing we encourage respondents who could not be present during the interview (as well as all other respondents) to leave appropriate documentation related to their income (e.g. payslips, tax returns, etc.) with the person who responds on their behalf. Furthermore, in cases when the proxy interviewer is unsure about how to answer certain questions, we instruct interviewers to call the households back at a later date when they can

check about the missing information with the person concerned. These are still recorded as proxy interviews.

2.5. **Imputation procedure**

According with EU-SILC Framework Regulation, "Member States shall transmit to the Commission (Eurostat) in the form of micro-data files weighted cross-sectional and longitudinal data which has been checked, edited and imputed in relation to the income".

Countries should implement imputation procedure for their income variables but flexibility is given to them in order to let them choose the one which is the most appropriate in their case.

The objective of imputation and micro-simulation is to convert full or partial item non-response into more complete values. These procedures are applied primarily to income variables. The completed values are of course still subject to measurement (collection and processing) errors, at least in part as a result of the imputation and modelling involved in their creation. *Thus these processes can be seen as a link between item non-response and measurement errors.*

As regard item non-response, particular attention needs to be paid to the income variables. Missing income data have been dealt by imputation, filling in nearly all missing values by imputed ones. It has to be kept in mind that imputed values are not values actually observed, but are based on some models and assumptions, though trying to make the best use of available data. Imputation can have a significant effect on the overall accuracy; furthermore, variance estimates assuming that imputed values are exact ones will generally be biased.

The impact of imputation on the EU-SILC data is difficult to assess, though some useful information has been provided in the 'imputation flags' which have been constructed for each income variable.

The item non-response is structurally high for some income components which are difficult to collect through interview (capital income, self employment income), or which have not been collected fully because they can be reconstructed using auxiliary information (e.g., child allowance).

In a vast majority of the cases, the Eurostat database provides full income records because countries are asked to impute missing component. However the Commission Regulation requests for income variables, to provide for each wave of the EU-SILC longitudinal component the following information:

- percentages of households (per income components collected or compiled at household level)/persons (per income components collected or compiled at personal level) having received an amount for each income component,
- percentage of missing values for each income component collected or compiled at household/personal level, and
- percentage of partial information for each income component collected or compiled at household/personal level.

The above information could be computable for each income component from the micro database if the imputation flag available for each income component would be able to distinguish the real imputation due to missing data from gross/net conversion. Actually the degree of collected income is controlled through imputation index attached to each value which records the collected amount divided by the recorded amount. The impact of imputation on the data is difficult to assess as the imputation flags are of limited reliability in some cases.

Next table indicates the types of imputation techniques used by countries, as reported in the national quality reports. Some more detailed notes, summarised from the national reports, follow the table.

Table 13: Imputation techniques used

	Mean/median imputation	Regression model	Hot deck	Cold deck	Other methods
Belgium	Y	Y	Y	N	Y
Czech Republic	N	N	Y	N	N
Denmark	not reported/not done				
Germany	Y	Y	N	N	Y
Estonia	Y	N	Y	N	Y
Ireland	not reported/not done				
Greece	not reported/not done				
Spain	N	Y	N	N	N
France	N	Y	N	N	N
Italy	N	N	Y	N	N
Cyprus	N	N	N	N	Y
Latvia	N	N	Y	N	N
Lithuania	N	N	N	N	Y
Luxembourg	N	Y	N	N	Y
Hungary	N	N	N	N	Y
Malta	Y	Y	Y	N	Y
The Netherlands	N	N	N	N	Y
Austria	N	Y	N	N	Y
Poland	N	Y	Y	N	Y
Portugal	N	Y	N	N	N
Slovenia	N	N	N	N	Y
Slovakia	not reported/not done				
Finland	N	N	Y	N	Y
Sweden	not reported/not done				
United Kingdom	N	N	Y	N	Y
Iceland	N	N	N	N	Y
Norway	N	N	N	N	Y

Source: National Quality Reports 2006.

Belgium

There was implicitly more emphasis on the regression model techniques of imputation. The 'conventional' box-plot method was also used, and checks were also made for outliers via controls on a case to case basis. In order to correct and impute data for different variables they relied as much as possible on internal information present in the data itself; and on formal and legal sources of information.

Estonia

Mean/Median imputation was used only when single values were missing. Hot deck (random donor) was used when proportion of missing values was very small. When the exact value was missing but the respondent provided an interval, the values were imputed with hot-deck method within this interval. Other methods used included 'logical deduction of value, based on other data in questionnaire', and random regression using "IVE-ware".

Greece

"No imputation procedure was applied".

Spain

Regression model was used, based on the statistical imputation software "IVE-ware". "For each variable the best regression method was chosen according to the nature of the variable being imputed.

France

Regression model based on different equations taking into account new and old households and type of income imputing was used.

Italy

A hot deck imputation procedure for each quantitative variable was implemented by using the IMPUTE module of the software IVE-ware, as recommended by Eurostat.

Cyprus

"In the very few cases where imputation required, the method used was deductive imputation. Imputation was necessary in the cases where only net income was collected and in the cases of personal refusals (18 cases). Net income was converted to gross by applying the existing tax system and social insurance contributions rules. Personal refusals were imputed using existing data from previous waves as the starting point".

Lithuania

'Other methods' used included: deterministic methods that were used for PY010G, PY050G (mean/median imputation); PY130G, HY090G (distance matching); and deductive methods that were used for HY050G, HY140G (deductive imputation).

Latvia

Hot deck procedure was used for imputing the data on household and personal level. A hot-deck method is used for both imputations procedures. The main principle of the hot deck method is to use the current data (donors) to provide imputed values for records with missing values.

Hungary

'Other methods' used included deterministic method covering the cases, when the missing values could be determined by available background information at the given record. Practically it was used for social incomes and benefits.

Malta

'Other methods' used included the following. Item-non response in essential variables was tackled through estimations by means of auxiliary variables and the use of register information where available.

Austria

Regression model was used for the cross sectional imputation. As to the use of other methods, the longitudinal imputation procedure was based on the row-and-column-method of Little and Su. As suggested by the name, the method uses the row effects and the column effects of the data to identify an appropriate donor case.

Portugal

"Income data is collected can be provided by respondents either in gross values or in net values. The net series was obtained by the application of a specific gross-to-net micro simulation model". Concerning regression models: "The IVE-ware is applied in situations of total absence of data for a specific income variable".

Slovenia

For incomes variable several stages of imputation were used. Other methods used included Hot-deck method (or Nearest Neighbour version) with different imputation cells defined; Trimmed average method with different imputation cells and different trim-threshold defined, and Logical imputations.

Finland

Other methods included the following. For HY030G, the stratification method as a deterministic method was used to impute market rents values to households' equivalent dwellings from an external data source. For HY100N, HY022 and HY023, deductive imputation was used.

United Kingdom

The imputation process was supported by statistical tools and used standard statistical techniques for panel data, including - SAS (Statistical Analysis System) – to facilitate deductive imputation.

Iceland

'Other methods' used included imputation was applied when dealing with amounts or working hours and we knew that these amounts were paid or received but did not have the amount or the number. Not imputing would systematically underestimate the amount".

Norway

'Other methods' used included the following. "In the estimation of HH070, imputations are made on item non-response for items apart from rent and interests on mortgage Insurance Average values based on dwelling size (7 groups based on sqm) are imputed for those with item non-response".

2.6. **Imputed rent**

The imputed rent (HY030) refers to the value that shall be imputed for all households that do not report paying full rent, either because they are owner-occupiers or they live in accommodation rented at a lower price than the market price or because the accommodation is provided rent free.

This variable is only mandatory from 2007 operation, but four countries already provided it in 2006: Estonia and France for the cross-sectional and longitudinal data, and Greece and Finland only for the cross-sectional data. There are other countries collecting the data but not yet transmitting it to Eurostat.

2.7. **Company cars**

The variable PY020 has two parts: 1. "Imputed income from private use of company car", which is compulsory; and 2. "Other non-cash employee income", which is mandatory from 2007.

For 2006 is not possible to make a comparison of this variable among countries because some countries provided the two parts of the variable while others only the mandatory one. Part 1 was collected in all countries with the exception of France and Iceland, while part 2 was not collected by any country with the exception of Lithuania and Finland. In Austria the part 1 of the variable was collected but payments in kind of the private use of a company car were included in PY010. Information by country can be found in the annex, in the table on the individual income components.

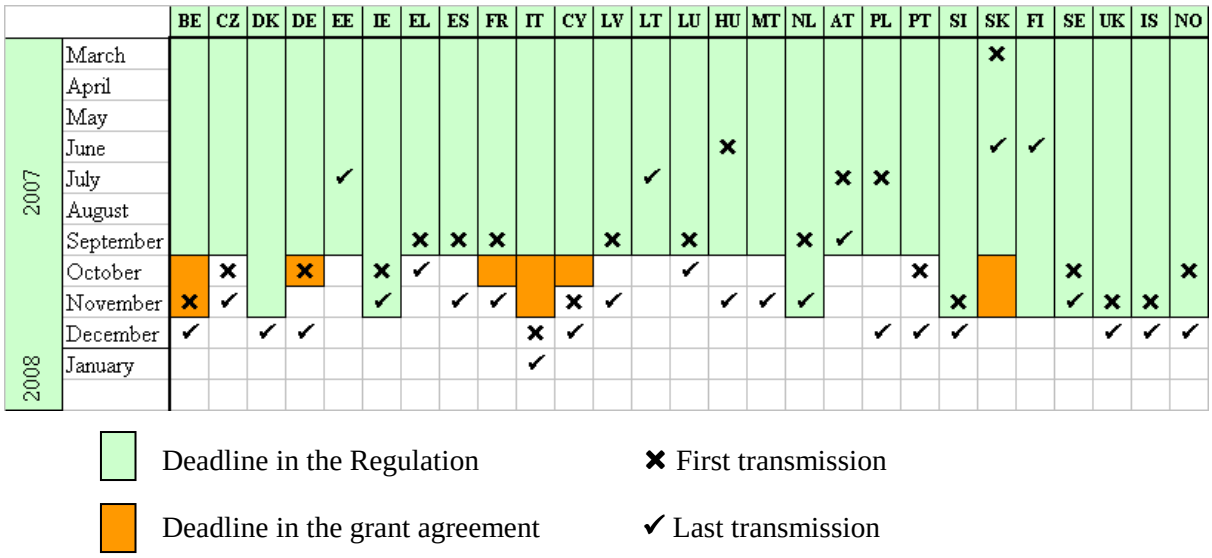
From 2007 operation onwards, we can compare the value of this variable[16].

3. TIMELINESS AND PUNCTUALITY

Cross-sectional data[17]

Regulation (EC) No 1177/2003 states that "(…) The extreme deadline for the transmission of micro-data to Eurostat shall be 30 November (N+1) for Member States where data are collected at the end of year N or through a continuous survey or through registers and 1 October (N+1) for other Member States". Nevertheless, the grant agreements between Eurostat and Member States might have a different deadline. The first cross-sectional microdata for the 2006 operation were received in Eurostat on March 2007. Reception of microdata extended up to January 2008. Key indicators were released on Eurostat website by 17 December 2007 with five exceptions: Belgium, Germany, Sweden and the United Kingdom on the 11 January 2008 and Iceland on the 06 February 2008.

Figure 5: Follow-up microdata (cross-sectional 2006)



Source: eDamis.

The deadline in the grant agreement is only presented for those countries where the date in the agreement is later to the date in the Regulation (EC) No 1177/2003.

Regulation deadlines (or contractual deadlines when there were posterior to the Regulation deadlines) were respected for the majority of the countries, at least for the first transmission of data, with the exception of Czech Republic, Denmark, Italy and Cyprus. Nevertheless, for the reception of the validated version of the data only eight countries (Estonia, Ireland,

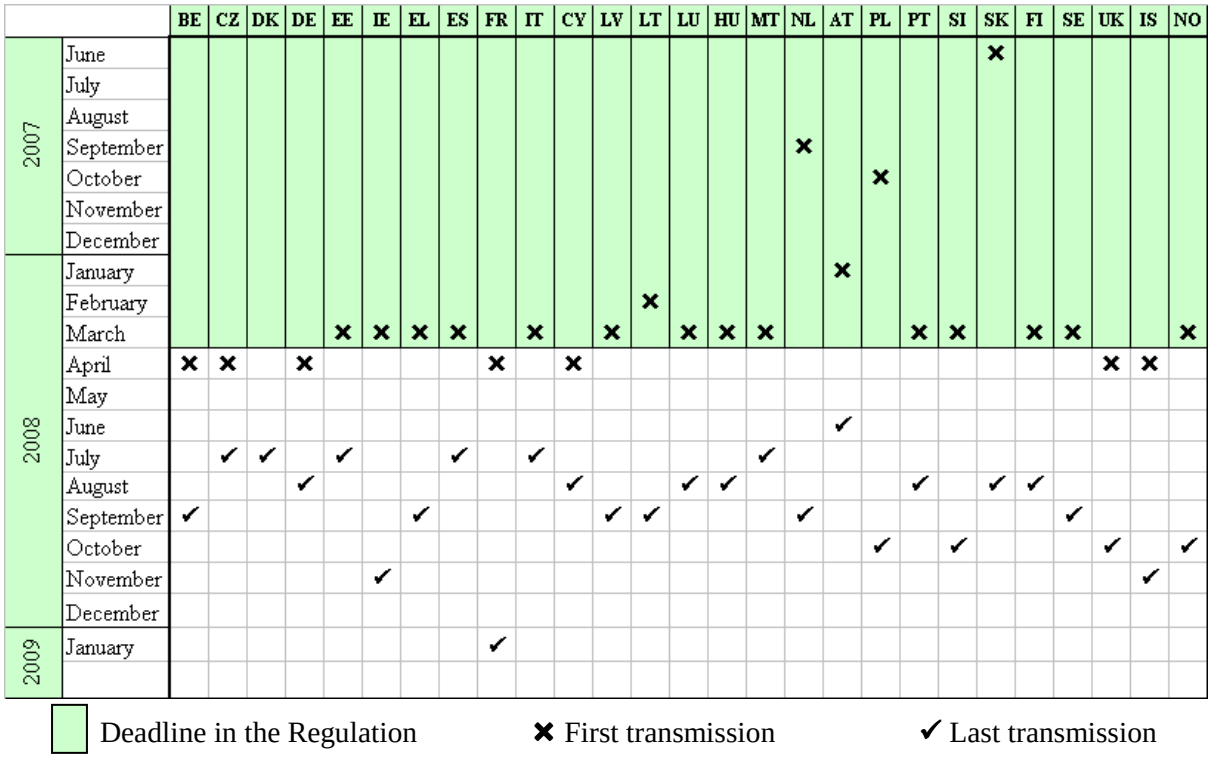
[16] From 2007 on: PY020 refers to "Other non-cash employee income" and PY021 to "Income from private use of company car".
[17] More detailed information on the dates and deadlines for transmission of data and quality reports can be found in the annex.

Lithuania, The Netherlands, Austria, Slovakia, Finland and Sweden) delivered the data before the specified deadline.

Longitudinal data[17]

For the longitudinal component, the Regulation (EC) No 1177/2003 states the following "(…) The mandatory deadline for the transmission of micro-data to Eurostat shall be the end of March (N+2), each year starting from the second year of EU-SILC". Contracts with Member States had different deadlines but all of them were earlier to the one in the Regulation. The first longitudinal microdata for the 2006 operation were received by Eurostat on June 2007. Reception of the first version of the microdata extended up to April 2008.

Figure 6: Follow-up microdata (longitudinal 2006)



Source: eDamis.
Deadline refers to the one presented in the Regulation (EC) No 1177/2003.

The deadline stated in the Regulation was respected, for the first transmission of the data, by the great majority of the countries. There were only seven exceptions (Belgium, Czech Republic, Germany, France, Cyprus, United Kingdom and Iceland) with a delay shorter than three weeks. Nevertheless, the reception of different versions of the data extended for several months for most of the countries, with five transmissions per country on average[18].

[18] Most countries were sending longitudinal data for the first time and they experienced problems mainly related to the weighting procedure.

Quality reports[17]

The deadline established in the Regulation (EC) No 1177/2003 for the transmission of the national final quality reports is end of year N+2 and almost all countries met the deadline. The only exceptions were Belgium, The Netherlands, Finland, United Kingdom and Norway.

Indicators

Regulation (EC) No 1177/2003 states that "Together with the micro-data files, Member States shall transmit social cohesion indicators based on the cross-sectional sample of year N which will be included in the annual Spring report of year (N+2) to the European Council".

Overarching indicators were transmitted in time for inclusion in the Joint Report on Social Cohesion and Social Protection to be released for the spring meeting of the European Council.

4. ACCESSIBILITY AND CLARITY

In accordance with Commission Regulation 831/2002, the Commission has released 2006 SILC anonymized microdata via CD-ROM to researchers. The UDB (User database) with the cross-sectional microdata was sent to countries and contractors[19] in March 2008. Longitudinal microdata were sent to countries and contractors in November 2008. 74 contractors received the 2006 cross-sectional data and 66 the 2006 longitudinal data.

Data aggregated tables in the form of predefined tables or of multidimensional tables are available free of charge on Eurostat website and can be explored via the data navigation tree. Public information on data coding as well as methodological description of EU-SILC is available at http://circa.europa.eu/Public/irc/dsis/eusilc/home. Moreover, in May 2008 Eurostat included a new dedicated section on the website containing key information on Living condition and social protection statistics including information on EU-SILC.

In addition, 2006 data was used in the following publications: Europe in figures. Eurostat yearbook 2008, Pocketbook: Key figures on Europe. 2009 edition, Pocketbook: Living conditions in Europe. Data 2003-06, European Sustainable Development Strategy (SDS) monitoring report, the Social Situation Report, the Joint report and the Annual Progress Report to the Spring Council.

5. COMPARABILITY

Comparability refers to a common set of concepts and definitions that shall be applied by the countries when designing the survey and collecting the data. It encompasses both basic definitions (reference population, private household, household membership…) and income concepts (employee income, self-employment income…).

[19] The term "contractors" includes universities, research institutes and some other bodies.

Commission Regulation 1980/2003 establishes the framework for comparability, which has set out standard definitions as accurately as possible to cover most of the cases that might be encountered in practice. Some degree of flexibility is allowed regarding the definitions but countries have to report on deviations and their estimated impact in the national quality report.

Some countries carry out specific studies on the characteristics of the survey to analyse the impact on comparability.

Czech Republic

The project comprise analytical studies assessing two aspects of the national implementation of the survey: dealing with attrition in the panel component and methodology used for estimation of imputed rent. These two topic areas require the development of methodology for national implementation, using the current best practices from other participating countries, with necessary adaptations to particularities of our national context.

Denmark

The main purpose of the action is to study consequences for comparability:

- of using income data based on administrative registers as an alternative to using data collected by a traditional survey and
- of using register information and a sample of persons instead of a sample of households when selecting and delimiting the households.

Estonia

The action covers:

- the coherence and accuracy analysis of EU-SILC income data;
- the evaluation of an alternative sample selection scheme for EU-SILC and their impact of sample selecting on comparability.

5.1. **Basic concepts and definitions**

Two summary tables on different aspects that can hamper comparability can be found in the annex. A first table covers the adherence/deviation to the standard definition on the reference population, the private household and the household membership. A second table presents the reference period for income, for taxes on income and social insurance contributions and for taxes on wealth. The main conclusions from these tables are the following:

Reference population

Countries do not report any difference with the standard definition in the quality reports.

Private household

Three countries inform about some differences with the standard definition:

- Italy: Cohabitants related through marriage, kinship, affinity, patronage and affection constitute the private household.

- Austria: Private households were generally defined as a person living alone or a group of persons living in the same dwelling. All persons at the dwelling form the household as shared expenses were assumed.

- United Kingdom: A household is defined as a single person or a group of people who have the address as their only or main residence and who either share one meal a day or share the living accommodation. A group of people is not counted as a household solely on the basis of a shared kitchen or bathroom.

Household membership

Some deviations from the standard definition are described by five countries:

- Spain: The quality report provides comparative tables to illustrate the differences between the national and the standard definitions of household membership. In short, the following persons, provided they share the expenses of the household and intend to stay at least 6 months, are not considered as household members in the Spanish SILC (but should be under the EU standard definition) so long as they have another address which they regard as their usual residence: resident boarders, lodgers, tenants, visitors or domestic servants (live-in domestic employees, au-pair).

- Italy: Live-in domestic personal (au pairs) are not included as household members. Concerning these persons, only some socio-demographic information is collected (date of birth, sex, marital status, and duration of stay in the household). The number of these persons included in the sample was 35 (0.1% with respect to the total number of households and 0.06% w.r.t. interviewed individuals).

- Austria: Household membership is described as follows: 1. All Persons who are actually living in the dwelling unit. The question whether these residents have their main residence in this particular dwelling, is not relevant. 2. Lodgers, visitors, au-pairs and guests are considered members of the household if they stay or intend to stay 6 months or more in the household, or if they do not have any other home address. 3. Persons who are temporarily away for less than 6 month and are not members of other private households. 4. Household members who are absent for 6 months or more who are not members of other private households and/or are children or partners of actual household members.

- Portugal: Contrary to the EU-SILC concept, persons absent for long periods, but having household ties (persons working away from home) are not considered as household members if the absence is for more than 6 months (the income obtained from them is considered as a private transfer).

- United Kingdom: A person is in general regarded as living at an address if he or she (or the informant) considers the address to be his or her main residence. There are however, certain rules which take precedent over this criterion. Children aged 16 or over who live away from home for the purposes of either work or study and come home only for holidays

are not included at the parental address under any circumstances. Children of any age away from the home in a temporary job and children under 16 at boarding school are always included in the parental household. Anyone who has been away from the address continuously for 6 months or longer is excluded. Anyone who has been living continuously at the address for 6 months or longer is included even if she has his or her main residence elsewhere. Addresses used only as second homes are never counted as a main residence.

Income reference period

The income reference period for most of countries is the calendar year previous to the survey year, i.e. 2005; with two exceptions: 1. In Ireland the income reference period is the last twelve months. 2. In United Kingdom the current income is annualised and aims to refer the current calendar year, i.e. weekly estimates are multiplied by 52, monthly by 12…

Reference period for taxes on income and social insurance contributions

For almost all countries 2005 was the reference period for taxes on income and social insurance contributions with the exception of those countries that do not collect taxes on income (Italy, Latvia[20] and Portugal) and for Ireland and United Kingdom with reference period similar to the income's one.

To evaluate the income taxes two possibilities are envisaged: A. Income tax paid/received during the income reference period and B. Income tax paid/received related to the total income received during the income reference period. In 2006:

- Fourteen countries followed definition A: Belgium, Estonia, Greece, Spain, France, Cyprus, Lithuania, Luxembourg[21], Hungary, Malta, Austria, Poland, Slovenia and Slovakia.

- Ten countries followed definition B: Czech Republic, Denmark, Germany, Ireland, The Netherlands, Finland, Sweden, United Kingdom, Iceland and Norway.

Reference period for taxes on wealth

Most of the countries used as reference period 2005 with the exception of United Kingdom (April 2006-March 2007). Five countries do not have taxes on wealth: Belgium, Ireland, Malta, The Netherlands and Austria.

Time lag

The lag in months between income reference period and current variables differs from country to country, from Ireland and United Kingdom with no time lag to Sweden with up to 12 months lag.

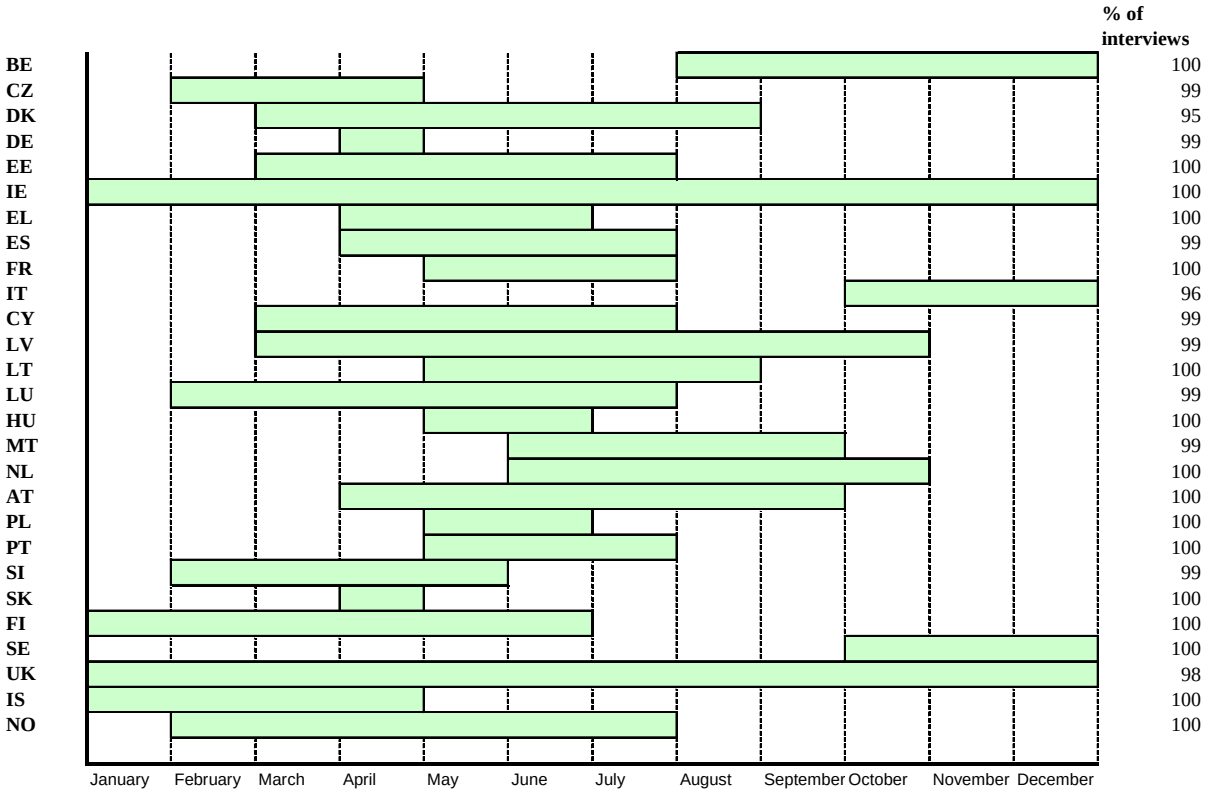
[20] Latvia is authorized to not deliver any gross income data before 2007. Thus, no data on income tax and on social contributions was collected.
[21] When data are imputed taxes only concern the income 2005.

Fieldwork duration

The duration in months of the data collection also varies greatly from country to country: Hungary, Poland, and Slovakia less than three months; Belgium, Czech Republic, Germany, Estonia, Greece, Spain, France, Cyprus, Lithuania, Malta, The Netherlands, Portugal, Slovenia, Finland, Sweden and Iceland between three and six months; Denmark, Italy, Austria, Luxembourg and Norway between six and nine months; and Ireland, Latvia and United Kingdom over nine months.

The following chart summarizes the fieldwork period by country; figures correspond to the information on the month of the household interview (HB050).

Figure 7: Fieldwork period for the 2006 operation



Source: Micro-database (March 2009).
Notes to the figure: (1) Last column presents the percentage of interviews that were carried out in the months presented in this graph by country. (2) In Italy 3% of the interviews were carried out in January, February and March 2007. In United Kingdom 2% of the interviews took place from January to April 2007.

It can be concluded that most of the countries finished the fieldwork period by July, with the exception of: Denmark and Lithuania that finished in August; Latvia, Malta, The Netherlands and Austria that finished by October; Belgium, Ireland and Sweden that completed the fieldwork at the end of the year; and Italy and United Kingdom that ended the fieldwork period in 2007.

5.2. **Components of income**

Regarding the components of income some flexibility has been allowed to the definitions, particularly for taking into account national constraints. Countries report on any differences between the national definitions and the standard EU-SILC definition. Two summary tables by country and income component can be found in the annex, one on household income components and one on personal income components.

A summary table can be found in the annex about the source or procedure used for the collection of income variables. It can be highlighted that 'register countries' used administrative data except in Slovenia and Iceland where interviews were also conducted to collect income variables. Most of the 'survey countries' obtained the information from the interview with only three exceptions, Ireland, Latvia and Lithuania, where registers were also used to collect some income components.

As noted, up to 2006, countries could report income components in either gross or net form; some countries report only gross components, others report net and gross at least for some components.

The following table ("Mode of collection and recording of self-employment income") summarises the form in which one of the income components – namely income from self-employment - has been recorded in the country micro data.

In this table are included all members of the household of the individuals interviewed in both years. Are also consequently included some individuals who were not interviewed in both waves but at least one other person from their household was interviewed.

Column (5) of the table shows the percentage of individuals (aged 16+) who received income from self-employment during the income reference year. The figures are for the 2006 survey year, 2-year longitudinal data set.

Columns (6)-(8) show how the numbers receiving self-employment income are distributed according to the form of recording of the information: only as gross; or only as net; or in both gross and net forms. Nearly half the countries reported these income components in both forms. A few reported only net (EL, IT, LV, PT) and the remaining only gross.

This may be contrasted with the situation as it emerged in 2007, from the time income components had to be recorded in the gross form. The right side of the table shows similar information for the full cross-sectional sample of 2007 for comparison. Here all countries have to report gross amounts, but a majority (over 60%) reported the net amount as well.

In situations where income is collected net but recorded gross, the amounts need to be converted into the required form. This is normally done on the basis of some micro-simulation procedure. Micro-simulation has similarity to imputation in that both involve some form of modelling whether explicit or implicit (micro-simulation tends to be more dependent on external data and relationships, while imputation often depends more on relationships between variables observed in the dataset itself).

Hence the extent of net-gross conversion involved should also be noted in the context of discussion on item non-response.

The table "Breakdown of self-employment income recorded as gross and net according to its mode of collection" shows the form of collection and recording of one important income component, namely self-employment income. The table has three panels. The first panel shows the number of individuals receiving self-employment income, missing cases (where it could not be imputed and/or converted to gross amount), and the number receiving and recording the amount.

The second panel of the table shows, among the case where self-employment income has been received and recorded, the percentage where the income is *recorded gross*, followed by the breakdown of that percentage according to the form in which the amount was *collected*. This indicates the extent and form of net-gross conversion, normally involving micro-simulation.

The third panel shows the form of collection where the net amount has been recorded. It shows, among the case where self-employment income has been received and recorded, the percentage where the income is *recorded net*, followed by the breakdown of that percentage according to the form in which the amount was *collected*.

In fact, the net collection as well as net recording can both be in different forms (such as after different types of deductions), and this information is also provided in this part of the table. When the form of net collected differs from the form in which the net is recorded, some conversion normally on the basis of modelling or micro-simulation is involved.

Finally, the table "Percent distribution according to value of the imputation flag" provides information on the 'imputation factor' for net and gross self-employment income. Actually, this flag gives the combined effect of imputation and micro-simulation (net-to-gross conversion). It is defined as the ratio of the amount collected to the amount recorded in the data base for the component concerned. A value of 1.0 means that the information recorded is exactly the same as was collected. A value of zero means that the information recorded is completed imputed; no amount was collected. A value between these limits means that the recorded amount exceeds the amount actually collected. This may be the result of imputation of missing subcomponents within the component concerned, and/or the result of conversion from net to gross amount on the basis of micro-simulation or other modelling (processes which in fact involve 'imputation' of taxes and other deductions).

Table 14: Mode of collection and recording of self-employment income

	Longitudinal 2-year sample 2006								Cross-sectional sample 2007						
				received & recorded		form of recording (%)						received & recorded			
	total persons	not receiving	not stated	number	%	only gross	only net	both	total persons	not receiving	not stated	number	%	also recorded net_(%)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(1)	(2)	(3)	(4)	(5)	(6)	
BE	6642	6204	0	438	6.59			100	12322	11554		768		100	**BE**
CZ	7717	7132	21	564	7.31	100			19384	17914	36	1434	7.58		**CZ**
DK	6420	4804	0	1616	25.17	100			11610	8921		2689	23.16		**DK**
DE	18486	17229	0	1257	6.80	100			26291	24698		1593	6.06		**DE**
EE	9110	8393	0	717	7.87			100	11971	11163		808	6.75	100	**EE**
IE	6089	5485	0	604	9.92			100	10892	9760		1132	10.39	100	**IE**
EL	8607	7155	0	1452	16.87		100		12346	9935		2411	19.53	86.19	**EL**
ES	19358	17840	0	1518	7.84			100	28656	26529		2127	7.42	100	**ES**
FR	12358	11818	0	540	4.37			100	20357	19477		880	4.32	100	**FR**
IT	32348	26946	0	5402	16.70		100		44629	37210		7419	16.62	100	**IT**
CY	6499	5820	0	679	10.45	99.71		0.29	8470	7510		960	11.33	0.42	**CY**
HU	11030	10323	0	707	6.41	100			18490	16610		1880	10.17		**HU**
MT															**MT**
NL	12997	11783	0	1214	9.34	100			19623	17745		1878	9.57		**NL**
AT	7662	6951	0	711	9.28			100	13391	12094		1297	9.69	100	**AT**
LV	6087	5793	0	294	4.83		100		9270	8869		401	4.33	100	**LV**
LT	6599	5984	0	615	9.32			100	10913	9881		1032	9.46	100	**LT**
LU	6533	6217	0	316	4.84	3.48		96.52	7913	7520		393	4.97	97.71	**LU**
PL	26639	24003	0	2636	9.90			100	34888	31199		3689	10.57	84.28	**PL**
PT	7398	6699	0	699	9.45		100		9947	8917		1030	10.35	100	**PT**
SI	18936	15964	0	2972	15.69			100	24730	20828		3902	15.78	100	**SI**
SK	9294	8812	0	482	5.19	100			12573	11952		621	4.94		**SK**
FI	10287	8117	0	2170	21.09	100			21773	17140		4633	21.28		**FI**
SE	8669	7478	0	1191	13.74			100	14204	12294		1910	13.45	100	**SE**
UK	12441	11536	0	905	7.27			100	17484	16212		1272	7.28		**UK**
IS	3905	3479	0	426	10.91	100			6567	5854		713	10.86		**IS**
NO	5382	4784	23	575	10.68	100			11706	10396	109	1201	11.19		**NO**

Source: Data from the 2007 UDB (1st March 2009).

Table 15: Breakdown of self-employment income recorded as gross and net according to its mode of collection (Longitudinal 2-year sample 2006)[22]

	(A) Whether received and recorded					(B) Recorded gross: mode of collection (PY050g_f)					(C) Recorded 'net': mode of collection and recording (PY050n_f)										
	total	not received	received	missing	received & recorded	total	1	3	4	5	total	11	22	31	33	41	42	51	52	55	
AT	7.662	6.951	711	0	711	100,0			100,0		100,0	100,0									**AT**
BE	6.642	6.204	438	0	438	100,0	53,7		46,3		100,0	100,0									**BE**
CY	6.499	5.820	679	0	679	100,0			100,0		0,3	0,3									**CY**
CZ	7.717	7.132	585	21	564	100,0	19,0		81,0		0,0										**CZ**
DE	18.486	17.229	1.257	0	1.257	100,0			100,0		0,0										**DE**
DK	6.420	4.804	1.616	0	1.616	100,0			100,0		0,0										**DK**
EE	9.110	8.393	717	0	717	100,0	45,9		49,1	5,0	100,0	45,9				49,1		5,0			**EE**
ES	19.358	17.840	1.518	0	1.518	100,0			100,0		100,0	100,0									**ES**
FI	10.287	8.117	2.170	0	2.170	100,0			100,0		0,0										**FI**
FR	12.358	11.818	540	0	540	100,0		100,0			100,0				100,0						**FR**
GR	8.607	7.155	1.452	0	1.452	0,0					100,0	100,0									**GR**
HU	11.030	10.323	707	0	707	100,0			100,0		0,0										**HU**
IE	6.089	5.485	604	0	604	100,0			100,0		100,0	34,3				65,7					**IE**
IS	3.905	3.479	426	0	426	100,0			100,0		0,0										**IS**
IT	32.348	26.946	5.402	0	5.402	0,0					100,0	100,0									**IT**
LT	6.599	5.984	615	0	615	100,0			100,0		100,0					100,0					**LT**
LU	6.533	6.217	316	0	316	100,0			100,0		96,5	96,5									**LU**
LV	6.087	5.793	294	0	294	0,0					100,0									100,0	**LV**
NL	12.997	11.783	1.214	0	1.214	100,0			100,0		0,0										**NL**
NO	5.382	4.784	598	23	575	100,0			100,0		0,0										**NO**
PL	26.639	24.003	2.636	0	2.636	100,0	100,0				100,0	100,0									**PL**
PT	7.398	6.699	699	0	699	0,0					100,0	11,0	0,9			21,0	1,7	52,5	12,9		**PT**
SE	8.669	7.478	1.191	0	1.191	100,0		100,0			100,0			100,0							**SE**
SI	18.936	15.964	2.972	0	2.972	100,0			100,0		100,0	100,0									**SI**
SK	9.294	8.812	482	0	482	100,0			100,0		0,0										**SK**
UK	12.441	11.536	905	0	905	100,0			100,0		100,0	100,0									**UK**

Source: UDB Cross-sectional and Longitudinal 2006

PY050g_f

ected (1st digit)

1 net of tax on income at source and social contributions
2 net of tax on income at source
3 net of tax on social contributions
4 gross
5 unknown

PY050n_f

Collected (1st digit)

1 net of tax on income at source and social contributions
2 net of tax on income at source
3 net of tax on social contributions
4 gross
5 unknown

Recorded (2nd digit)

1 net of tax at source and social contributions
2 net of tax on income at source
3 net of tax on social contributions

[22] Data from the 2007 UDB (1st March 2009).

Table 16: Percent distribution according to value of the 'imputation flag'. Cash benefits or losses from self-employment

Longitudinal sample 2-year duration (L2)
Percent disribution according to value of imputation flag (PY050N_I)

	0	0+0.25	0.25-0.5	0.5-0.75	0.75-1	1	>1	all	(T)
(A) NET amount (PY050N): percent disribution according to value of imputation flag (PY050N_I)									
AT	14	0	1	0	1	83	0	100	711
BE	31	0	0	0	0	69	0	100	438
EE	14	0	0	0	0	81	4	100	713
ES	27	7	6	10	16	23	11	100	1.518
FR	0	2	0	0	0	90	8	100	540
GR	0	0	0	0	0	100	0	100	1.452
IE	69	0	0	0	0	10	21	100	604
IT	15	0	0	0	0	83	2	100	5.402
LT	2	0	0	0	0	90	7	100	615
LU	66	0	1	0	0	33	0	100	305
LV	5	0	0	0	0	95	0	100	294
PL	32	0	0	1	2	64	0	100	2.344
PT	65	0	0	0	0	12	23	100	699
SE	0	0	0	0	0	100	0	100	1.191
SI	0	3	2	1	2	90	2	100	2.972
UK	16	0	0	0	0	0	84	100	905
(B) GROSS amount (PY050N): percent disribution according to value of imputation flag (PY050G_I)									
AT									
BE									
DE	7	2	2	2	1	86	1	100	1.257
EE	15	11	1	8	5	56	4	100	717
ES	16	13	0	16	21	31	4	100	1.518
FR	0	0	4	5	0	0	91	100	540
HU	30	0	0	0	0	70	0	100	707
IE	69	0	0	0	0	31	0	100	604
LT	2	0	0	1	13	82	1	100	615
LU	46	0	0	1	0	53	0	100	316
NL	0	0	0	0	0	100	0	100	1.214
PL	17	8	6	7	9	51	2	100	2.636
SE	0	0	0	0	0	100	0	100	1.191
SI	0	3	2	1	2	90	2	100	2.972
SK	0	0	0	0	0	100	0	100	482
UK	16	0	0	0	0	0	84	100	905

Longitudinal sample 3-year duration (L3)
Percent disribution according to value of imputation flag (PY050N_I)

	0	0+0.25	0.25-0.5	0.5-0.75	0.75-1	1	>1
AT	14	1	1	0	1	83	0
BE	32	0	0	0	0	68	0
EE	14	1	0	0	0	81	4
ES	25	6	6	10	16	26	11
FR	0	2	0	0	0	90	8
GR	0	0	0	0	0	100	0
IE	66	0	0	0	0	13	21
IT	14	0	0	0	0	84	1
LT							
LU	65	0	1	0	0	35	0
LV							
PL							
PT	65	0	0	0	0	12	22
SE	0	0	0	0	0	100	0
SI							
UK							
AT	45	0	0	1	0	47	6
BE	54	1	0	0	0	45	0
DE							
EE	15	11	1	8	5	56	4
ES	14	11	0	17	21	32	4
FR	0	0	4	5	0	0	91
HU							
IE	66	0	0	0	0	34	0
LT							
LU	45	0	0	1	0	54	0
NL							
PL							
SE	0	0	0	0	0	100	0
SI							
SK							
UK							

(T) = Total persons receiving self-employment income

Imputation factor = collected value / recorded value

Countries where the flag has not been coded are not included in the table concerned.

Source: Data from the 2007 UDB (1st March 2009).

A summary table (based on the information presented in the national quality reports complemented by bilateral e-mail exchanges) on the form in which income variables at component level have been obtained and the method used for obtaining income target variables in the required form (i.e. as gross values) can be found in the annex.

5.3. **Tracing rules**

Tracing rules are defined in Commission Regulation EC 1982/2003. Most of the countries follow the common rules, and some of them report in detail the procedure, in particular, Belgium and most of the 'register countries'[23].

- Belgium: Although the "tracing rules" from Eurostat say that sample households non-enumerated the first year of the panel "may be dropped", some households who did not participate in 2004 were contacted in 2005. These cases concern households who were not interviewed in 2004 because they were temporarily away, unable to respond due to illness or due to other reason (DB130=22 to 24).

- Finland: The tracing rules for the follow-up of sample persons, sample households and co-residents have been followed in the longitudinal survey according to the EU-SILC requirements framework. Because of the sampling design and the sampling unit definition used (the selected individuals); only the initial sample persons of the first wave are followed over the survey years. Households are constructed and household members are defined (mostly co-residents, see the household membership definition) around these sample persons. Household members include the ones who were currently living in the households containing the initial sample person or who were temporarily absent from that household at the end of the income reference period (31 December). Membership status is checked in the each wave.

- Slovenia: Due to the fact that in Slovenia use a sample of persons and each household has only one selected person, only the selected person is traced. These persons are at least 16 years old. Such person is traced, if he/she moves in the territory of Slovenia. If the sample person moved permanently into institution or collective household, such household is excluded from survey. Were excluded from the survey also the households where the sampled person died. In the case that sampled person moved, interviewers (CAPI) had to fill in special form, where they wrote new address, if they found it from persons who live in the address or from neighbours. They sent to the office these forms with new addresses and in the office is prepared an additional list of sampled persons which is sent to the appropriate interviewer. In the case that move person who was interviewed by phone,

[23] In the selected respondent data model used by register countries only the selected respondent is a panel individual and has to be traced out for the following waves. If the household splits, only information on the household of the selected individuals will be extracted from register. This model leads thus to the systematic drop out of non selected respondent in case of household split. This situation holds also for children less than 16 when the household splits. No information is obtained for children who move with the non selected respondent and these are thus virtually out of the panel. Although the weighting system ensures that the panel remains representative of the target population, structural drop out from the initial sample could decrease the size of the sample for the individual at stake (children and split household). This is one of the reasons why the sample size for the selected respondent data model has been increased in the Framework Regulation.

interviewer wrote the new address into the computer program and after the CATI interviewing period was finished, all lists are sent to the appropriate interviewers. In the case that interviewer could not get a new address, the Statistical office tried to find the new address from other sources. This way all selected persons and their households who moved are interviewed face to face under condition that the new address is available.

- Slovakia: 1. If the whole household has moved out, the interviewer had to find out its new address by all available sources. This information could be obtained from neighbours or relatives, municipal/communal office and others. Interviewer provide new address of household, name and surname of the head of the household in relevant form and also filled ID number of household and this form gave to coordinator of the Regional Office in period at least 3 days. Consequently coordinator decided on another procedure to continue in this circumstance. 2. Similarly interviewer proceeded in the case of one or more selected persons moved out. Basic source of information on place of moving of selected person/persons was information received from other household members. For each person moved out interviewer completed relevant form, where was listed new address of this person again, his/her name and surname, household ID and personal ID. 3. In the case if interviewer was entrusted to collect data for household or person moved out, needed information was received from coordinator of the relevant Regional Office.

- Iceland: In Iceland a respondent is selected from the national register. Whoever lives with the selected respondent is also included in the survey. If the composition of the households of the selected respondent changes between waves the other household members are not traced. Only the selected respondent is traced and if he or she has new household-partners they will be included in the survey. The information used for tracing are received from the national register, information on phone numbers are received from the largest phone company in Iceland. Information from former household members is also used to help to locate the selected respondents if the selected respondent has moved. All data are collected through telephone.

- Norway: In the Norwegian EU-SILC the respondent (person/selected respondent) is selected from the population register. All household members of the selected respondent are included in the survey. If the household composition changes between waves, household members are not traced. We only trace the selected respondent. Tracing is done by using updated data from the population register, data from the previous data collection and by searching for phone numbers. The interviewer can also apply different ad-hoc methods to trace respondents.

There is a lack of information on the tracing rules in the national quality reports of Denmark, Ireland and Sweden.

6. COHERENCE

There is a variety of sources to analyse coherence of EU-SILC with these other sources. The main sources used by the countries are: EU-SILC data from previous operations (considered as an analysis of the comparability of the data), Household Budget Survey (HBS), Labour Force Survey (LFS), National Accounts (NA) and administrative sources. In each survey or

administrative data variables similar to those in EU-SILC can be found and then the definitions and data can be compared taking as starting point EU-SILC variables. A summary table with the coherence studies carried out by the countries can be found in the annex.

In addition to the information in the national quality reports, some countries carry out specific studies on comparability. The following table summarizes the comparability studies carried out by the countries:

Table 17: EU-SILC impact studies on comparability of national implementations

Greece	1) Review of comparability of EU-SILC wages using multivariate analysis (factor analysis, multiple liner regression, generalized liner regression, etc.) and using data from structural earning surveys and labour cost surveys. 2) Analysis of quality of trend data from EU-SILC. 3) Comparability of EU-SILC social benefits and administrative data (ESSPROS).
France	The aim of the study is to match EU-SILC data and fiscal data in France and to study the comparability of the two sources.
Poland	The study covers: 1. Evaluation of the effect of income data imputation on the value and accuracy of income and poverty indicators. The comparative analysis of EU-SILC results with and without income data imputation will require: - the designing of a weighting system for the group of households which did not reveal missing components of income data; - the estimation of income categories (variables) and of differentiation and poverty indicators for the group of households which did not reveal missing components of income data; - the accuracy estimation of selected indicators. 2. Evaluation of the impact of including the imputed rent category on income level and poverty indicators. The comparative analysis of imputed rents estimated in EU-SILC and in HBS will be taken into account. 3. The comparative analysis of EU-SILC and HBS results. Before EU-SILC implementation household budget surveys provided the main source of data on income of the population in Poland. It was also on the basis of that survey that Laeken indicators of income and poverty were determined.
Norway	The main purpose of the action is to evaluate the data quality (comparability) of the housing cost variable collected by EU-SILC survey comparing: - different fieldwork arrangement using direct experiment on EU-SILC sample (split sample); - EU-SILC with different surveys (Norwegian consumer expenditure survey and Survey on housing conditions and housing costs) with special attention on the impact of the different collection mode used on non response bias; - EU-SILC survey data and register information.

7. ANNEXES

Annex 1: Sampling design

Belgium

The Belgian EU-SILC survey is a stratified two-stage sampling. There is no clustering of sampling units. The stratification is done by NUTS2 region (10 provinces plus the Brussels Capital region).

- Primary units: the municipalities (or part thereof in the larger ones) with probability proportional to size.
- Secondary units: private households by systematic sampling.

Czech Republic

A sample of dwellings is selected using a stratified two-stage design. The stratification of the Census Enumerations Units (CEUs-small geographical units) is done by region (NUTS4) and by number of residents in the municipality.

- At the first stage, CEUs are sampled as primary sampling units (PSU) with probability proportional to their size.
- In the second stage, 10 dwellings are sampled in each sampled CEU by simple random sampling without replacement.

All the households and the individuals living in the selected dwellings are then eligible for interview.

Denmark

The sampling design is simple random sampling. The sample is a one stage sampling being the sampling unit the individual person. The sampling frame is all individuals aged 14 or more but only households where the selected person is 16 or more at the beginning of the survey year are included in the indicators computation of that year.

Germany

In 2005 the survey started with three quota samples and one random sample. Each year one quota sample is replaced by a further random sample. The sampling frame for the random subsamples is the permanent sample (DSP), a sampling frame recruited among former participants of the German Microcensus.

All the individuals living in the selected addresses are eligible for interview.

Estonia

The design used is one-stage stratified unequal probability sampling of household, with a household selected with probability proportional to the number of persons aged 14 and more in it. The EU-SILC sample is selected according to the following sampling procedure:

- Stratification by county level into three strata by the population size: "big" counties, "small" counties and the Hiiu County, which forms a separate stratum as the smallest county in terms of population size.
- A sample of persons aged 14 and more is selected with equal probabilities within strata.

All the households of the selected persons are identified and all eligible persons in the household are interviewed.

Ireland

In 2004, the Irish EU-SILC sample is selected according to a stratified two-stage selection. The stratification is done by County and degree of urbanisation.

- At the first stage, simple random selection of dwelling blocks.
- At the second stage, simple random selection of households.

Greece

In 2003, a sample of addresses is drawn according to a stratified two-stage selection. The stratification is done by NUTS2 region and degree of urbanisation.

- At the first stage, a sample of blocks is selected with probability proportional to the number of dwellings.
- At the second stage, households are systematically selected within each block.

All the persons living in the selected addresses are then interviewed in order to obtain information at personal level.

Spain

A sample of dwellings is drawn according to a stratified two-stage selection. The stratification of the Census sections is done by administrative region and number of dwellings.

- At the first stage, selection of Census sections with probability proportional to the number of dwellings.
- At the second stage, systematic selection of dwellings within each section.

All the persons living in the selected dwellings are eligible for interview.

France

The type of sampling design is a stratified three-stage sampling. In 2004, a sample of dwellings is drawn from the 1999 Master Sample updated for the "new" dwellings (i.e. the units that came out after the 1999 Census). The selection is done so as to make the sample self-weighted.

- At the first stage, selection by groups of municipalities proportional to size (stratified according geographical criteria as NUTS2 and degree of urbanisation).
- At the second stage, the systematic selection is of dwellings for the urban areas and ad-hoc groups of municipalities for the rural areas.
- The third stage only exists for the rural areas and the dwellings are selected by systematic sampling.

All the households and the individuals living in the selected dwellings are interviewed.

Italy

In 2004, a sample of households is drawn according to a stratified two-stage selection. The stratification of the municipalities is done by administrative region and number of residents.

- At the first stage, selection of four municipalities with probability proportional to the number of residents.
- At the second stage, systematic selection of households within each municipality.

All the persons living in the selected households are then eligible for interview.

Cyprus

The sample design is one-stage stratification. The sampling units are private household which are selected by simple random sampling within each stratum (9 strata based on District).

All the individuals that are current members of the selected households are eligible for interview.

Latvia

The Latvian EU-SILC sample is according to a stratified two-stage design. The stratification is based on the degree of urbanisation.

- At the first stage, the primary sampling units (PSU, Population Census counting areas) are selected in each stratum with probability proportional to the number of households.
- At the second stage, a simple random sample of units (addresses) is selected within each area.

In Latvia several households can be registered in one address. All households and individuals living in the selected address are included in the survey.

Lithuania

The new subsample of households is selected by stratified sample design. The stratification is based on degree of urbanisation into seven strata.

- A simple random sample of non-institutional persons aged 16 and over is selected in each stratum from the Population Register.

Households where the selected persons live are surveyed.

Luxembourg

The type of sampling design is stratified simple random sampling. In 2003, first year of the survey, two samples are drawn independently:

- A sample of "tax" households, which are in fact a group of persons who depends on the same Social Security system.

- A sample of dwellings wherein none of the members depends on Luxembourgish Social Security system.

A "tax household" is basically a group of persons living in the same dwelling and who depend on the same Luxembourgish Social Security system.

The samples are selected by stratified simple random sampling.

Hungary

EU-SILC sample is selected by a stratified two-stage sampling in one part of the population and by a stratified one-stage sampling in the other part. Localities are stratified by General Election Districts and size (in terms of number of dwellings).

- In the first part of the population, one locality is selected with probability proportional to the number of dwellings. Within each selected locality, a systematic selection of dwellings is done.

- In the other part of the population, a systematic selection of dwellings is done in each stratum.

The final sampling units are the dwellings and, in each of them, every household is observed.

Malta

The sampling design involves simple random sampling of dwellings from the Census of Population and Housing database, which served as the sampling frame for this survey. Consequently, these dwellings have served as the best possible proxy to the household population that were targeted for this survey.

All the persons living in the selected dwellings are then interviewed in order to obtain information at personal level.

The Netherlands

The EU-SILC sample is composed of the addresses that took part in the Labour Force Survey (LFS) and are willing to cooperate to EU-SILC. The LFS sample is selected according to a stratified three-stage sampling design. The stratification of the municipalities is done by geographical criteria (COROP and interviewer region).

- At the first stage, municipalities are selected with a probability proportional to the number of addresses and according to the above mentioned stratification. At the second stage, there is a simple random selection of addresses within each municipality.

- At a third stage, persons of 16 and older are selected by simple random sampling.

The LFS has a panel structure with five rotational groups. When the first wave (face-to-face interviews) has been completed, addresses with all residents aged over 64 are removed from the sample. In order to get full covering of the target population, an additional sample of addresses with all residents aged 65 and over is drawn for the EU-SILC sample.

All the households and the individuals living at the selected addresses are then eligible for interview. Then, in each sampled household, a respondent is chosen to be fully interviewed, the information on other members of the household being obtained via the registers

Austria

The sampling design is simple random sampling without stratification.

All the households and the individuals living in the eligible addresses are interviewed.

Poland

The Polish EU-SILC sample is selected according to a stratified two-stage design. The stratification is based on NUTS2 region and degree of urbanisation.

- At the first stage, Census areas are selected with probability proportional to the number of dwellings.

- At the second stage, a simple random sample of dwellings is selected.

All the households and the individuals living in the selected dwellings are eligible for contact.

Portugal

The EU-SILC sample follows a stratified two-sage cluster sampling design.

- At the first stage, Census sections are systematically selected. Primary Sampling Units are the areas of the Master Sample (made of census enumeration areas) and they are stratified by a regional criterion.

- At the second stage, a simple random sample of households is selected in each Census section.

All the persons living in the same dwelling are interviewed.

Slovenia

The sample for the Slovenian EU-SILC is selected according to a stratified two-stage design. The strata are defined according to the size of the settlement and its proportion of agricultural households.

- In each stratum, Primary Sampling Units (PSU) are firstly systematically selected.
- In the second phase, seven persons aged 16 and over are selected in each PSU.

Finally, all the households the selected persons belong to are eligible for contact.

Slovakia

One-stage stratified sampling is used in EU-SILC. Stratification is based on geographical criteria (NUTS3 region and degree of urbanisation).

The proportional number of households is selected by simple random sampling in individual strata.

All the households and the individuals living in the selected dwellings are contacted.

Finland

The sampling design of the Finnish EU-SILC survey is a two-phase sampling design. In the first phase, a master sample is drawn by systematic sampling from the Population Register. Then, dwellings units are constructed by adding to the master sample all the persons sharing the same domicile code as the selected persons. The Master Sample is stratified by using a socio-economic categorisation of the dwelling units. In the second phase, a simple random sample of dwelling units is selected in each stratum of the master sample. Households are defined later on in the interview stage.

Sweden

A systematic sample of persons aged 16 and over is drawn from the Population Register (RTB). The final EU-SILC sample also includes a panel of persons that was drawn in 1980 and are re-interviewed every 8 year. In order to cover the whole target population, this panel has been supplemented every 8 year with a systematic sample of immigrants and a systematic sample of individuals aged 16-23.

Finally, all the households the selected persons belong to are then interviewed.

United Kingdom

Data is collected from two sources. First, data is collected by the Office of National Statistics (ONS), using the General Household Survey. Second, a sample of 300 households is collected by NISRA (Northern Ireland Statistics and Research Agency) as part of the "Living conditions survey".

EU-SILC uses a probability, stratified two-stage sample design. Households are sampled from the small users Postcode Address File (PAF).

- The postcode sectors are the Primary Sampling Units. The Postcode address file is ordered by postcode sector, which are similar in size to a UK electoral ward area.
- The Secondary Sampling Units are addresses within those sectors.

All adults aged 16 or over from every household at the sampled address are interviewed.

Iceland

The sampling design is one-stage simple random sample without stratification. The sampling units are persons aged 16 years and more living in private households selected from the Population Register.

All the households the selected persons belong to are then interviewed.

Norway

The EU-SILC in Norway comprises two parts. First, an "old" panel was drawn in 1997 according to a stratified two-stage design. Municipalities are stratified by socio-economic criteria and municipalities are drawn with probability proportional to the population size.

- A systematic sample of registered persons aged 16 and over is selected in each municipality so as to make the final sample self-weighted.

For the "new" part, the sample units are the persons aged 16 years and over that are registered in the Central Population Register. The sample is systematically drawn within one-year groups so as to maintain self-weighting.

All the households the selected persons belong to are then interviewed.

Annex 2: Sampling errors

Sampling errors: the concept

The particular units that happen to be selected into a particular sample depends on chance, the possible outcomes being determined by the procedures specified in the sample design. This means that, even if the required information on every selected unit is obtained entirely without error, the results from the sample are subject to a degree of uncertainty due to these chance factors affecting the selection of units. Sampling error is a measure of this uncertainty.

While survey data are subject to errors from diverse sources, information on sampling errors is of crucial importance in proper interpretation of the survey results, and for the purpose of evaluating and improving the sample design, including sample size. The importance of including information on sampling errors in survey reports cannot be over-emphasised.

Of course, sampling error is only one component of the total error in survey estimates, and not always the most important component. By the same token, it is the lower (and the more easily estimated) bound of the total error: *a survey will be useless if this component alone becomes too large for the survey results to add useful information with any measure of confidence to what is already known prior to the survey*.

Furthermore, survey estimates are typically required not only for the whole population but also separately for many subgroups in the population. Generally, the relative magnitude of sampling error compared to that of other types of errors increases as we move from estimates for the total population to estimates for individual subgroups and comparison between subgroups. *Information on the magnitude of sampling errors is therefore essential in deciding the degree of detail with which the survey data may be meaningfully tabulated and analysed*.

Similarly, sampling error information is needed for sample design and evaluation. While the design is also determined by many other considerations (such as costs, availability of sampling frames, the need to control measurement errors), rational decisions on the choice of sample size, allocation, clustering, stratification, estimation procedures etc. can only be made on the basis of detailed knowledge of their effect on the magnitude of sampling errors of statistics obtained from the survey.

Technical methodology

This section describes practical procedures for estimating sampling errors with particular reference to the Jack-knife Repeated Replication method, officially adopted by Eurostat. It also describes the standardisation in the procedures utilised by Eurostat for providing a standard tool which *can be used unchanged for any country and any survey year* for the statistics specified for the Final Quality Report.

Variance computation procedures

Large scale household surveys are generally based on multi-stage, stratified and otherwise complex designs. A typical survey is multi-purpose in several respects: it involves many types of interrelated variables; many types of estimates such as proportions, means, ratios and differences of ratios, and more complex statistics; various types of units of analysis such as households and individuals; various levels of disaggregation of the sample; and diverse and

numerous subclasses (subpopulations) for which estimates of levels, differences and other relationships are required. Practical procedures for estimating sampling errors therefore: (i) must take into account the actual, complex structure of the design; (ii) should be flexible enough to be applicable to diverse designs; (iii) should be suitable and convenient for large-scale application, and for producing results for diverse statistics and subclasses; (iv) should be robust against departure of the design in practice from the ideal `model' assumed in the computation method; (v) should have desirable statistical properties such as small mean-squared error of the variance estimator; (vi) should be economical in terms of the effort and cost involved; and (vii) suitable computer software should be available for application of the method.

The theory of 'simple replicated variance estimators' provides the basis for most practical approaches to variance estimation, though in application to complex situations, additional assumptions and approximations may be involved. Drawing on this basic idea, two broad practical approaches to the computation of sampling errors may be identified:

(1) Computation from comparisons among estimates for replications of the sample, each of which reflects the structure of the full sample, including its clustering and stratification.

(2) Computation from comparisons among certain aggregates for primary selections or replicates within each stratum of the sample, also known as linearization method.

The Jack-knife Repeated Replication is a commonly used method which belongs to class (1). This is the method adopted and developed for application in EU-SILC at the EU level and also in countries that chose to use it.

Repeated replication procedures

JRR is one of the classes of practical methods for variance estimation in complex samples based on *measures of observed variability among replications of the full sample*. The basic requirement is that the full sample is composed of a number of subsamples or replications, each with the same design and reflecting complexity of the full sample, enumerated using the same procedures. However, as the replications are not independent, and special procedures are required in constructing them to avoid bias in the resulting variance estimates.

A replication differs from the full sample only in size. But its own size should be large enough for it to reflect the structure of the full sample, and for any estimate based on a single replication to be close to the corresponding estimate based on the full sample. At the same time, the number of replications available should be large enough so that comparison among replications gives a stable estimate of the sampling variability in practice. The various re-sampling procedures available differ in the manner in which replications are generated from the parent sample and the corresponding variance estimation formulae evoked (such as the Balanced Repeated Replication (BRR) and the bootstrap, apart from JRR).

Compared to the 'linearisation' method, repeated re-sampling methods tend to involve heavier computational work. However, they have the major advantage of not requiring an explicit expression for the variance of each particular statistic. They are also more encompassing: by repeating the entire estimation procedure independently for each replication, the effect of various complexities, such as each step of a complex weighting procedure, can be incorporated into the variance estimates produced.

Jack-knife Repeated Replication (JRR)

The basic model of the JRR for application in the context described above may be summarised as follows. Consider a design in which two or more primary sampling units (PSUs) have been selected independently from each stratum in the population. Within each primary selection (PS), sub-sampling of any complexity may be involved, including weighting of the ultimate units.

In the "standard" version, each JRR replication can be formed by eliminating one sample PSU from a particular stratum at a time, and increasing the weight of the remaining sample PSUs in that stratum appropriately so as to obtain an alternative but equally valid estimate to that obtained from the full sample.

The above involves creating as many replications as the number of primary units in the sample. The computational work involved is sometimes optimised by reducing the number of replications required. For instance, by grouping PSUs within strata, or by forming JRR replications by eliminating a whole group at a time. This is possible only when any stratum contains several units. One situation in which some grouping of units is unavoidable is when the sample or a part of it is a direct sample of ultimate units or of small clusters, so that the number of replications under "standard" JRR is too large to be practical. Alternatively, or in addition, the groupings of units may be cut across strata. It is also possible to define the replications in the standard way ("delete one-PSU at a time Jack-knife"), but actually construct and use only a subsample of those.

Briefly, the standard JRR involves the following.

Let z be a full-sample estimate of any complexity, and $z_{(hi)}$ be the estimate produced using the same procedure after eliminating primary unit i in stratum h and increasing the weight of the remaining (a_h-1) units in the stratum by an appropriate factor g_h (see below). Let $z_{(h)}$ be the simple average of the $z_{(hi)}$ over the a_h sample units in h. The variance of z is then estimated as:

$$\mathrm{var}(z) = \Sigma_h \left[(1 - f_h) . g_h . \Sigma_i \left(z_{(hi)} - z_{(h)} \right)^2 \right]$$

A major advantage of a procedure like the JRR is that, under quite general conditions for the application of the procedure, the same and relatively simple variance estimation formula holds for z of any complexity. Normally, the factor g_h is taken as $g_h = a_h / (a_h - 1)$, but for reasons noted below, it is preferable to use $g_h = w_h / (w_h - w_{hi})$, where $w_h = \Sigma_i w_{hi}$, $w_{hi} = \Sigma_j w_{hij}$, the sum of sample weights of ultimate units j in primary selection i. The latter form retains the total weight of the included sample cases unchanged across the replications created. With the sample weights scaled such that their sum is equal (or proportional) to some external more reliable population total, population aggregates from the sample can be estimated more efficiently, often with the same precision as proportions or means.

The JRR variance estimates take into account the effect on variance of aspects of the estimation process which are allowed to vary from one replication to another.

Variance estimation of measures based on subpopulations

(1) Normally, variance estimation for subpopulations does not involve any new procedures: the same formulae apply except that sample elements not members of the subpopulation of interest are simply disregarded. The only complication which sometimes arises is that, considering only the subpopulation members, some strata and PSUs may become empty. This would normally require some re-definition of the sample structure for the purpose of variance estimation. This is true whether the linearisation or the JRR method is being used.

(2) In the context of poverty and inequality, the subpopulation measures of interest are usually of a special type: in these, *all (or some) of the parameters involved in the definition of the measure are estimated from the full sample, while the measure itself is being estimated for the subpopulation concerned*. The most important example is the at-risk-of-poverty rate for a subpopulation, but with an individual's poverty status defined in relation to the poverty line determined from income distribution of the whole population. The JRR methods can be easily adapted for this purpose on the following lines.

In the Final Quality Report, however, the statistics for which sampling errors are asked for all concern total household income, and individual income components received at household and personal levels. Computing sampling errors for subpopulations for such statistics does not involve the complications mentioned in (2) above. The normal, simpler procedures (1) for subpopulations apply.[24]

Standardisation of the variance computation procedure

Here a standard procedure for the computation of sampling errors in the EU-SILC is described. The programs implemented provide a standard tool which *can be used unchanged for any country and any survey year* for the statistics specified for the Final Quality Report. The sample design of course varies from one country to another, and can also vary in detail from one survey year to another in the same country. Two steps must be completed before application of the standardised SAS programs for variance estimation. These are:

(1) The definition of the units to be included in the dataset.

(2) Definition for each unit of the 'computational' variables. The definition of computational strata and primary sampling units can be a technically complex task requiring sampling expertise, as well as knowledge of details concerning the sample design, selection and implementation – details which are country- and possibly even wave-specific. Figure 1 shows the overall structure of the recommended variance computation procedure.

[24] Complications of type (2) however arise for the type of statistics covered in the Intermediate Quality Report.

Figure 8: Country-specific and standardised aspect of the variance estimation procedure

Country- and application-specific aspects: →

(1) Creation of the data set, comprising of units to be included in the computations

(2) Creation of the sample structure variables (stratum, PSU and sample weight) for each unit (planned report describing basic principals of the procedures)

…forming input to the standard SAS programs, same for all countries and waves: →

(3) 'Creation structure for JRR' program

(4) 'JRR shell' of the SAS programs

(5) Variable-specific macros called from the JRR shell

Note that the 'JRR shells' referred to above are highly standardised: they are not specific to country nor to individual variable or statistic. They only require some limited variation from one *class* of statistics to another, such as between the production of the set of statistics required for the Final Quality Report.

Useful measures of sampling errors - some basic concepts

The magnitude of standard error of a statistic depends on a variety of factors such as:

- the nature of the estimate
- its units of measurement (scale) and magnitude
- variability among elements in the population (population variance)
- sample size
- the nature and size of sampling units
- sample structure; sampling procedures
- estimation procedures.

Consequently, the value of standard error for a particular statistic is specific to the statistic concerned. To relate standard error of one statistic to that of another, it is necessary to decompose the error into components from which the effect of some of the above factors has been removed; that is, into components which are more stable or 'portable' from one type of statistic or design to another statistic or design. The standard error of a statistic such as a mean is written in the following equations in several forms, in terms of measures which are more portable in the above sense.

Relative standard error, rse

$$se(\bar{y}) = \bar{y}.rse(\bar{y})$$

This refers to standard error of an estimate, divided by the value of the estimate. It removes the effect on standard error of the magnitude and scale of measurement of the estimate, but still depends on other factors such as sample size and design.

Standard error in an equivalent simple random sample (SRS); population variance

Standard error of a statistic estimated from a complex sample can be factorised into two parts:

(1) sr standard error which would have been obtained in a simple random sample of the same size;

(2) deft the design factor, summarising the effect of design complexities.

$$se(\bar{y}) = deft.sr(\bar{y})$$

The second component (*sr*) is independent of the sample design and relates to the sample size in a very simple way:

$$sr(\bar{y}) = s/\sqrt{n},$$

where *s*, standard deviation, is a measure of variability in the population, independent of sample design or size. (Population variance refers to the square of s.) The scale of measurement can also be removed by considering the coefficient of variation, *cv*:

$$s = \bar{y}.cv$$

Standard deviation and cv are a useful and highly portable measures. Furthermore, they can be estimated in a simple way irrespective of complexities of the design in most practical situations. For example for a proportion *p*

$$s^2 = \frac{n}{n-1}.p(1-p) \approx p(1-p),$$

while more generally, for a weighted ratio *r* we have:

$$s^2 = \left(\frac{n}{n-1}\right).\sum_i w_i z_i^2 / \sum_i w_i;\ z_i = (y_i - r.x_i)/\bar{x}$$

where

$$r = \sum_i w_i y_i / \sum_i w_i x_i;\quad \bar{x} = \sum_i w_i x_i / \sum_i w_i$$

The coefficient of variation is more portable, but it is not so useful when the denominator in its definition is close to zero, as for example may happen for estimates of differences between subclasses. Also, there is generally *no advantage in going from s to cv in the case of proportions; in fact the former is preferable* since it is symmetrical (the same) for a proportion (p) and its complement (1-p). In fact in many social surveys, most statistics of interest are likely to be in the form of *proportions* rather than means or general ratios. For proportions or percentages, it is important to keep a clear distinction between the error expressed in relative terms (as % of the proportion p), and in terms of *absolute percentage points*. (Example: A poverty rate of 22% differs from a rate of 20% by 10% in relative terms, but by 2 percentage points in absolute terms.) Both forms are relevant. For large proportions, the error is often

better expressed in relative terms, while for very small proportions expression in terms of absolute percentage points is often more meaningful.

The above is also true of measures which are similar to proportions, such as the at-risk-of-poverty rate, which is the main statistics presented in the intermediate quality report. Indeed, the at-risk-of-poverty rate is the central statistic of interest in EU-SILC. This is a complex statistic, but in certain respects it is similar to a simple proportion.

The design effect

The *design effect*, *deft*^{2}, (or its square-root, *deft*, which is sometimes called the *design factor*) is a comprehensive summary measure of the effect on sampling error of various complexities in the design. By taking the ratio of actual to simple random sample (SRS) standard error, deft removes the effect of factors common to both, such as size of the estimate and scale of measurement, population variance and overall sample size. However, for a given variable, its magnitude still depends on other features of the design.

A major factor determining the deft value is the size of the sample taken per PSU, the nature of the units used as PSUs and the sub-sampling procedures within those units.

In practice, design effect for a statistic is computed by estimating its variance (i) under the actual sample design, and (ii) assuming a simple random sample of the same size. The ratio of these two quantities gives *deft*^{2}.

Effective sample size

As already noted, sampling precision is determined by size of the sample, as well as by its design, that is, its efficiency or design effect. Both of these factors are specific to the statistic being considered. It is helpful to keep separate the issue of design effect. The precision requirements are more clearly expressed and understood in terms of the "effective" rather than actual sample size. By effective *sample size* of a sample with complex design, we mean the *size of a simple random sample of analysis units which has the same precision as the complex design*. The effective size of a complex sample of size n with design effect deft^2 is:

$$n_{eff} = \frac{n}{deft^2}$$

In place of the value of standard error, the required level of precision is sometimes expressed in terms of the "95% confidence interval", which corresponds to an interval 2 standard errors wide around the estimated value.

It can be easily seen that the effective sample size can also be expressed in terms of *cv*, the coefficient of variation and *rse*, the relative standard error:

$$n_{eff} = \left(\frac{cv}{rse} \right)^2$$

Further note on design effect

As defined above, design effect the ratio of the variance under the given sample design, to the variance under a simple random sample of the same size: $se = se_R \cdot deft$

Proceeding from estimates of sampling error to estimates of design effects (ratio of actual sampling error to that under equivalent simple random sampling, SRS) is essential for understanding the patterns of variation in and the determinants of magnitude of the error, for smoothing and extrapolating the results for diverse statistics and population subclasses, and for evaluating the performance of the sampling design. Computing design effects requires the additional step of estimating sampling errors under simple random sampling. The standard SAS programs provided by Eurostat for variance estimation implement a practical procedure for achieving this using JRR methodology.

Design effect itself can be decomposed into three components.

(1) the effect of sample weights on variance

(2) the effect of clustering, stratification and aspects other than weighting, and

(3) if applicable, the effect of clustering of persons within households.

Factor (1) does not depend on the structure of the sample, other than the presence of unequal sample weights for the elementary units of analysis. The main effect is the variability of these weights in the sample. The effect is also influenced by the extent to which the variable being estimated is correlated with the sample weights.

Factor (2) is the design effect resulting from stratification and clustering, i.e. sample structure factors other than the sample weights. For income variables which are of interest, it is normally computed on the basis of comparison of the actual (generally clustered and stratified) sample with the results from a simple random sample of *households*. This is because income is essentially measured at the household level: total household income, even if obtained from incomes of individual members, is then equalised, and the equalised amount ascribed to each member in a uniform way. Note that the above consideration applies also to 'register countries', since in those countries as well, the household remains the basic unit for the collection of the *income variables*.

The design effects presented in this report are the product of Factors (1) and (2). Factor (3) has not been included.

Factor (3), deft from clustering of persons within households, can arise when the units of analysis are different from households. For the variables included for sampling errors in the Final Quality Report, the situation is as explained in the table below.

- For variables concerning total household income and net and gross income components at household level, the units of analysis are households and Factor (3) in the design effect is 1.0 by definition.
- Every member in a household is assigned the household equivalised income. This income is identical for all members of a household. For such indicators, the comparison for the purpose of design effect is shifted from a simple random sample of households to a simple

random sample of persons. It can be seen from theory that *for income variables* (which are constant for all members of a household) *this additional factor in the design effect approximately equals square-root the average number of persons per household.* This applies to the variable equivalised disposable income for classes by household characteristics such as size, for which sampling errors are included in the final quality reports.

- This factor is smaller when particular subgroups of persons are considered, such as a particular age and sex group. It is the average number per household of individuals of the particular category of interest which matters in determining Factor (3) affecting the design effect. This applies to the variable equivalised disposable income for subpopulations for which sampling errors are included in the final quality reports.

- The situation is somewhat more complex in the case of net and gross income components collected at personal level. The individual responses can be expected to be correlated within a given household, but are not identical for different persons. This gives a design effect which is less than the number of adults per household receiving the income component concerned. In principle, the within-household correlation can be even negative, giving Factor (3) under 1.0.

Table 18: Units of analysis and the effect of clustering of persons within households ("Factor (3)")

	Unit of analysis	**Deft^2 from clustering within households**
Total household income	All households	= 1.0
Income components		
Net income components at household level Gross income components at household level	Households receiving the component	= 1.0
Net income components at personal level Gross income components at personal level	Persons receiving the component concerned	< (the average number of persons receiving the component per household)
Equivalised disposable income		
Subclasses by household characteristics (e.g. size)	Person	= (household size)
Subclasses by personal characteristics (e.g. age group, sex)	Person	= (average number of persons in the particular category per household)

The longitudinal sample

In the EU-SILC Users' Data Base (UDB) the "longitudinal data set" has been constructed simply on the basis of rotation groups. More precisely, it consists of (i) the rotation groups present in both the current (Y) and the previous (Y-1) waves; and (ii) and of those rotation groups, the ones present in each of the preceding two waves, if any. Some of the statistics presented in the national final quality reports are based on this data set. However, for the purpose of constructing longitudinal indicators such as on sampling errors, non-response and follow-up rates, it is necessary to extract from this data base a properly longitudinal sample.

The longitudinal samples are identified on the basis of *continuous presence of individual persons in the survey* for the specified number of most recent years, for two most recent years for the 2-year longitudinal sample; three most recent years for the 3-year longitudinal sample,

and so on. An "expansion" of the longitudinal sample base is required to ensure inclusion of *whole* households with all their members, as required for computation and analysis of income variables. The final set of units for inclusion in the computation can be identified in terms of the above as follows.

- Longitudinal persons, i.e. individuals present in the last two, three and four waves provide the basis for the longitudinal samples of corresponding durations.
- At each wave of the "longitudinal component", the household included are those which contain at least one longitudinal person at that wave.
- All persons present in the above mentioned set of household at each wave are included in the analysis of the longitudinal sample. This applies also to adults (including if applicable 'selected respondents') who provide the interview and/or information on their income.

The structure of the longitudinal sample is shown in Figure 2 (for a sample with standard design involving four rotational groups, after 4 or more years of its operation).

Figure 9: On the construction of the longitudinal sample

1. Cross-sectional sample (year Y)	**2. Longitudinal sample: present last two years**	**3. Longitudinal sample: present last three years**	**4. Longitudinal sample: present last four years**

Technical specification of the statistics analysed in the final quality reports

The statistics to be analysed in the final quality reports are those reported in the Commission Regulation (EC) No. 28/2004, summarised in the introduction to this section. Sampling errors for the specified (cross-sectional) measures are computed for (up to) four overlapping sample bases:

- full cross-sectional sample for the current year (Y);
- longitudinal sample over the two years (Y , Y-1) and (Y-2);
- longitudinal sample over three years (Y-2, Y-1, Y);
- and ultimately, also the longitudinal sample over four years (Y-3 to Y).

The reference data for each of the above sample basis is always fro the current year (Y). The units of analysis are all or particular categories of households or persons, depending on the type of statistic as defined above in the table, namely:

- total household income variables – all households (including household with zero reported income);
- income components collected at household level – households receiving non-zero income from the component concerned;
- income components collected at personal level – persons receiving non-zero income from the component concerned;
- equivalised disposable income classified by household size – all households; and
- equivalised disposable income classified by age and sex – all persons.

The longitudinal weight variables in EU-SILC UDB, with slight modification, provide the weights to be used in the calculations.

AUSTRIA

Variable		(1) Sample base: Full cross-sectional sample Source: Cross-sectional data 2006				(2) Sample base: 2-years longitudinal sample (2005-06) Source: Longitudinal data 2006				(3) Sample base: 3-years longitudinal sample (2004-05-06) Source: Longitudinal data 2006			
		estimate	standard error	sample size	deft	estimate	standard error	sample size	deft	estimate	standard error	sample size	deft
		(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Total household income													
Total household gross income	**HY010**	41,716	400	6,028	1.07	42,266	527	3,809	1.11	41,442	614	2,102	1.13
Total disposable household income	**HY020**	31,534	286	6,028	1.09	31,823	383	3,809	1.11	31,357	449	2,102	1.12
... excluding transfers except pensions	**HY022**	28,233	289	6,028	1.09	28,791	374	3,809	1.11	28,517	458	2,102	1.12
... excluding all transfers	**HY023**	20,834	316	6,028	1.06	21,114	413	3,809	1.10	20,262	463	2,102	1.11
Household level income components													
Property income	**HY040n**	9,615	1,170	226	1.02	11,242	2,466	148	1.15	12,753	3,148	85	1.16
	HY040g	9,240	1,282	225	1.01	10,588	2,448	147	1.10	12,100	3,064	84	1.11
Family/Children allowances	**HY050n**	4,703	70	2,120	1.12	4,747	92	1,381	1.16	4,536	121	804	1.07
	HY050g	4,703	70	2,120	1.12	4,747	92	1,381	1.16	4,536	121	804	1.07
Other social exclusions	**HY060n**	3,075	603	120	1.32	3,122	544	71	1.02	2,056	515	34	1.03
	HY060g	3,075	603	120	1.32	3,122	544	71	1.02	2,056	515	34	1.03
Housing allowances	**HY070n**	1,470	71	204	1.09	1,477	91	123	1.10	1,457	138	66	1.08
	HY070g	1,470	71	204	1.09	1,477	91	123	1.10	1,457	138	66	1.08
Inter-household transfers received	**HY080n**	4,704	293	410	1.20	4,715	317	256	1.06	4,242	360	139	1.06
	HY080g	4,704	293	410	1.20	4,715	317	256	1.06	4,242	360	139	1.06
Capital income	**HY090n**	338	21	4,588	1.01	398	33	2,917	1.09	417	48	1,619	1.08
	HY090g	422	27	4,588	1.01	498	41	2,917	1.09	521	60	1,619	1.08
Mortgage interest	**HY100n**												
	HY100g												
Children's income	**HY110n**	1,661	192	53	1.11	1,654	270	28	1.07	1,495	287	21	1.00
	HY110g	1,784	216	53	1.12	1,672	271	28	1.07	1,506	290	21	1.01
Regular taxes on wealth	**HY120n**												
	HY120g												
Inter-household transfers paid	**HY130n**	3,748	175	392	1.05	3,766	221	242	1.09	3,809	283	132	1.10
	HY130g	3,748	175	392	1.05	3,766	221	242	1.09	3,809	283	132	1.10
Tax	**HY140n**												
	HY140g	10,174	151	5,923	1.06	10,379	186	3,749	1.09	9,984	224	2,076	1.12
Tax adjustment	**HY145n**	-239	32	2,499	1.11	-258	40	1,683	1.09	-209	48	943	1.10

Austria (cont.)

Variable		estimate	standard error	sample size	deft	estimate	standard error	sample size	deft	estimate	standard error	sample size	deft
		(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Personal level income components													
Employee cash or near cash income	**PY010n**	17,212	163	6,254	1.08	17,519	198	3,978	1.12	17,185	272	2,209	1.13
	PY010g	24,478	266	6,254	1.07	24,958	325	3,978	1.11	24,373	438	2,209	1.14
Non-Cash employee income	**PY020n**												
	PY020g												
Contributions to private pension	**PY035n**	1,090	28	2,732	1.13	1,120	45	1,815	1.34	1,129	69	1,029	1.37
	PY035g	1,090	28	2,732	1.13	1,120	45	1,815	1.34	1,129	69	1,029	1.37
Self-employment income	**PY050n**	14,476	578	1,098	1.08	13,996	631	711	1.06	13,860	842	405	1.08
	PY050g	18,707	714	1,098	1.06	18,380	840	711	1.06	17,853	1,104	405	1.10
Production for own consumption	**PY070n**	239	18	259	1.05	228	25	164	1.04	219	31	88	1.01
	PY070g	239	18	259	1.05	228	25	164	1.04	219	31	88	1.01
Pension from private plans	**PY080n**	3,663	1,022	29	1.05	3,472	1,248	19	0.99	2,745	930	9	0.85
	PY080g	3,936	1,269	29	1.06	3,852	1,608	19	1.00	2,749	930	9	0.85
Unemployment benefits	**PY090n**	4,512	160	724	1.14	4,222	179	465	1.04	4,132	316	237	1.03
	PY090g	4,588	170	724	1.10	4,341	198	465	1.02	4,295	353	237	1.01
Old-age benefits	**PY100n**	15,385	179	3,045	1.10	15,801	238	1,943	1.19	16,031	361	1,098	1.17
	PY100g	18,816	254	3,045	1.08	19,433	342	1,943	1.18	19,750	527	1,098	1.17
Survivor' benefits	**PY110n**	8,481	538	105	1.12	8,723	536	72	1.04	9,080	655	37	1.03
	PY110g	10,467	715	105	1.12	10,689	734	72	1.06	11,059	952	37	1.04
Sickness benefits	**PY120n**	2,771	337	181	1.01	3,078	521	106	1.04	2,563	483	62	1.01
	PY120g	3,392	392	181	1.03	3,681	572	106	1.04	3,085	599	62	0.99
Disability benefits	**PY130n**	12,547	402	366	1.18	12,712	454	241	1.08	12,990	521	131	1.03
	PY130g	14,773	551	366	1.20	14,979	620	241	1.10	15,258	693	131	1.04
Education-related allowances	**PY140n**	3,078	323	178	1.05	2,968	697	102	1.22	2,903	965	67	1.22
	PY140g	3,078	323	178	1.05	2,968	697	102	1.22	2,903	965	67	1.22
Employees' gross monthly earnings	**PY200g**	1,919	17	5,682	1.08								
Equivalised mean income by household size													
1 household member	**HX090**	17,947	262	1,755	1.10	18,466	360	1,040	1.12	18,289	489	544	1.13
2 household members	**HX090**	21,555	306	1,823	1.07	21,801	428	1,189	1.15	21,673	501	653	1.14
3 household members	**HX090**	21,239	400	1,053	1.13	21,295	465	676	1.11	20,911	576	376	1.20
4 and more	**HX090**	18,367	250	1,397	1.10	18,543	297	904	1.10	18,685	411	529	1.10
all households	**HX090**	19,594	147	6,028	1.10	19,905	195	3,809	1.14	19,748	251	2,102	1.16
Equivalised mean income by age class and by gender													
<25	**HX090**	18,028	209	4,513	1.11	18,292	242	3,063	1.10	18,349	322	1,805	1.12
25 to 34	**HX090**	19,526	305	1,677	1.10	19,703	388	1,058	1.07	19,780	511	572	1.11
35 to 44	**HX090**	20,439	322	2,382	1.14	20,505	421	1,520	1.19	20,150	517	841	1.24
45 to 54	**HX090**	21,164	283	2,121	1.08	21,264	411	1,363	1.13	20,681	519	762	1.18
55 to 64	**HX090**	22,400	335	1,859	1.09	22,821	345	1,249	1.13	22,454	538	689	1.12
65+	**HX090**	18,655	246	2,331	1.07	19,253	336	1,471	1.17	19,385	511	844	1.15
Male	**HX090**	20,030	184	7,178	1.10	20,298	228	4,716	1.14	20,288	298	2,649	1.16
Female	**HX090**	19,334	156	7,705	1.11	19,562	199	5,008	1.14	19,355	252	2,864	1.17
all persons	**HX090**	19,674	156	14,883	1.10	19,918	197	9,724	1.14	19,800	250	5,513	1.17

DENMARK

Variable		(1) Sample base: Full cross-sectional sample Source: Cross-sectional data 2006				(2) Sample base: 2-years longitudinal sample (2005-06) Source: Longitudinal data 2006				(3) Sample base: 3-years longitudinal sample (2004-05-06) Source: Longitudinal data 2006			
		estimate	standard error	sample size	deft	estimate	standard error	sample size	deft	estimate	standard error	sample size	deft
		(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Total household income													
Total household gross income	**HY010**	51,748	562	5,711	1.07	54,433	779	3,210	1.15	52,177	1,193	1,997	1.28
Total disposable household income	**HY020**	33,895	344	5,711	1.07	35,522	494	3,210	1.19	34,262	778	1,997	1.39
... excluding transfers except pensions	**HY022**	28,482	361	5,711	1.08	30,337	496	3,210	1.17	29,149	781	1,997	1.35
... excluding all transfers	**HY023**	24,452	367	5,711	1.04	26,202	542	3,210	1.13	25,267	803	1,997	1.28
Household level income components													
Property income	**HY040n**												
	HY040g	2,432	266	134	1.06	2,595	399	85	1.14	2,717	421	53	1.04
Family/Children allowances	**HY050n**												
	HY050g	2,850	44	2,236	1.25	2,938	73	1,263	1.56	2,833	67	798	1.31
Other social exclusions	**HY060n**												
	HY060g												
Housing allowances	**HY070n**												
	HY070g	2,263	69	693	1.15	2,383	116	295	1.19	2,225	176	177	1.41
Inter-household transfers received	**HY080n**												
	HY080g	2,074	98	333	1.27	2,169	98	160	1.12	2,261	165	95	1.18
Capital income	**HY090n**												
	HY090g	539	106	5,640	0.99	514	119	3,179	0.94	296	148	1,980	0.88
Mortgage interest	**HY100n**												
	HY100g	4,771	62	3,338	1.08	4,792	81	1,993	1.11	4,774	100	1,243	1.15
Children's income	**HY110n**												
	HY110g	1,261	143	540	1.32	1,375	184	318	1.40	1,244	109	213	1.08
Regular taxes on wealth	**HY120n**												
	HY120g	1,000	14	4,100	1.08	984	17	2,434	1.12	938	21	1,530	1.23
Inter-household transfers paid	**HY130n**												
	HY130g	2,308	111	263	1.20	2,105	106	124	1.04	2,145	181	76	1.07
Tax	**HY140n**												
	HY140g	17,392	228	5,678	1.04	18,377	301	3,196	1.07	17,512	444	1,987	1.12
Tax adjustment	**HY145n**												

Denmark (cont.)

Variable		estimate	standard error	sample size	deft	estimate	standard error	sample size	deft	estimate	standard error	sample size	deft
		(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Personal level income components													
Employee cash or near cash income	**PY010n**												
	PY010g	32,078	297	8,443	1.21	33,004	391	4,815	1.18	32,376	528	2,975	1.23
Non-Cash employee income	**PY020n**												
	PY020g	2,649	156	680	1.11	2,677	196	404	1.07	2,752	303	242	1.15
Contributions to private pension	**PY035n**												
	PY035g												
Self-employment income	**PY050n**												
	PY050g	7,528	678	2,842	0.96	7,541	631	1,616	0.99	6,494	1,008	997	1.17
Production for own consumption	**PY070n**												
	PY070g												
Pension from private plans	**PY080n**												
	PY080g												
Unemployment benefits	**PY090n**												
	PY090g	9,384	214	2,061	1.23	9,053	281	1,160	1.21	8,696	397	723	1.31
Old-age benefits	**PY100n**												
	PY100g	18,669	287	1,574	0.99	18,409	347	916	1.00	17,638	437	608	1.06
Survivor' benefits	**PY110n**												
	PY110g	7,234	810	75	1.15	5,133	927	39	1.12	6,792	1,749	19	1.31
Sickness benefits	**PY120n**												
	PY120g	4,365	241	1,059	1.25	3,918	283	558	1.23	3,888	505	327	1.49
Disability benefits	**PY130n**												
	PY130g	16,392	432	688	1.11	16,217	480	369	1.09	15,448	743	223	1.21
Education-related allowances	**PY140n**												
	PY140g	4,611	137	797	1.29	4,744	233	407	1.52	4,756	286	258	1.47
Employees' gross monthly earnings	**PY200g**												
Equivalised mean income by household size													
1 household member	**HX090**	18,973	332	1,109	1.06	19,411	468	563	1.11	18,828	992	350	1.51
2 household members	**HX090**	26,115	300	2,239	1.01	26,207	387	1,288	1.02	25,009	481	785	1.02
3 household members	**HX090**	26,272	419	895	1.09	26,946	587	498	1.08	27,419	738	329	1.05
4 and more	**HX090**	24,583	361	1,468	1.09	24,397	422	861	1.27	24,306	518	533	1.21
all households	**HX090**	22,768	200	5,711	1.18	23,322	255	3,210	1.23	22,623	482	1,997	1.62
Equivalised mean income by age class and by gender													
<25	**HX090**	22,295	279	4,874	1.18	22,910	321	2,708	1.30	23,116	419	1,686	1.30
25 to 34	**HX090**	23,055	399	1,602	1.26	23,018	474	828	1.33	22,154	648	479	1.23
35 to 44	**HX090**	25,279	405	2,286	1.36	25,606	471	1,345	1.55	24,114	1,207	848	2.62
45 to 54	**HX090**	28,631	390	2,276	1.17	29,109	443	1,307	1.15	28,565	648	837	1.18
55 to 64	**HX090**	28,452	428	2,004	1.13	28,779	672	1,216	1.16	28,245	708	736	1.01
65+	**HX090**	18,925	268	1,634	0.99	18,710	271	949	1.02	17,980	312	630	1.10
Male	**HX090**	24,369	228	7,323	1.21	24,628	272	4,198	1.32	24,074	486	2,632	1.74
Female	**HX090**	23,666	197	7,353	1.13	24,084	250	4,155	1.14	23,650	303	2,584	1.15
all persons	**HX090**	24,013	190	14,676	1.17	24,353	224	8,353	1.24	23,861	350	5,216	1.51

ICELAND

Variable		(1) Sample base: Full cross-sectional sample Source: Cross-sectional data 2006				(2) Sample base: 2-yars longitudinal sample (2005-06) Source: Longitudinal data 2006				(3) Sample base: 3-yars longitudinal sample (2004-05-06) Source: Longitudinal data 2006			
		estimate	standard error	sample size	deft	estimate	standard error	sample size	deft	estimate	standard error	sample size	deft
		(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Total household income													
Total household gross income	**HY010**	77,428	1,118	2,838	1.04	77,162	1,872	1,722	1.14	77,466	2,279	1,109	1.14
Total disposable household income	**HY020**	54,520	794	2,838	1.04	54,533	1,377	1,722	1.16	54,876	1,640	1,109	1.16
... excluding transfers except pensions	**HY022**	49,924	791	2,838	1.04	50,701	1,370	1,722	1.16	51,116	1,663	1,109	1.16
... excluding all transfers	**HY023**	44,718	829	2,838	1.03	44,074	1,453	1,722	1.18	44,456	1,801	1,109	1.18
Household level income components													
Property income	**HY040n**												
	HY040g	6,949	672	150	0.93	6,767	1,246	87	1.09	7,005	1,532	56	1.15
Family/Children allowances	**HY050n**												
	HY050g	4,441	221	1,005		3,632	280	619		3,266	263	386	
Other social exclusions	**HY060n**												
	HY060g	3,587	642	65	1.32	2,570	577	29	1.01	2,584	995	15	0.99
Housing allowances	**HY070n**												
	HY070g	2,039	38	1,010	1.09	2,006	59	632	1.13	1,984	56	413	1.08
Inter-household transfers received	**HY080n**												
	HY080g	4,557	195	448	1.20	4,593	265	255	1.14	4,443	313	150	1.17
Capital income	**HY090n**												
	HY090g	7,616	935	1,870	1.04	8,255	1,777	1,174	1.21	9,513	2,046	743	1.20
Mortgage interest	**HY100n**												
	HY100g	6,949	185	2,015	1.08	6,831	339	1,225	1.17	7,171	476	774	1.19
Children's income	**HY110n**												
	HY110g	1,318	105	379	1.11	1,416	166	238	1.24	1,436	227	137	1.28
Regular taxes on wealth	**HY120n**												
	HY120g	1,260	25	2,473	1.02	1,255	32	1,515	1.26	1,267	42	981	1.08
Inter-household transfers paid	**HY130n**												
	HY130g	3,426	150	419	1.07	3,314	237	252	1.14	3,108	272	162	1.13
Tax	**HY140n**												
	HY140g	21,360	395	2,836	1.02	21,096	557	1,721	1.08	21,085	772	1,108	1.05
Tax adjustment	**HY145n**												

Iceland (cont.)

Variable		estimate	standard error	sample size	deft	estimate	standard error	sample size	deft	estimate	standard error	sample size	deft
		(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Personal level income components													
Employee cash or near cash income	**PY010n**												
	PY010g	36,459	465	5,497	1.10	36,839	573	3,228	1.13	37,082	689	2,026	1.06
Non-Cash employee income	**PY020n**												
	PY020g												
Contributions to private pension	**PY035n**												
	PY035g												
Self-employment income	**PY050n**												
	PY050g	12,739	761	678	1.03	12,691	857	426	1.09	12,817	1,187	255	1.08
Production for own consumption	**PY070n**												
	PY070g												
Pension from private plans	**PY080n**												
	PY080g												
Unemployment benefits	**PY090n**												
	PY090g	3,251	214	211	1.12	3,044	318	117	1.18	2,875	360	76	1.00
Old-age benefits	**PY100n**												
	PY100g	18,942	429	740	1.08	18,605	493	479	1.09	18,434	505	313	1.04
Survivor' benefits	**PY110n**												
	PY110g	8,238	1,043	313	1.00	10,668	2,211	170	1.11	8,459	2,243	93	1.15
Sickness benefits	**PY120n**												
	PY120g	1,383	209	12	0.84	1,669	360	10	1.01	1,551	594	5	0.95
Disability benefits	**PY130n**												
	PY130g	17,403	898	298	1.17	16,894	922	176	0.97	16,688	1,227	113	0.96
Education-related allowances	**PY140n**												
	PY140g	3,491	782	216	0.97	3,036	1,076	126	1.41	3,862	1,709	76	1.33
Employees' gross monthly earnings	**PY200g**	4,102	62	4,477	1.11				0.00				0.00
Equivalised mean income by household size													
1 household member	**HX090**	28,209	1,000	383	1.05	28,129	1,586	238	1.24	26,189	1,513	158	1.02
2 household members	**HX090**	34,209	1,211	829	1.04	36,064	2,157	508	1.15	37,780	2,520	347	1.17
3 household members	**HX090**	31,644	702	573	1.02	33,157	910	352	1.03	34,525	1,158	230	1.03
4 and more	**HX090**	31,498	568	1,053	1.10	31,778	715	624	1.10	30,953	699	374	1.07
all households	**HX090**	31,480	444	2,838	1.08	32,309	840	1,722	1.25	32,462	962	1,109	1.19
Equivalised mean income by age class and by gender													
<25	**HX090**	29,311	438	3,489	1.13	30,225	500	2,076	1.10	30,212	629	1,484	1.05
25 to 34	**HX090**	30,069	631	1,051	1.05	31,180	844	651	1.07	31,480	876	412	1.09
35 to 44	**HX090**	31,989	655	1,202	1.08	32,017	709	727	1.13	30,637	769	461	1.08
45 to 54	**HX090**	37,142	1,201	1,258	1.11	36,073	1,051	707	1.08	37,004	1,193	436	1.06
55 to 64	**HX090**	40,902	2,000	791	1.13	44,426	4,661	489	1.23	47,837	5,099	329	1.23
65+	**HX090**	27,738	700	796	1.12	27,556	1,157	513	1.48	26,262	1,114	341	1.02
Male	**HX090**	32,220	438	4,300	1.11	33,076	670	2,593	1.29	33,301	760	1,747	1.23
Female	**HX090**	31,031	425	4,287	1.08	31,968	755	2,570	1.21	31,903	797	1,716	1.22
all persons	**HX090**	31,630	402	8,587	1.09	32,520	677	5,163	1.25	32,594	749	3,463	1.22

SWEDEN

Variable		(1) Sample base: Full cross-sectional sample Source: Cross-sectional data 2006				(2) Sample base: 2-yars longitudinal sample (2005-06) Source: Longitudinal data 2006				(3) Sample base: 3-yars longitudinal sample (2004-05-06) Source: Longitudinal data 2006			
		estimate	standard error	sample size	deft	estimate	standard error	sample size	deft	estimate	standard error	sample size	deft
		(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Total household income													
Total household gross income	**HY010**	39,859	378	6,803	1.03	42,243	483	4,449	1.04	42,844	664	2,903	1.07
Total disposable household income	**HY020**	27,434	225	6,803	1.05	29,222	300	4,449	1.04	29,595	405	2,903	1.06
... excluding transfers except pensions	**HY022**	23,233	224	6,803	1.03	25,191	295	4,449	1.02	25,639	395	2,903	1.04
... excluding all transfers	**HY023**	18,694	237	6,803	1.02	19,674	348	4,449	1.04	19,947	427	2,903	1.06
Household level income components													
Property income	**HY040n**	1,322	125	114	1.06	1,749	268	67	1.05	1,837	397	44	1.05
	HY040g	1,888	179	114	1.06	2,499	383	67	1.05	2,625	567	44	1.05
Family/Children allowances	**HY050n**	3,527	77	2,256	1.07	3,870	86	1,416	1.02	3,737	93	927	0.97
	HY050g	3,991	99	2,256	1.06	4,373	110	1,416	1.02	4,188	117	927	0.96
Other social exclusions	**HY060n**	4,058	316	211	1.03	4,373	435	103	0.98	3,688	546	56	0.98
	HY060g	4,058	316	211	1.03	4,373	435	103	0.98	3,688	546	56	0.98
Housing allowances	**HY070n**	2,192	63	626	1.08	2,278	90	372	1.12	2,189	118	225	1.11
	HY070g	2,192	63	626	1.08	2,278	90	372	1.12	2,189	118	225	1.11
Inter-household transfers received	**HY080n**	1,959	83	347	1.14	2,095	118	211	1.19	2,063	127	147	1.14
	HY080g	1,959	83	347	1.14	2,095	118	211	1.19	2,063	127	147	1.14
Capital income	**HY090n**	906	100	5,203	1.11	1,438	222	3,555	1.06	1,420	268	2,363	1.15
	HY090g	1,294	143	5,203	1.11	2,054	317	3,555	1.06	2,029	383	2,363	1.15
Mortgage interest	**HY100n**	841	18	3,387	1.09	837	18	2,358	1.09	841	28	1,585	1.12
	HY100g	1,201	25	3,387	1.09	1,196	26	2,358	1.09	1,201	40	1,585	1.12
Children's income	**HY110n**	161	26	1,194	1.04	169	27	849	1.02	192	39	569	1.03
	HY110g	199	31	1,194	1.04	207	32	849	1.02	232	46	569	1.03
Regular taxes on wealth	**HY120n**	1,264	46	4,581	1.06	1,290	80	3,053	1.23	1,324	99	1,997	1.33
	HY120g	1,264	46	4,581	1.06	1,290	80	3,053	1.23	1,324	99	1,997	1.33
Inter-household transfers paid	**HY130n**	1,501	87	140	1.08	1,442	119	86	1.09	1,464	150	62	1.11
	HY130g	1,501	87	140	1.08	1,442	119	86	1.09	1,464	150	62	1.11
Tax	**HY140n**	11,944	147	6,681	0.99	12,320	175	4,410	1.02	12,504	235	2,879	1.05
	HY140g	11,944	147	6,681	0.99	12,320	175	4,410	1.02	12,504	235	2,879	1.05
Tax adjustment	**HY145n**												

Sweden (cont.)

Variable		estimate	standard error	sample size	deft	estimate	standard error	sample size	deft	estimate	standard error	sample size	deft
		(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Personal level income components													
Employee cash or near cash income	**PY010n**	15,496	107	9,640	1.07	16,254	131	6,315	1.08	16,391	166	4,129	1.08
	PY010g	22,588	182	9,640	1.07	23,576	215	6,315	1.07	23,818	279	4,129	1.08
Non-Cash employee income	**PY020n**	826	31	2,148	1.04	855	126	1,490	1.04	770	49	996	1.02
	PY020g	1,290	51	2,148	1.03	1,408	266	1,490	1.04	1,210	79	996	1.02
Contributions to private pension	**PY035n**	730	14	4,157	1.06	753	20	2,744	1.07	746	25	1,833	1.03
	PY035g	730	14	4,157	1.06	753	20	2,744	1.07	746	25	1,833	1.03
Self-employment income	**PY050n**	3,647	261	1,839	1.03	3,781	388	1,191	0.99	3,835	438	791	0.99
	PY050g	5,437	393	1,839	1.02	5,417	503	1,191	1.01	5,454	552	791	1.00
Production for own consumption	**PY070n**												
	PY070g												
Pension from private plans	**PY080n**	2,486	130	974	1.17	2,395	133	692	1.08	2,398	178	458	1.08
	PY080g	3,638	210	974	1.18	3,478	204	692	1.08	3,484	268	458	1.08
Unemployment benefits	**PY090n**	4,577	120	1,239	1.13	4,601	157	785	1.11	4,646	173	528	1.09
	PY090g	6,195	166	1,239	1.13	6,172	212	785	1.12	6,217	229	528	1.09
Old-age benefits	**PY100n**	10,628	124	2,728	1.06	11,105	148	1,930	1.03	11,165	171	1,267	1.03
	PY100g	14,775	188	2,728	1.05	15,390	234	1,930	1.03	15,476	275	1,267	1.03
Survivor' benefits	**PY110n**	4,537	233	118	1.13	4,669	284	65	1.14	4,602	374	40	1.12
	PY110g	6,294	351	118	1.13	6,414	430	65	1.12	6,229	550	40	1.13
Sickness benefits	**PY120n**	2,231	82	2,523	1.11	1,974	95	1,704	1.07	1,858	106	1,140	1.07
	PY120g	3,104	116	2,523	1.12	2,738	135	1,704	1.08	2,561	147	1,140	1.07
Disability benefits	**PY130n**	7,867	140	929	1.06	7,903	165	611	1.05	7,764	194	406	1.03
	PY130g	10,615	200	929	1.06	10,599	236	611	1.07	10,412	272	406	1.04
Education-related allowances	**PY140n**	2,834	76	2,053	1.22	2,644	85	1,292	1.21	2,593	124	806	1.21
	PY140g	2,843	77	2,053	1.23	2,654	86	1,292	1.20	2,609	126	806	1.20
Employees' gross monthly earnings	**PY200g**												
Equivalised mean income by household size													
1 household member	**HX090**	15,424	195	1,681	0.99	16,353	268	1,201	1.03	16,500	364	773	1.05
2 household members	**HX090**	21,734	255	2,389	1.05	23,178	322	1,521	0.99	23,518	469	1,009	0.99
3 household members	**HX090**	19,488	257	1,010	1.05	21,047	306	636	1.05	20,830	380	412	1.09
4 and more	**HX090**	17,976	235	1,723	1.31	20,231	298	1,091	1.08	20,118	284	709	1.05
all households	**HX090**	18,219	127	6,803	1.09	19,499	165	4,449	1.06	19,659	242	2,903	1.08
Equivalised mean income by age class and by gender													
<25	**HX090**	16,500	181	6,162	1.23	18,425	184	4,808	1.06	18,541	220	2,985	1.06
25 to 34	**HX090**	18,595	214	2,059	1.09	20,711	287	1,285	1.06	20,829	340	862	1.08
35 to 44	**HX090**	19,008	256	2,388	1.29	21,222	321	1,547	1.06	21,182	305	991	1.08
45 to 54	**HX090**	21,410	285	2,475	1.23	23,197	401	1,550	1.12	23,042	311	1,013	1.13
55 to 64	**HX090**	24,241	413	2,044	1.04	25,098	558	1,365	0.99	25,443	780	905	1.00
65+	**HX090**	16,364	206	2,021	1.09	16,571	293	1,369	1.04	16,832	442	888	1.03
Male	**HX090**	18,894	156	8,452	1.16	20,348	164	6,012	1.05	20,399	209	3,851	1.06
Female	**HX090**	18,504	141	8,697	1.14	19,759	161	5,912	1.10	19,941	241	3,793	1.13
all persons	**HX090**	18,694	135	17,149	1.15	20,050	145	11,924	1.08	20,167	207	7,644	1.10

ESTONIA

Variable		(1) Sample base: Full cross-sectional sample. Source: Cross-sectional data 2006 (note: not computed)				(2) Sample base: 2-years longitudinal sample (2005-06). Source: Longitudinal data 2006				(3) Sample base: 3-years longitudinal sample (2004-05-06). Source: Longitudinal data 2006			
		estimate	standard error	sample size	deft	estimate	standard error	sample size	deft	estimate	standard error	sample size	deft
		(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Total household income													
Total household gross income	**HY010**					8,106	160	3,897	1.44	8,126	155	3,336	1.28
Total disposable household income	**HY020**					6,762	121	3,897	1.42	6,780	118	3,336	1.27
... excluding transfers except pensions	**HY022**					6,349	120	3,897	1.41	6,379	118	3,336	1.29
... excluding all transfers	**HY023**					5,202	127	3,897	1.42	5,222	124	3,336	1.27
Household level income components													
Property income	**HY040n**					811	200	63	1.29	852	212	57	1.26
	HY040g					1,067	263	63	1.29	1,121	279	57	1.26
Family/Children allowances	**HY050n**					643	23	1,598	1.26	615	21	1,378	1.15
	HY050g					695	29	1,598	1.29	657	26	1,378	1.16
Other social exclusions	**HY060n**					82	25	27	1.12	82	29	21	1.06
	HY060g					82	25	27	1.12	82	29	21	1.06
Housing allowances	**HY070n**					391	34	97	0.97	389	39	86	1.19
	HY070g					391	34	97	0.97	389	39	86	1.19
Inter-household transfers received	**HY080n**					915	140	135	1.82	809	101	123	1.71
	HY080g					915	140	135	1.82	809	101	123	1.71
Capital income	**HY090n**					90	32	1,107	1.01	73	24	955	0.88
	HY090g					109	35	1,107	1.26	108	37	955	1.27
Mortgage interest	**HY100n**					798	58	321	1.38	793	63	280	1.43
	HY100g					798	58	321	1.38	793	63	280	1.43
Children's income	**HY110n**					167	35	63	0.96	184	40	56	0.88
	HY110g					169	36	63	0.93	186	41	56	0.85
Regular taxes on wealth	**HY120n**					31	1	2,522	1.21	32	1	2,172	1.19
	HY120g					31	1	2,522	1.21	32	1	2,172	1.19
Inter-household transfers paid	**HY130n**					835	58	185	1.17	850	61	158	1.16
	HY130g					823	57	185	1.17	850	61	158	1.16
Tax	**HY140n**												
	HY140g					1,806	51	2,875	1.67	1,803	50	2,486	1.48
Tax adjustment	**HY145n**					-149	6	1,346	1.07	-150	7	1,152	1.07

Estonia (cont.)

Variable		estimate	standard error	sample size	deft	estimate	standard error	sample size	deft	estimate	standard error	sample size	deft
		(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Personal level income components													
Employee cash or near cash income	**PY010n**					4,422	80	4,932	1.48	4,412	81	4,279	1.52
	PY010g					5,517	107	4,932	1.53	5,501	106	4,279	1.52
Non-Cash employee income	**PY020n**					1,052	71	165	1.38	1,037	79	142	1.46
	PY020g					1,423	99	165	1.34	1,387	108	142	1.45
Contributions to private pension	**PY035n**					322	17	479	1.33	319	18	413	1.32
	PY035g					322	17	479	1.33	319	18	413	1.32
Self-employment income	**PY050n**					572	50	713	1.08	532	57	631	1.07
	PY050g					769	70	717	1.14	729	82	634	1.13
Production for own consumption	**PY070n**												
	PY070g												
Pension from private plans	**PY080n**					1,351	1,295	4	1.16	1,905	1,799	3	1.05
	PY080g					1,501	1,439	4	1.16	2,116	1,999	3	1.05
Unemployment benefits	**PY090n**					470	64	133	1.38	473	77	110	1.26
	PY090g					559	80	133	1.38	561	96	110	1.25
Old-age benefits	**PY100n**					2,064	15	2,334	1.17	2,064	12	1,983	0.91
	PY100g					2,093	26	2,334	1.28	2,084	15	1,983	0.88
Survivor' benefits	**PY110n**					737	49	84	1.16	720	48	76	1.16
	PY110g					737	49	84	1.16	720	48	76	1.16
Sickness benefits	**PY120n**					235	26	594	1.49	240	30	511	1.56
	PY120g					276	32	594	1.50	281	37	511	1.58
Disability benefits	**PY130n**					1,122	25	596	1.23	1,118	27	514	1.07
	PY130g					1,122	25	596	1.23	1,118	27	514	1.07
Education-related allowances	**PY140n**					708	222	182	2.66	728	271	156	2.87
	PY140g					708	222	182	2.66	728	271	156	2.87
Employees' gross monthly earnings	**PY200g**												
Equivalised mean income by household size													
1 household member	**HX090**					3,250	139	783	1.47	3,306	152	673	1.58
2 household members	**HX090**					4,225	133	1,113	1.54	4,158	112	914	1.40
3 household members	**HX090**					4,811	160	834	1.40	4,913	187	729	1.39
4 and more	**HX090**					4,475	88	1,164	1.21	4,423	96	1,018	1.21
all households	**HX090**					4,063	76	3,894	1.61	4,069	68	3,334	1.50
Equivalised mean income by age class and by gender													
<25	**HX090**					4,285	83	4,089	1.57	4,311	94	3,566	1.59
25 to 34	**HX090**					5,544	207	1,074	1.31	5,492	214	929	1.30
35 to 44	**HX090**					4,569	114	1,466	1.48	4,564	121	1,250	1.36
45 to 54	**HX090**					4,357	94	1,599	1.18	4,403	102	1,382	1.20
55 to 64	**HX090**					4,278	139	1,261	1.35	4,310	114	1,086	1.39
65+	**HX090**					3,132	48	1,781	1.11	3,112	49	1,511	1.21
Male	**HX090**					4,477	83	5,277	1.65	4,476	84	4,534	1.62
Female	**HX090**					4,137	60	5,993	1.48	4,146	60	5,190	1.38
all persons	**HX090**					4,292	66	11,270	1.55	4,296	65	9,724	1.52

LITHUANIA

Variable		(1) Sample base: Full cross-sectional sample. Source: Cross-sectional data 2006 (note: not computed)				(2) Sample base: 2-years longitudinal sample (2005-06). Source: Longitudinal data 2006				(3) Sample base: 3-years longitudinal sample (2004-05-06). Source: Longitudinal data 2006			
		estimate	standard error	sample size	deft	estimate	standard error	sample size	deft	estimate	standard error	sample size	deft
		(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Total household income													
Total household gross income	**HY010**					5,807	119	2,971	1.10				
Total disposable household income	**HY020**					4,797	90	2,971	1.10				
... excluding transfers except pensions	**HY022**					4,496	90	2,971	1.11				
... excluding all transfers	**HY023**					3,624	94	2,971	1.13				
Household level income components													
Property income	**HY040n**					549	92	154	1.18				
	HY040g					618	106	154	1.18				
Family/Children allowances	**HY050n**					636	38	419	1.06				
	HY050g					675	44	419	1.05				
Other social exclusions	**HY060n**												
	HY060g					316	47	96	1.14				
Housing allowances	**HY070n**												
	HY070g					90	6	130	0.87				
Inter-household transfers received	**HY080n**												
	HY080g					739	61	215	1.17				
Capital income	**HY090n**					699	223	98	1.05				
	HY090g					800	262	98	1.05				
Mortgage interest	**HY100n**												
	HY100g					831	149	71	1.32				
Children's income	**HY110n**					155	99	4	1.20				
	HY110g					155	99	4	1.20				
Regular taxes on wealth	**HY120n**												
	HY120g					19	1	582	0.99				
Inter-household transfers paid	**HY130n**												
	HY130g					685	53	248	1.15				
Tax	**HY140n**												
	HY140g					1,526	46	1,941	1.13				
Tax adjustment	**HY145n**					-191	9	424	0.93				

Lithuania (cont.)

Variable		estimate	standard error	sample size	deft	estimate	standard error	sample size	deft	estimate	standard error	sample size	deft
		(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Personal level income components													
Employee cash or near cash income	**PY010n**					3,153	59	3,119	1.17				
	PY010g					4,113	83	3,119	1.13				
Non-Cash employee income	**PY020n**												
	PY020g					493	55	66	1.16				
Contributions to private pension	**PY035n**												
	PY035g					156	22	92	1.11				
Self-employment income	**PY050n**					2,238	132	615	1.15				
	PY050g					2,346	139	615	1.19				
Production for own consumption	**PY070n**												
	PY070g												
Pension from private plans	**PY080n**												
	PY080g												
Unemployment benefits	**PY090n**					520	44	101	1.18				
	PY090g					532	46	101	1.17				
Old-age benefits	**PY100n**												
	PY100g					1,584	14	1,789	1.12				
Survivor' benefits	**PY110n**												
	PY110g					524	33	160	1.14				
Sickness benefits	**PY120n**												
	PY120g												
Disability benefits	**PY130n**												
	PY130g					1,161	27	446	0.90				
Education-related allowances	**PY140n**												
	PY140g					374	45	225	1.22				
Employees' gross monthly earnings	**PY200g**												
Equivalised mean income by household size													
1 household member	**HX090**					2,060	73	630	1.23				
2 household members	**HX090**					3,161	100	917	1.20				
3 household members	**HX090**					3,281	117	635	1.21				
4 and more	**HX090**					2,935	87	789	1.44				
all households	**HX090**					2,799	46	2,971	1.13				
Equivalised mean income by age class and by gender													
<25	**HX090**					2,846	63	2,726	1.18				
25 to 34	**HX090**					3,454	129	732	1.16				
35 to 44	**HX090**					2,956	81	1,143	1.04				
45 to 54	**HX090**					3,342	91	1,264	1.17				
55 to 64	**HX090**					3,037	93	996	1.13				
65+	**HX090**					2,301	48	1,344	0.92				
Male	**HX090**					3,008	53	3,819	1.13				
Female	**HX090**					2,866	49	4,386	1.12				
all persons	**HX090**					2,932	47	8,205	1.11				

Annex 3: Non-sampling errors

Types of errors in survey data

It is necessary to begin with an explanation of concepts concerning errors in survey data.

All statistical data, from whatever source and whatever the manner of their collection, are potentially subject to errors of various types. It is important that the results of surveys are accompanied by descriptions of their quality and limitations.

Firstly, knowledge about data quality is required for their *proper use and interpretation.* This knowledge is essential in determining whether and with what degree of confidence the patterns observed in the results are real, and not merely products of the variability and deficiency inherent in the data. Information on the nature and magnitude of errors can also be useful for making appropriate corrections to the data or adjustments in their interpretation.

Secondly, measures of data quality are important for the evaluation and improvement of survey design and procedures. A detailed investigation of the sources, magnitude and impact of errors is necessary to identify how survey design and procedures may be improved and resources allocated more efficiently among various aspects of the survey operation.

Continued monitoring and improvement of data quality is particularly important major continuous or repeated surveys such as EU-SILC.

The objective of a sample survey is to make estimates or inferences of general applicability for a study population, derived from observations made on a limited number (a sample) of units in the population. We can distinguish between two groups of errors affecting this process:

(a) Errors in measurement

> These arise from the fact that what is measured on the units included in the survey can depart from the actual (true) values for those units. Errors in measurement centre on *substantive content of the survey*: definition of the survey objectives and questions; ability and willingness of the respondent to provide the information sought; the quality of data collection, coding editing, processing etc..

(b) Errors in estimation

> These are errors in the process of extrapolation from the particular units enumerated to the entire study population for which estimates or inferences are required. These centre on the *process of sample design and implementation*, and include errors of coverage, sample selection, sample implementation and non-response, as well as sampling errors and estimation bias.

Group (a) concerns the accuracy of measurement at the level of individual units enumerated in the survey: how the value has reported by the respondent, and recorded, coded, edited, corrected, imputed and tabulated by the survey workers, may depart from the actual value for the unit concerned. This group of errors can be studied in relation to the various stages of the survey operation: data collection, processing, analysis etc..

Group (b), which concerns the legitimacy of generalisation from the units observed to the target population, includes sampling variability and various biases associated with sample selection and implementation, such as coverage, selection and non-response errors.

The above categorisation, based on operational considerations, is more fundamental than the distinction usually made between *sampling* and *non-sampling* errors. Each group of errors may be further classified in as much detail as possible to identify *specific* sources of error, so as to facilitate their assessment and control.

Variable error and bias

The distinction between variable error and bias is useful because the two components differ in their sources, methods of assessment and control, and impact on the survey results.

Some of the conditions under which the survey is taken are 'essential' to the situation. In addition, survey results are also influenced by transient or chance factors. On this basis it is useful in practice to distinguish between two components into which any particular type of error may be decomposed: (i) a *variable component*, and (ii) *bias*. The underlying idea is that of possible repetitions of the same procedure or operation under essentially the same conditions. The result of the repetitions are affected by random factors, as well as by systematic factors which arise from the conditions under which repetitions are undertaken and affect the results of all repetitions in essentially the same way. The *variable component* of an error arises from chance factors affecting different samples and repetitions of the survey differently. *Bias* arises from factors which are a part of the essential conditions and affect all repetitions in more or less the same way.

Types of errors in surveys

Errors in measurement

1 conceptual errors

- errors in basic concepts, definitions, and classifications
- errors in putting them into practice (questionnaire design, interviewers training and instructions)

2 response errors

- response bias
- simple response variance
- correlated response variance

3 processing errors

- editing errors
- coding and data entry errors
- programming errors, etc.

Mixed category

4 item non-response

- don't knows
- refusals, etc.

Errors in estimation

5 coverage and related errors

- under-coverage
- over-coverage
- sample selection errors

6 unit non-response

- refusals
- inaccessible
- not-at-homes, etc.

7 sampling error

- sampling variance
- estimation bias

Non-sampling errors = 1 to 6

The concepts of unit and item non-response

The term non-response encompasses a wide variety of reasons for non-observation. Non-response means failure to obtain a measurement on one or more study variables for one or more sample units. Non-response errors occur when the survey fails to get a response to some or all of the questions. Non-response causes both an increase in variance, due to the decrease in the effective sample size and/or due to the use of imputation and, more importantly, causes bias as the non-respondents and respondents generally differ with respect to the characteristic of interest.

Non-response is a potential source of bias particularly if the missing data mechanism is not what has been termed as 'Missing At Random'. For instance, one might expect persons with high incomes to be more reluctant to give income information in an interview, thus rendering the upper income class under-represented in the sample and the estimates downwardly biased.

Two categories of non-response can be distinguished:

Unit non-response:

> This refers to the type of non-response in which no information is available from eligible sample units for such reasons as: "impossible to contact", "not at home" (in these two cases contact with the selected element is never established), "unable to answer", "incapacity", "hard core refusal", "inaccessible", or "unreturned questionnaire". It may also happen that a person in a household refuses to co-operate although the household interview has been accepted ('individual' non-response).

Item non-response:

> This refers to the type of non-response in which sufficient information has been provided in the interview for it to be retained in the data base, but the required

information is missing on some particular items. Often this happens in questions the interviewee does not answer because he/she considers them personal or not easily understandable.

Item non-response is the intermediate category between 'errors in measurement' and 'errors in estimation' as defined above. Like other measurement errors, item non-response is subject-matter specific – it occurs to different degrees in different types of questions. At the same time, item non-response is simply additive to the unit non-response in any analysis involving the item concerned. The two together constitute the total non-response level for the item.

Methods of assessment

Indicators or measures of quality of survey data may be obtained by a variety of methods. Some procedures can yield quantitative information on the magnitude and impact of specific types of error, while others provide only qualitative indicators. Though the appropriateness of a method will depend on the specific source and type of error, the various phases of a survey are closely related. Therefore *errors cannot always be attributed to a particular type or source*. The same or similar methods of assessment/control may indeed be suitable for measuring more than one type of error, and some of the indicators obtained may provide no more than general or overall measures of data accuracy without being able to identify specific sources and types of error.

Scope of this report

The following sections provide summary information on main components of non-sampling errors in EU-SILC longitudinal data for 2006. We begin in sub-section 1 with coverage and related errors related to the sampling frame. This information tends to be stable over years except in the very rare situation when major changes are introduced in ongoing EU-SILC operations. Much of this information has already been reported in previous quality reports, in particular the Intermediate Quality Report for 2006. Therefore the presentation below will be brief. The longitudinal data can cover up to 4 years 2003-2006. Since most surveys began a year later, the data generally cover 3 years 2004-2006, and only 2005-2006 in countries which started in 2005. Sub-section 2 clarifies the structure of the longitudinal sample. This is particularly important in the discussion of unit non-response in sub-section 3. This section considers in some detail this major potential source of 'estimation error' (as defined above). Both cross-sectional and longitudinal rates of unit non-response are discussed. Next we consider item non-response in Section sub-section 4, which is also a major problem, especially concerning income variables in countries where this information is obtained through personal interviews. In the context of item non-response, we discuss some information on the procedures and extent of imputation, as well as on the net-to-gross conversion of income components. The procedures of imputation and micro-simulation applied to missing income data provide a link between item non-response and measurement errors. National information on these is presented in sub-section 5.

Relatively limited information has been recorded on measurement errors. These, including data collection and processing errors, are described in sub-section 6. Any methodological studies undertaken in order to assess the magnitude or impact of response and processing errors are noted.

Coverage errors

The target population is the set of elements for which estimates are required while the frame population is composed of the units which are eligible for inclusion through a given sampling procedure. Coverage errors arise from discrepancies between the target and the frame populations, and also from errors in selecting the sample from the frame. The condition of 'probability sampling' is violated if: (a) the survey population is not fully and correctly represented in the sampling frame; (b) the selection of units from the frame into the sample is not according to procedures specified in the sample design; or (c) not all the units selected into the sample are successfully enumerated.

Coverage error concerns primarily (a), but also (b). Errors of coverage arise in circumstances like the following:

- Some units in the target population are missing from the frame. This is under-coverage: the missed units have no chance of being selected into any sample.
- Some units in the sampling frame are not in the target population. This results in over-coverage, unless such units can be identified and eliminated after selection.
- Some units in the target population appear more than once in the frame ('duplication').

In a multi-stage sample, coverage error can arise at any of the stages. For example, while the list of area units in the frame can be expected to be complete, serious coverage error can arise in the delineation of *boundaries* of area units. New units and units in sparsely populated areas may be left out of the frame. Errors in list of ultimate sampling units arise because of changes in those units. List of addresses are less durable than frames of area units, and lists of households less durable than addresses, dwellings or other structural units, and lists of persons even less so. The most common problem with list frames concerns *under-coverage*. *Over-coverage* can also occur (though less commonly than under-coverage) if (a) some units appear in the list more than once (without being so identified for appropriate correction of selection probabilities); or (b) units out-of-scope of the survey are included, but not identified as such and removed during fieldwork; (c) units outside the boundaries of sample are included.

The bias resulting from under-coverage may be summarised as follows:

1. In estimating population total counts, the effect of coverage error is direct and of similar relative magnitude.
2. In estimating total values, the effect will depend on the relative value of the units missed: it will be proportionately larger if the units with above-average values tend to be missed more often, and vice versa.
3. The effects are usually less drastic when estimating statistics such as proportions, means, rates and ratios: here the resulting bias depends on the differences in characteristics of the units covered and the units not covered.
4. Regarding differences and comparisons between population subgroups, the resulting bias depends on the net algebraic difference in the biases for the groups being compared: biases can cancel out to the extent they are common or similar.

Neither the magnitude nor the effect of coverage errors is easy to estimate because it requires information not only external to the sample but also, by definition, *external to the sampling frame*.

Sample selection and implementation errors are distinguished from coverage errors proper in that the latter concern shortcomings of the frame and what remains outside the frame, while sample selection and implementation errors refer to losses and distortions within the sampling frame. Examples are incorrect application of the selection procedures and selection probabilities, and more importantly, inappropriate substitution of the selected units by others during field work.

Common problems with list frames

Completeness of the frame is a most critical requirement (and perhaps also the most common problem) of list frames. Occasionally it is also important that the list contains pertinent and accurate information on the size and other characteristics of individual units so as to permit efficient stratification and control of the selection process. Problems can arise in the absence of one-to-one correspondence between listings used for sample selection, and the elementary units which are interest in the survey. The lack of correspondence can arise in several forms, such as the following.

- Presence of blanks in the lists, that is listing representing no real units.
- Duplications in the list, meaning that the same unit is represented by more than one listing.
- Clustering of elements, meaning that more than one unit may be represented by the same listing.
- Under-coverage, referring to units not represented in the frame; this is the most serious and difficult problem and biases the results of many surveys. No simple or cheap solutions to the problem of under-coverage exist.
- Failure to locate units - the failure to identify which unit(s) a selected listing represents.
- Changes in units and unit characteristics.

Errors in measurement

As noted before, the broad range of 'errors in measurement' refers to the problem that what is measured on the units included in the survey can depart from the true values for those units. These errors centre on *substantive aspects* such as definition of the survey objectives, formulating questions, ability and willingness of the respondent to provide the information sought, and the quality of data collection and processing. These affect the accuracy of measurement at the level of individual units enumerated in the survey. This group of errors can be studied in relation to the various stages of the survey operations. From the point of survey operation and methods of assessing and controlling these errors, it is useful to divide them into two categories: the so-called 'measurement errors' concerning the process of data collection, and 'processing errors' concerning the subsequent process of transforming the data to the form of a micro database suitable for analysis. This distinction is made in the Commission Regulation on quality reports.

Despite this operational distinction, the two classes of error have great conceptual similarity. In this section, we first discuss the conceptual basis common to both these classes of 'errors in measurement'.

Measurement biases

Measurement biases refer to the more or less systematic errors in obtaining the required information. They arise from shortcomings *affecting the whole survey operation*: basic conceptual errors in defining and operationalising the survey content; any incorrect instruction affecting all the survey workers; errors in the coding frame or programs for processing the data, etc. They also arise from inherent difficulties - more or less independent of the specific technical design and procedures of the survey - in collecting certain types of information (such as income in EU-SILC interviews), given the general social situation and the type of respondents involved.

The assessment of measurement biases requires analysis of internal and external consistency of the data, comparison with models and other sources, with measurements using alternative and improved procedures, and in general terms, a thorough understanding of the subject matter and practical conditions of data collection of the survey. The first step in identifying bias is through logical and substantive analysis of consistency and relationships in the data, against external standards and prior knowledge of the subject.

Beyond that, the assessment requires comparison with more accurate data: from some existing external source, and/or collected with special, improved methods. There are several possibilities in connection with such assessment. For instance, the study response bias may involve two interviews on a subsample following the original interview. These would consist of a *re-interview*, which is an independent replication of the original interview and is aimed at measuring response variance; followed in discrepant cases by a *reconciliation interview* aimed at establishing correct responses and identifying biases and their sources.

Measurement variance

These refer to variable errors in data collection (response or interviewer variance), and similar errors in data processing (coding, data entry etc.). The following discussion in terms of response variance also applies to other sources of measurement variance.

Two components of response variance may be distinguished: *simple response variance*; and *correlated response variance*. The decomposition of the total response error into components is based on the following concept.

(i) A part of the error is common to the work of all interviewers; this is the *response bias*.

(ii) In addition, each interviewer has his/her own particular bias, which affects the interviewer's whole work load; this is the correlated response variance component. By definition, its expected value averaged over all interviewers (of the type employed in the survey) is zero. This is the *correlated response variance.*

(iii) The third component - simple response variance - is random, not correlated with any particular interviewer.

This distinction is useful because the components differ in nature and method of assessment and control. This is the *simple response variance.*

As already noted, the bias component is a product of the basic survey design, procedures and conditions.

Correlated variance indicates lack of uniformity and standardisation in the interviewers' work. Its high value indicates the need for better training and supervision of survey work. Its magnitude also depends on the number of interviewers engaged in the survey (just as sampling error depends on the size of the sample). And just like the computation of sampling error, the estimation of its magnitude requires comparisons between different replications of the sample, here the basic unit of comparison being the individual interviewer work loads, just as the sample areas may form the basic components in computing sampling error.

Simple response variance, by contrast, is an indicator of the inherent instability of particular items in the questionnaire: it indicated that the information obtained is not sufficiently repeatable, hence not reliable. Its measurement requires comparisons between independent repetitions of the survey under the same general conditions. There is no way, in a single survey, to distinguish between variation among the true values of units (which gives rise to sampling error), and the additional variability arising from random factors affecting individual responses. In fact, the usual procedures for estimating sampling error automatically include the full effect of the simple response variance component. Separate estimation of this component requires a re-interview survey, independent of the original survey but under the same conditions and using the same procedures.

Technical note on the computation of longitudinal response and follow-up rates

In order to calculate the *household response rates* required for the longitudinal component we should have the distribution of the following set of key variables for the second and following wave of the EU-SILC longitudinal component[25]:

- DB110 household status
- DB120 contact at address
- DB130 household questionnaire result
- DB135 household interview acceptance

In fact, comparing the result code of these variables from wave t and (t-1) we can define the dimension of the following groups of household:

- The household passing from wave t to (t+1);
- The newly created or added in t passing to wave (t+1)
- The household no to be passed from wave t to (t+1)

With this clear definition we could compute the following measures:

- Wave response rate: Percentage of households successfully interviewed (DB135=1) which were passed on to wave t (from wave t-1) or newly created or added during wave t, excluding those out of scope (under the tracing rules) or non-existent.
- Longitudinal follow-up rate: Percentage of households which are passed on to wave t+1 for follow-up within the households received into wave t from wave t-1, excluding those out of scope (under the tracing rules) or non-existent.

[25] Variables DB120, DB130 and DB135 are not a part of the UDB disseminated to the EU-SILC research community.

- Follow-up ratio: Number of households passed on from wave t to wave t+1 in comparison to the number of households received for follow-up at wave t from wave t-1
- Achieved sample size ratio: Ratio of the number of households accepted for the database (DB135=1) in wave t to the number of households accepted for the database (DB135=1) in wave t-1.

In order to calculate the personal response rates required for the longitudinal component it is needed the distribution of the following set of key variables for each panel and wave of EU-SILC longitudinal component:

- RB110 membership status of the person
- RB120 person moving out
- RB250 respondent status

In fact, comparing the personal interview outcome in wave t for the sample persons forwarded from last wave t-1 the following can be computed.

- Wave response rate: Percentage of sample persons successfully interviewed (RB250=11, 12, 13) among those passed on to wave t (from wave t-1) or newly created or added during wave t, excluding those out of scope (under the tracing rules).
- Percentage of co-residents selected in wave 1 successfully interviewed (RB250=11, 12, 13) among those passed on to wave t (from wave t-1).
- Longitudinal follow-up rate: Percentage of sample persons successfully interviewed (RB250=11, 12, 13) in wave t out of all of sample persons selected, excluding those who have died or been found ineligible (out of scope), breakdown by causes of non-response.
- Achieved sample size ratio: Ratio of the number of completed personal interviews (RB250=11, 12, and 13) in wave t to the number of completed personal interviews in wave t-1.
- This ratio will be defined for sample persons and for all persons including non-sample persons aged 16+ and for co-residents aged 16+ selected in first wave
- Response rate for non-sample persons: Ratio of the number of completed personal interviews (RB250=11, 12, 13) of non-sample persons aged 16+ in wave t to all non-sample persons aged 16+ listed in the households accepted for the database (DB135=1) in wave t or listed in the most recently conducted household interviews for households, which were forwarded from wave t-1 to wave t for follow-up, but could not be successfully interviewed in wave t.

Annex 4: Mode of data collection

Table 19: Mode of data collection (2006)

	PAPI		CAPI		CATI		Self-administered	
	X	L	X	L	X	L	X	L
Belgium	0	0	100	100	0	0	0	0
Czech Republic	99.46	99.33	0	0	0	0	0.54	0.67
Denmark	0	0	0	0	94.61	93.82	5.39	6.18
Germany	0	0.02	0	0	0	0	100	99.98
Estonia	1.71	36.32	98.08	63.00	0.17	0.22	0.04	0.47
Ireland	0	0	100	100	0	0	0	0
Greece	72.01	65.66	25.43	32.15	2.57	2.12	0	0.06
Spain	0.00	28.07	93.25	66.57	6.75	4.58	0	0.78
France	0	0	100	100	0	0	0	0
Italy	100	100	0	0	0	0	0	0
Cyprus	0.03	0.21	99.97	99.79	0	0	0	0
Latvia	13.37	60.73	81.88	36.13	4.65	2.51	0.09	0.63
Lithuania	95.70	96.38	0	0	2.67	1.93	1.63	1.69
Luxembourg	100	100	0	0	0	0	0	0
Hungary	100	100	0	0	0	0	0	0
Malta	0	6.30	100	93.70	0	0	0	0
The Netherlands	0	0	0	0	100	100	0	0
Austria	0	0	99.35	97.62	0.65	2.38	0	0
Poland	100	100	0	0	0	0	0	0
Portugal	8.56	8.63	91.44	91.37	0	0	0	0
Slovenia	0	57.49	47.40	9.27	52.60	33.24	0	0
Slovakia	99.46	99.42	0	0	0	0	0.54	0.58
Finland	0	0	2.99	2.60	97.01	97.40	0	0
Sweden	0.05	0.03	0	0	99.95	99.97	0	0
United Kingdom	0	98.74	100.00	1.26	0	0	0	0
Iceland	0	0	0	0	100	100	0	0
Norway	0	0	0.63	0.51	99.37	99.49	0	0

Source: Micro-database (March 2009).

PAPI: Paper Assisted Personal Interview; CAPI: Computer Assisted Personal Interview; CATI: Computer Assisted Telephone Interview; X: Cross-sectional; L: Longitudinal.

Annex 5: Timeliness and punctuality

Table 20: Follow-up cross-sectional data (2006)

	Regulation deadline	Contractual deadline	First transmission	Number of transmissions	Last transmission
Belgium	01/10/2007	30/11/2007	23/11/2007	2	18/12/2007
Czech Republic	01/10/2007	30/09/2007	**04/10/2007**	2	12/11/2007
Denmark	30/11/2007	01/10/2007	**03/12/2007**	2	04/12/2007
Germany	01/10/2007	30/10/2007	08/10/2007	8	07/12/2007
Estonia	01/10/2007	30/08/2007	23/07/2007	2	31/07/2007
Ireland	30/11/2007	30/10/2007	01/11/2007	1	08/11/2007
Greece	01/10/2007	01/10/2007	01/10/2007	2	08/10/2007
Spain	01/10/2007	30/08/2007	**18/09/2007**	2	07/11/2007
France	01/10/2007	30/10/2007	28/09/2007	5	12/11/2007
Italy	01/10/2007	30/11/2007	**07/12/2007**	3	29/01/2008
Cyprus	01/10/2007	30/10/2007	**26/11/2007**	2	05/12/2007
Latvia	01/10/2007	01/10/2007	01/10/2007	2	02/11/2007
Lithuania	01/10/2007	01/08/2007	31/07/2007	1	31/07/2007
Luxembourg	01/10/2007	30/09/2007	20/09/2007	3	26/10/2007
Hungary	01/10/2007	15/06/2007	18/06/2007	4	12/11/2007
Malta	01/10/2007	30/09/2007	**09/11/2007**	2	16/11/2007
The Netherlands	30/11/2007	01/10/2007	02/10/2007	2	19/11/2007
Austria	01/10/2007	31/07/2007	31/07/2007	2	14/09/2007
Poland	01/10/2007	31/07/2007	31/07/2007	4	11/12/2007
Portugal	01/10/2007	30/09/2007	**02/11/2007**	3	12/12/2007
Slovenia	30/11/2007	30/11/2007	23/11/2007	2	12/12/2007
Slovakia	01/10/2007	30/11/2007	30/03/2007	2	01/06/2007
Finland	30/11/2007	30/08/2007	11/06/2007	1	11/06/2007
Sweden	30/11/2007	30/11/2007	01/10/2007	3	01/11/2007
United Kingdom	30/11/2007	30/11/2007	29/11/2007	3	18/12/2007
Iceland	30/11/2007	30/11/2007	05/11/2007	4	07/12/2007
Norway	30/11/2007	30/11/2007	04/10/2007	4	03/12/2007

Source: eDamis and Regulation (EC) No 1177/2003.

Table 21: Follow-up longitudinal data (2006)

	Regulation deadline	Contractual deadline	First transmission	Number of transmissions	Last transmission
Belgium	31/03/2008	31/03/2008	**14/04/2008**	6	29/09/2008
Czech Republic	31/03/2008	31/03/2008	**08/04/2008**	5	29/07/2008
Denmark	31/03/2008	31/03/2008	**13/05/2008**	3	13/07/2008
Germany	31/03/2008	01/02/2008	**03/04/2008**	7	10/08/2008
Estonia	31/03/2008	31/03/2008	31/03/2008	3	14/07/2008

	Regulation deadline	Contractual deadline	First transmission	Number of transmissions	Last transmission
Ireland	31/03/2008	31/12/2007	**31/03/2008**	5	28/11/2008
Greece	31/03/2008	01/03/2008	**21/03/2008**	2	25/09/2008
Spain	31/03/2008	31/03/2008	28/03/2008	6	29/07/2008
France	31/03/2008	31/03/2008	**11/04/2008**	11	13/01/2009
Italy	31/03/2008	31/03/2008	31/03/2008	3	07/07/2008
Cyprus	31/03/2008	31/03/2008	**10/04/2008**	6	11/08/2008
Latvia	31/03/2008	31/03/2008	31/03/2008	4	10/09/2008
Lithuania	31/03/2008	29/02/2008	29/02/2008	3	26/09/2008
Luxembourg	31/03/2008	31/03/2008	03/03/2008	4	27/08/2008
Hungary	31/03/2008	31/03/2008	31/03/2008	5	18/08/2008
Malta	31/03/2008	No Date	03/03/2008	4	18/07/2008
The Netherlands	31/03/2008	01/10/2007	01/10/2007	5	30/09/2008
Austria	31/03/2008	31/01/2008	01/02/2008	3	20/06/2008
Poland	31/03/2008	31/10/2007	31/10/2007	5	08/10/2008
Portugal	31/03/2008	31/03/2008	31/03/2008	5	05/08/2008
Slovenia	31/03/2008	31/03/2008	31/03/2008	5	07/10/2008
Slovakia	31/03/2008	30/07/2007	21/06/2007	4	18/08/2008
Finland	31/03/2008	31/03/2008	31/03/2008	5	12/08/2008
Sweden	31/03/2008	31/03/2008	31/03/2008	5	16/09/2008
United Kingdom	31/03/2008	31/03/2008	**11/04/2008**	5	22/01/2009
Iceland	31/03/2008	31/03/2008	**21/04/2008**	10	04/11/2008
Norway	31/03/2008	31/03/2008	31/03/2008	6	10/10/2008

Source: eDamis and Regulation (EC) No 1177/2003.

Table 22: Follow-up final national quality reports (2006)

	Regulation deadline	Contractual deadline	First version	Last version
Belgium	31/12/2008	31/12/2008	**25/02/2009**	
Czech Republic	31/12/2008	31/12/2008	09/01/2009	
Denmark	31/12/2008	31/12/2008	15/01/2009	
Germany	31/12/2008	31/12/2008	29/12/2008	
Estonia	31/12/2008	31/12/2008	30/12/2008	
Ireland	31/12/2008	**31/10/2007**	**18/06/2008**	19/03/2009
Greece	31/12/2008	31/12/2008	09/12/2008	
Spain	31/12/2008	31/12/2008	15/12/2008	
France	31/12/2008	31/12/2008	06/01/2009	
Italy	31/12/2008	31/12/2008	07/01/2009	
Cyprus	31/12/2008	31/12/2008	02/01/2009	23/01/2009
Latvia	31/12/2008	31/12/2008	29/12/2008	
Lithuania	31/12/2008	31/12/2008	31/12/2008	

	Regulation deadline	Contractual deadline	First version	Last version
Luxembourg	31/12/2008	**30/11/2008**	**29/12/2008**	19/01/2009
Hungary	31/12/2008	31/12/2008	16/12/2008	
Malta	31/12/2008	31/12/2008	22/12/2008	16/02/2009
The Netherlands	31/12/2008	31/12/2008	**13/02/2009**	
Austria	31/12/2008	**31/10/2008**	**19/11/2008**	
Poland	31/12/2008	**30/11/2008**	01/12/2008	
Portugal	31/12/2008	31/12/2008	31/12/2008	
Slovenia	31/12/2008	31/12/2008	30/12/2008	21/01/2009
Slovakia	31/12/2008	**30/07/2008**	17/07/2008	28/01/2009
Finland	31/12/2008	31/12/2008	**05/02/2009**	09/02/2009
Sweden	31/12/2008	31/12/2008	22/12/2008	
United Kingdom	31/12/2008	31/12/2008	**02/02/2009**	
Iceland	31/12/2008	31/12/2008	12/12/2008	06/02/2009
Norway	31/12/2008	31/12/2008	**15/01/2009**	10/02/2009

Source: eDamis and e-mails.

Annex 6: Basic concepts and reference periods

Table 23: Basic concepts and definitions: are the standard EU-SILC definitions used? (2006)

	Reference population	Private household definition	Household membership
Belgium	F	F	F
Czech Republic	F	F	F
Denmark	F	F	F
Germany	F	F	F
Estonia	F	F	F
Ireland	F	F	F
Greece	F	F	F
Spain	F	F	L
France	F	F	F
Italy	F	L	L
Cyprus	F	F	F
Latvia	F	F	F
Lithuania	F	F	F
Luxembourg	F	F	F
Hungary	F	F	F
Malta	F	F	F
The Netherlands	F	F	F
Austria	F	L	L
Poland	F	F	F
Portugal	F	F	L
Slovenia	F	F	F
Slovakia	F	F	F
Finland	F	F	F
Sweden	F	F	F
United Kingdom	F	L	L
Iceland	F	F	F
Norway	F	F	F

Source: National Quality Reports 2006.
F (fully comparable); L (largely comparable); P (partly comparable); N (not comparable).

Table 24: Reference period (2006)

	Income reference period	Reference period for taxes on income and social insurance contributions	Reference period for taxes on wealth
Belgium	2005	2005	NA
Czech Republic	2005	2005	2005
Denmark	2005	2005	2005
Germany	2005	2005	2005
Estonia	2005	2005	2005
Ireland	12 months prior interview date	12 months prior interview date	NA
Greece	2005	2005	2005
Spain	2005	2005	2005
France	2005	2005	2005
Italy	2005	NA	2005
Cyprus	2005	2005	2005
Latvia	2005	Not collected	2005
Lithuania	2005	2005	2005
Luxembourg	2005	2005	2005
Hungary	2005	2005	2005
Malta	2005	2005	NA
The Netherlands	2005	2005	NA
Austria	2005	2005	NA
Poland	2005	2005	2005
Portugal	2005	NA	2005
Slovenia	2005	2005	2005
Slovakia	2005	2005	2005
Finland	2005	2005	2005
Sweden	2005	2005	No information
United Kingdom	Centred around interview date	Centred around interview date	Financial year Apr 06-March 07
Iceland	2005	2005	2005
Norway	2005	2005	2005

Source: National Quality Reports 2006.
NA: this tax does not exist in the country.

Annex 7: Components of income

Table 25: Household income components: are the standard EU-SILC definitions used? (2006)

	HY010	HY020	HY022	HY023	HY030	HY040	HY050	HY060	HY070	HY080	HY090	HY100	HY110	HY120	HY130
	Total hh gross income	Total disposable hh income	Total disposable hh income before social transfers other than old-age and survivors' benefits	Total disposable hh income before all social transfers	Imputed rent (1)	Income from rental of property or land	Family/ Children related allowances	Social exclusion payments not elsewhere classified	Housing allowances	Regular inter-hh cash transfers received	Interest, dividends, profit from capital investments in incorporated businesses	Interest paid on mortgage (1)	Income received by people aged under 16	Regular taxes on wealth	Regular inter-hh transfers paid
BE	F	F	F	F	NC - NA	F	L	L	L	F	F	F	F	NA	F
CZ	F	F	F	F	NC - NA	F	F	F	F	F	F	F	F	F	F
DK	F	F	F	F	P	F	F	F	F	F	L	F	F	F	F
DE	F	F	F	F	NC - NA	L	F	F	F	F	L	NC - NA	F	F	F
EE	F	F	F	F	F	F	F	F	F	F	L	F	L	F	F
IE	F	F	F	F	NC - NA	F	F	F	F	F	F	F	F	NA	F
EL	F	F	F	F	NC - NA	F	F	F	F	F	F	NC - NA	F	F	F
ES	F	F	F	F	NC - NA	F	F	F	F	F	F	NC - NA	F	F	F
FR	Not collected	F	F	F	F	F	F	F	F	L	F	F	F	F	L

	HY010	HY020	HY022	HY023	HY030	HY040	HY050	HY060	HY070	HY080	HY090	HY100	HY110	HY120	HY130
	Total hh gross income	**Total disposable hh income**	**Total disposable hh income before social transfers other than old-age and survivors' benefits**	**Total disposable hh income before all social transfers**	**Imputed rent (1)**	**Income from rental of property or land**	**Family/ Children related allowances**	**Social exclusion payments not elsewhere classified**	**Housing allowances**	**Regular inter-hh cash transfers received**	**Interest, dividends, profit from capital investments in incorporated businesses**	**Interest paid on mortgage (1)**	**Income received by people aged under 16**	**Regular taxes on wealth**	**Regular inter-hh transfers paid**
IT	Not collected	F	F	F	NC - NA	F	F	F	F	F	F	F	F	F	F
CY	F	F	F	F	NC - NA	F	F	F	F	F	F	NC - NA	F	F	F
LV	Not collected	F	F	F	NC - NA	F	F	F	F	F	F	NC - NA	F	F	F
LT	F	F	F	F	NC - NA	F	F	F	F	F	F	F	F	F	F
LU	F	F	F	F	NC - NA	F	F	F	F	F	F	NC - NA	F	F	F
HU	F	F	F	F	NC - NA	F	F	F	F	F	F	F	F	F	F
MT	F	F	F	F	NC - NA	F	F	F	F	F	F	F	F	Not collected	F
NL	L	L	L	L	F	F	L	F	F	L	F	F	F	Not collected	L
AT	F	F	F	F	NC - NA	F	F	F	F	F	F	NC - NA	F	NA	F

	HY010	HY020	HY022	HY023	HY030	HY040	HY050	HY060	HY070	HY080	HY090	HY100	HY110	HY120	HY130
	Total hh gross income	Total disposable hh income	Total disposable hh income before social transfers other than old-age and survivors' benefits	Total disposable hh income before all social transfers	Imputed rent (1)	Income from rental of property or land	Family/ Children related allowances	Social exclusion payments not elsewhere classified	Housing allowances	Regular inter-hh cash transfers received	Interest, dividends, profit from capital investments in incorporated businesses	Interest paid on mortgage (1)	Income received by people aged under 16	Regular taxes on wealth	Regular inter-hh transfers paid
PL	F	F	F	F	NC - NA	F	F	F	F	F	F	NC - NA	F	F	F
PT	Not collected	L	L	L	NC - NA	F	F	F	F	L	F	F	N	F	L
SI	F	F	F	F	NC - NA	F	F	F	F	F	F	F	F	F	F
SK	F	F	F	F	NC - NA	F	F	L	L	F	F	F	F	F	F
FI	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F
SE	F	F	F	F	NC - NA	F	F	F	F	L	F	F	F	F	L
UK	F	L	F	F	NC - NA	F	F	F	F	F	F	F	F	F	F
IS	L	F	F	F	NC - NA	L	F	F	F	F	F	F	F	F	F
NO	F	F	F	F	NC - NA	F	L	F	L	F	F	F	F	F(2)	F

Source: National Quality Reports 2006.

F (fully comparable); L (largely comparable); P (partly comparable); N (not comparable); NC – NA (Not collected-not applicable, country does not send this data but it is not compulsory).

(1) Mandatory from 2007 onwards.

(2) Included in HY140.

Table 26: Individual income components: are the standard EU-SILC definitions used? (2006)

	PY010	PY020	PY020	PY030	PY050	PY070	PY090	PY100	PY110	PY120	PY130	PY140	PY200
	Cash or near-cash employee income	Income from private use of company car	Other non-cash employee income (1)	Employers' social insurance contributions (1)	Cash profits or losses from self-employment	Value of goods produced for own consumption (1)	Unemployment benefits	Old-age benefits	Survivors' benefits	Sickness benefits	Disability benefits	Education-related allowances	Gross monthly earnings for employees (4)
BE	F	F	NC - NA	NC - NA	F	NC - NA	L	L	L	L	L	L	NC - NA
CZ	F	F	NC - NA	NC - NA	F	P	F	F	F	F	F	F	NC - NA
DK	F	F	NC - NA	F	F	F	F	F	F	F	F	F	NC - NA
DE	F	F	NC - NA	NC - NA	L	L	L	F	F	F	F	F	NC - NA
EE	F	F	NC - NA	F	F	NC - NA	F	F	L	F	F	F	NC - NA
IE	F	F	NC - NA	F	F	F	F	F	F	F	F	F	F
EL	F	F	NC - NA	NC - NA	F	NC - NA	F	F	F	F	F	F	F
ES	F	F	NC - NA	NC - NA	F	NC - NA	F	F	F	F	F	F	F
FR	L	Not collected	NC - NA	F	F	F	F	F	F	F	F	F	NC - NA
IT	F	F	NC - NA	NC - NA	F (3)	NC - NA	F	F	F	F (3)	F	F	F
CY	F	F	NC - NA	NC - NA	F	NC - NA	F	F	F	F	F	F	NC - NA
LV	F	F	NC - NA	NC - NA	F	NC - NA	F	F	F	F	F	F	NC - NA

	PY010	PY020	PY020	PY030	PY050	PY070	PY090	PY100	PY110	PY120	PY130	PY140	PY200
	Cash or near-cash employee income	Income from private use of company car	Other non-cash employee income (1)	Employers' social insurance contributions (1)	Cash profits or losses from self-employment	Value of goods produced for own consumption (1)	Unemployment benefits	Old-age benefits	Survivors' benefits	Sickness benefits	Disability benefits	Education-related allowances	Gross monthly earnings for employees (4)
LT	L	F	F	NC - NA	F	F (2)	F	F	F	F (5)	F	F	NC - NA
LU	F	F	NC - NA	NC - NA	F	NC - NA	F	F	F	F	F	F	NC - NA
HU	F	F	NC - NA	NC - NA	F	NC - NA	F	F	F	F	F	F	NC - NA
MT	F	F	NC - NA	NC - NA	F	NC - NA	F	F	F	F	F	F	NC - NA
NL	L	F	NC - NA	NC - NA	F	NC - NA	L	F	F	F	F	F	NC - NA
AT	F	F	NC - NA	NC - NA	F	NC - NA	F	F	F	F	F	F	F
PL	L	F	NC - NA	NC - NA	L	NC - NA	F	F	F	L	F	F	F
PT	F	F	NC - NA	NC - NA	F	NC - NA	F	F	F	F	F	F	F
SI	F	F	NC - NA	NC - NA	L	L	F	F	F	F	F	F	NC - NA
SK	F	F	NC - NA	NC - NA	L	L	F	F	F	F	F	F	NC - NA
FI	F	F	F	NC - NA	F	NC - NA	F	F	F	F	F	F	NC - NA
SE	F	F	NC - NA	NC - NA	F	NC - NA	F	F	F	F	F	F	NC - NA
UK	F	F	NC - NA	NC - NA	F	F (6)	F	F	F	F	F	F	F
IS	L	Not collected	NC - NA	NC - NA	F	NC - NA	F	F	F	L	F	F	NC - NA

	PY010	PY020	PY020	PY030	PY050	PY070	PY090	PY100	PY110	PY120	PY130	PY140	PY200
	Cash or near-cash employee income	Income from private use of company car	Other non-cash employee income (1)	Employers' social insurance contributions (1)	Cash profits or losses from self-employment	Value of goods produced for own consumption (1)	Unemployment benefits	Old-age benefits	Survivors' benefits	Sickness benefits	Disability benefits	Education-related allowances	Gross monthly earnings for employees (4)
NO	L	F	NC - NA	NC - NA	F	NC - NA	F	L	L	L	L	F	NC - NA

Source: National Quality Reports 2006.

F (fully comparable), L (largely comparable), P (partly comparable), N (not comparable), NC - NA (country does not send this data but it is not compulsory).

(1) Mandatory from 2007 onwards.

(2) Variable collected but not recorded in microdata file.

(3) Paid sickness leaves of employees are included in the dependent employment incomes; the same holds true for self-employment.

(4) Variable mandatory only for countries that send the gender pay gap.

(5) Sickness benefits could not be separated from cash or near cash employee income and are recorded under this variable.

(6) This component of income is assumed to be zero.

Table 27: Source or procedure used for the collection of income variables (2006)

	Source income variables
Belgium	Interview
Czech Republic	Interview
Denmark	Register
Germany	Self-administered questionnaire
Estonia	Interview
Ireland	Interview and register
Greece	Interview
Spain	Interview
France	Interview

Registers were used for all social transfer income variables.

	Source income variables	
Italy	Interview	Administrative data have been linked to sample data and used for checking pensions and self-employment incomes.
Cyprus	Interview	
Latvia	Interview and register	Mainly interview.
Lithuania	Interview and register	Administrative data were used for making the survey income data more accurate or for supplementing them.
Luxembourg	Interview	
Hungary	Interview	
Malta	Interview	
The Netherlands	Register	
Austria	Interview	
Poland	Interview	
Portugal	Interview	
Slovenia	Interview and register	Mainly register.
Slovakia	Interview	
Finland	Register	
Sweden	Register	
United Kingdom	Interview	
Iceland	Interview and register	Tax register has been used for collecting all income variables, except for HY080 and HY130 (Regular inter-household cash transfer received and paid) which have been collected through interview.
Norway	Register	

Source: National Quality Reports 2006.

Table 28: The form in which income variables at component level have been obtained and the method used for obtaining income target variables in gross (2006)

Gross or net

	Gross or net	
Belgium	Gross and net	For a limited number of monetary variables a limited number of respondents had given only a value for the gross variant of the variable (the opposite (only net is given) occurred much more). For these cases a net value was imputed on basis of the gross using the Belgian rules of taxation. A small number of net-pensions and unemployment benefits were imputed in this way.
Czech Republic	Gross or net	
Denmark	Gross	
Germany	Gross	For all income variables respondents were asked for gross values. Only sickness benefits were supposed to be reported as an amount net of taxes and social contributions.
Estonia	Gross or net	Where only net values were collected or only net or gross value was recorded, the corresponding net and gross values were calculated on the basis of those recorded values. Conversion algorithms were created on the basis of the national tax system. Information as to which taxes were paid on income components were also collected and taken into account in the conversions.
Ireland	Gross and net	
Greece	Mainly net	Only net amounts are obtained and sent. However, it is planned to design a model on net-gross and gross-net conversion of all income variables.
Spain	Net	Respondents had the option of reporting income gross or net (of tax on income at source and, if applicable, of social contributions) at component level. The interviewee normally states income net at source although in some cases gives too gross. The form in which the amount are recorded in database are net of tax on income at source and, if applicable, of social contributions. – Net amounts: Target income variables were reported net of tax on income at source and, where applicable, net of social contributions. – Gross amounts: Target gross income variables have also been obtained, reported directly by the respondent or using a net-to-gross conversion model.

	Gross or net	
France	Net of social contributions	The income variables are collected net of social insurance contributions, which means net plus taxes. Information on taxes is also collected, and the social contributions are imputed. This means that the aggregate gross and net variables, HY010, HY020, HY022, and HY023 can be estimated. At the component level, the available information is, strictly, neither net nor gross. However, in the report and survey data, it is presented as net amounts. Actually this is net of social insurance contributions, but gross of taxes.
Italy	Net	
Cyprus	Gross	In the very few cases where gross income was impossible to collect, net income was recorded. It was converted to gross by applying the existing tax system and social insurance contributions rules.
Latvia	Net	
Lithuania	Mainly gross	Income components were collected gross, except PY010, PY050, PY090 and HY050. Conversion algorithms were created on the bases of country tax system. All income variables that are subjected to taxation and/or social insurance contribution were recorded gross and net into the microdata files (except for variable PY120 which included into variable PY010). Other income variables were recorded gross.
Luxembourg	Gross and net	
Hungary	Gross	Gross income data were collected for the income items but in case of certain benefits according to tax law which were not considered to be belonging to the taxable income net value were asked, like old-age pension or family allowance.
Malta	Mainly gross	Information on income was collected through a number of sub-questions for each income component as follows: Number of payments during the 12 months; Gross income at each payment; Net income at each payment; Tax paid per payment received; National Insurance paid per payment received. Preceding these sub-divisions was a note specifying that the income reference period was 2005, and a description of the specific income component being treated in each question. A response was expected only for one of sub-divisions 2 (gross income at each payment) and 3 (net income at each payment). Preference for the collection of information on gross income (rather than net) was expressed during briefing sessions for interviewers and was also implied through the choice of ordering of the sub-questions mentioned above.
The Netherlands	Gross	

	Gross or net	
Austria	Gross and net	When either the gross value or the net value is given and the corresponding missing value is calculated by applying general rules. If an income variable was missing but either the gross or the net amount was declared, the corresponding missing value was computed according to a model based on Austrian tax data.
Poland	Net	The respondents were asked to give the net incomes and contributions (income tax prepayments and compulsory social insurance). Only in the case of income from rental of a property (HY040) the respondents were asked to give the gross income and the amount of tax paid. The gross income was obtained by summing up net value, income tax prepayments and compulsory social insurance contributions. If the information on tax and insurance contributions was missing, the amounts were imputed on the basis of the results obtained. Only in the case of income from rental of property, the tax paid was subtracted from the gross income.
Portugal	Gross or net	All the income variables are presented net of taxes and social security contributions (except income from company car). However, some of the respondents reported gross incomes. If so, procedures were developed to convert gross incomes to net. For the households reporting only gross incomes, a micro-simulation model of the Portuguese tax system similar to the Euromod model was used. In case of both gross and net incomes reported, ad-hoc procedures were applied.
Slovenia	Gross and net	Only for PY020 gross amount was converted into the net amount. It was taken into account 25% tax, which is usually paid in advance to tax authority.
Slovakia	Gross	
Finland	Gross	
Sweden	Gross	Gross but without employers' social contributions.
United Kingdom	Gross and net	
Iceland	Gross	The income register data only report gross income at component level. Total assessed taxes and contributions to social security are collected separately from tax registers.
Norway	Gross	

Source: National Quality Reports 2006.

Annex 8: Coherence studies

Table 29: Comparison EU-SILC versus 'other sources' (2006)

	Comparison with EU-SILC 2005	Comparison with Household Budget Survey	Comparison with Labour Force Survey	Comparison with National Accounts	Comparison with administrative sources	Comparison with other sources
Belgium	Y	N	N	N	N	N
Czech Republic	N	N	N	Y	Y	N
Denmark	N	N	N	N	N	N
Germany	N	Y	N	N	N	Y (Socio-economic panel)
Estonia	N	Y	Y	Y	Y	Y (Wage Statistics)
Ireland	N	N	N	N	Y	Y (Revenue Register; Teagasc; Department of Social and Family Affairs)
Greece	Y	Y	Y	N	Y	N
Spain	Y	N	Y	Y	Y	N
France	N	N	N	Y	Y	Y (Tax Statistics)
Italy	N	N	Y	Y	Y	N
Cyprus	Y	N	Y	N	N	N
Latvia	N	Y	Y	N	N	Y (Wage and Social Protection Statistics)
Lithuania	N	Y	N	N	Y	Y (Wage Statistics)
Luxembourg	N	N	N	N	N	N
Hungary	Y	Y	N	N	N	Y (Income Survey)
Malta	Y	N	Y	Y	Y	N

	Comparison with EU-SILC 2005	Comparison with Household Budget Survey	Comparison with Labour Force Survey	Comparison with National Accounts	Comparison with administrative sources	Comparison with other sources
The Netherlands	N	N	N	N	N	Y (Income Panel Survey and Continuous Survey on Living Conditions)
Austria	Y	Y	Y	Y	Y	Y (Wage Statistics)
Poland	Y	Y	N	N	N	N
Portugal	N	Y	N	N	N	N
Slovenia	N	Y	Y	N	N	N
Slovakia	N	N	Y	N	Y	N
Finland	N	N	Y	Y	Y	Y (Income Distribution Survey)
Sweden	N	N	N	N	N	N
United Kingdom	N	N	N	N	N	Y (Family Resources Survey; Expenditure and Food Survey)
Iceland	N	N	N	N	N	N
Norway	N	N	N	N	N	N

Source: National Quality Reports 2006.