



## **EU-SILC 2005 in Estonia: Final Quality Report**

Tallinn 2007

# CONTENTS

INTRODUCTION .....	3
1. COMMON LONGITUDINAL EUROPEAN UNION INDICATORS BASED ON THE LONGITUDINAL COMPONENT OF EU-SILC .....	3
2. ACCURACY .....	3
2.1. Sample design .....	3
2.1.1. Type of sampling design .....	3
2.1.2. Sampling units.....	3
2.1.3. Stratification and sub-stratification criteria .....	3
2.1.4. Sample size and allocation criteria .....	3
2.1.5. Sample selection schemes .....	3
2.1.6. Sample distribution over time.....	3
2.1.7. Renewal of sample: Rotational groups .....	3
2.1.8. Weightings .....	3
2.1.9. Substitution .....	4
2.2. Sampling errors .....	5
2.3. Non-sampling errors .....	7
2.3.1. Sampling frame and coverage errors.....	7
2.3.2. Measurement and processing errors .....	7
2.3.3. Non-response errors .....	9
2.4. Mode of data collection.....	19
2.5. Imputation procedure .....	21
2.6. Imputed rent.....	26
2.7. Company cars.....	26
2.8. Within-household non-response inflation factor .....	26
3. COMPARABILITY .....	26
3.1. Basic concepts and definitions .....	26
3.2. Components of income.....	26
3.2.1. Differences between the national definitions and standard EU-SILC definitions .....	26
3.2.2. The source or procedure used for the collection of income variables .....	26
3.2.3. The form in which income variables at component level have been obtained .....	26
3.2.4. The method used for obtaining income target variables in the required form .....	27
3.3. Tracing rules .....	27
4. COHERENCE.....	27
4.1. Comparison of income target variables and number of persons who receive income from each 'income component', with external sources.....	27
4.2. Comparison of other target variables with external sources .....	28

## INTRODUCTION

The EU-SILC survey in Estonia started in 2004. In the first year, a sample of 6000 households was selected for the survey. Households were randomly divided into 4 rotational groups. According to original rotational scheme, one of these groups was to be dropped in 2005, but due to lower than expected response rate, it was decided to keep all the rotational groups in the sample. New sub-sample was also introduced into the survey to ensure cross-sectional representability. Thus, in 2005 sample consists of 5 rotational groups.

Present report concerns mostly longitudinal part of the survey, i.e. with 4 rotational groups retained from 2004. Unless specially mentioned, all tables in the report use data of these 4 sub-samples only.

Report follows as much as possible recommendations of two documents: Regulation No 28/2004 as regards the detailed content of intermediate and final quality reports and Technical document on intermediate and final quality reports (EU-SILC 132/04).

## 1. COMMON LONGITUDINAL EUROPEAN UNION INDICATORS BASED ON THE LONGITUDINAL COMPONENT OF EU-SILC

Longitudinal indicators are not available, as no rotational group has yet been in the survey for four years.

## 2. ACCURACY

### 2.1. Sample design

#### 2.1.1. Type of sampling design

Not to be provided for the second wave.

#### 2.1.2. Sampling units

Not to be provided for the second wave.

#### 2.1.3. Stratification and sub-stratification criteria

Not to be provided for the second wave.

#### 2.1.4. Sample size and allocation criteria

Not to be provided for the second wave.

#### 2.1.5. Sample selection schemes

Not to be provided for the second wave.

#### 2.1.6. Sample distribution over time

Not to be provided for the second wave.

#### 2.1.7. Renewal of sample: Rotational groups

Not to be provided for the second wave.

#### 2.1.8. Weightings

Cross-sectional weights of 2005 were calculated according to document 157/05, "Cross-sectional weighting: from second year of the survey onwards". In 2006, there was a change in weighting procedure and base weights for 2005 had to be calculated over again, which serve as a basis for

longitudinal weights of 2005. Procedure follows recommendations of documents V. Verma „EU-SILC weighting procedures: an outline” and J.-M. Museux „Weighting and estimation for the EU-SILC rotational design”, with some peculiarities due to modified rotational scheme. Here we describe the way base and longitudinal weights for sub-samples 1–4 were calculated.

#### 2.1.8.1. Base weights

First, we need a base weight of year 2004. In year 2004 all households were freshly selected and thus cross-sectional weights of 2004 (at the personal level, RB050) are 2004 base weights, ones we start with.

Sample consist of 4 rotational groups, base weights need to be calculated for these groups independently, as they will not be treated uniformly in the future. Therefore, 2004 base weight for person  $i$  is  $w_i = 4 \cdot RB050_i$ .

Next step is to correct for attrition that happened in 2005. This correction is done on person-level, i.e. corrected weights of persons within one household no longer need to be constant. Prior to any corrections we need to exclude from consideration persons that became out-of-scope in 2005 as they are not regarded as non-response. Out-of-scope are persons that were dead by 2005, became institutionalized or had left the country for a longer period. After excluding out-of-scope persons, attrition was modelled using a logistic model. Auxiliary variables included tenure status, household equivalised income, urbanization status of place of residence, social status, age, sex, county, ethnic nationality and household's assessment to its ability to make ends meet. With logistic model we get an estimate of response probability  $r_i$  of person  $i$  for year 2005, given he/she had responded in 2004 (i.e. his/her household had responded in 2004). Base weight for year 2005 for a person  $i$  is thus

$$w_{2i} = w_i / r_i.$$

Note that base weights defined above do not cover all individuals:

- children born to sample women get the base weight of the mother;
- persons moving into sample household from outside the survey population receive the average of base weights of existing household members;
- persons moving into sample households from other non-sample households in the population receive zero base weight.

Target variable RB060 for year 2004 is computed as  $RB060_i = w_i / 4 = RB050_i$ , and for year 2005 as  $RB060_i = w_{2i} / 4$ .

Personal base weights PB050 are calculated for persons aged 16+ with personal interview completed. To correct for within-household non-response, an empirical response rate  $r_h$  was calculated for each household. Personal base weight of person  $i$  in household  $h$  is then calculated as  $PB050_i = RB060_i / \max(0.5, r_h)$ , i.e. response rate is not allowed to inflate base weight more than 2 times. Personal base weights are the calibrated by sex and one-year age groups to the totals of R-file.

#### 2.1.8.2. Longitudinal weights

Longitudinal sample in 2005 consists of 4 sub-samples (rotational groups), selected in 2004 and retained in 2005, with base weights  $w_{2i}$  computed according to 2.1.8.1. Longitudinal weight for person  $i$  in year 2005 is thus simply  $l_i^R = w_{2i} / 4 = RB060_i$  for all persons in responding households and  $l_i^P = PB050_i$  for those with personal interview completed.

#### 2.1.9. Substitution

No substitution was used.

## 2.2. Sampling errors

The following table reports the mean, the number of observations (before and after imputations) and the standard error for different income components.

Estimates and their standard errors in Tables 1 and 2 are computed with the cross-sectional weights of year 2005 and on the full sample of 2005, i.e. rotational group 5 is included in calculations.

Table 2.1. Number of observations and standard error of different income components, 2005

Income components	Mean <sup>1</sup>	Number of observations		Standard error
		Before imputation <sup>2</sup>	After imputation	
Total household gross income (HY010)	111,526	3555	4151	1804
Total disposable household income (HY020)	91,377	3551	4151	1314
Total disposable household income before social transfer other than old-age and survivors' benefits (HY022)	84,899	3606	4151	1293
Total disposable household income before social transfers including old-age and survivors' benefits (HY023)	69,995	3616	4151	1357
<b>Net income components at household level</b>				
Income from rental of a property or land (HY040N)	98	4169	4169	38
Family/ children related allowances (HY050N)	3564	4169	4169	173
Social exclusion not elsewhere classified (HY060N)	6	4169	4169	2
Housing allowances (HY070N)	165	4164	4169	19
Regular inter-household cash transfers received (HY080N)	680	4168	4169	101
Interest, dividends, profit from capital investments in incorporated business (HY090N)	194	4122	4169	55
Income received by people aged under 16 (HY110N)	32	4169	4169	8
Regular taxes on wealth (HY120N)	300	4128	4169	12
Regular inter-household cash transfers paid (HY130N)	831	4165	4169	86
Tax on income and social contributions, net (HY140N)	0	4165	4169	0
Repayments/ receipts for tax adjustment (HY145N)	-722	4169	4169	58
<b>Gross income components at household level</b>				
Income from rental of a property or land (HY040G)	133	4169	4169	51
Family/ children related allowances (HY050G)	3826	4169	4169	215

<sup>1</sup> Zeros are included in calculations.

<sup>2</sup> Imputation includes both fully and partially missing values of national components of the income variable. Net/gross conversion is not considered as imputation.

Income components	Mean <sup>1</sup>	Number of observations		Standard error
		Before imputation <sup>2</sup>	After imputation	
Social exclusion not elsewhere classified (HY060G)	6	4169	4169	2
Housing allowances (HY070G)	165	4164	4169	19
Regular inter-household cash transfers received (HY080G)	680	4168	4169	101
Interest, dividends, profit from capital investments in incorporated business (HY090G)	254	4122	4169	73
Income received by people aged under 16 (HY110G)	34	4169	4169	9
Regular taxes on wealth (HY120G)	300	4128	4169	12
Regular inter-household cash transfers paid (HY130G)	831	4165	4169	86
Tax on income and social contributions, gross (HY140G)	18,697	0	4169	495
<b>Net income components at personal level</b>				
Employee cash or near cash income (PY010N)	33,868	11,672	11,948	648
Non-cash employee income (PY020N)	354	11,789	11,948	37
Contributions to individual private pension plans (PY035N)	251	11,865	11,948	21
Cash benefits or losses from self employment (PY050N)	666	11,864	11,948	94
Pension from individual private plans (PY080N)	2	11,947	11,948	1
Unemployment benefits (PY090N)	160	11,943	11,948	23
Old-age benefits (PY100N)	7418	11,933	11,948	150
Survivors' benefits (PY110N)	113	11,947	11,948	16
Sickness benefits (PY120N)	153	11,900	11,948	12
Disability benefits (PY130N)	889	11,946	11,948	49
Education-related benefits (PY140N)	142	11,947	11,948	22
<b>Gross income components at personal level</b>				
Employee cash or near cash income (PY010G)	43,052	11,670	11,948	864
Non-cash employee income (PY020G)	478	11,789	11,948	50
Employer's social insurance contributions (PY030G)	14,098	0	11,948	290
Contributions to individual private pension plans (PY035G)	251	11,865	11,948	21
Cash benefits or losses from self employment (PY050G)	969	11,837	11,948	131
Pension from individual private plans (PY080G)	3	11,947	11,948	1
Unemployment benefits (PY090G)	203	11,943	11,948	31

Income components	Mean <sup>1</sup>	Number of observations		Standard error
		Before imputation <sup>2</sup>	After imputation	
Old-age benefits (PY100G)	7433	11,933	11,948	153
Survivors' benefits (PY110G)	113	11,947	11,948	16
Sickness benefits (PY120G)	185	11,900	11,948	15
Disability benefits (PY130G)	889	11,947	11,948	49
Education-related benefits (PY140G)	142	11,947	11,948	22

The following table provides the same information for the equivalised disposable income broken down by sex, age groups and household size.

Table 2.2. Number of observations and standard error of mean equivalised disposable income, 2005

	Mean	Number of observations		Standard error
		Before imputation <sup>3</sup>	After imputation	
<b>Subclasses by household size</b>				
1 household member	41,688	731	772	1668
2 household member	55,154	2160	2388	1076
3 household members	62,145	2250	2700	1044
4 and more	59,602	4657	6024	688
<b>Population by age group</b>				
<25	56,805	3364	4222	910
25-34	71,469	95	1217	2166
35-44	58,850	1296	1640	1268
45-54	57,458	1376	1704	1133
55-64	55,941	1170	1327	1170
65+	42,590	1638	1774	655
<b>Population by sex</b>				
Male	58,478	4536	5554	802
Female	55,363	5262	6330	654

## 2.3. Non-sampling errors

### 2.3.1. Sampling frame and coverage errors

Not to be provided for the second wave.

### 2.3.2. Measurement and processing errors

#### 2.3.2.1. Measurement errors

The measurement errors can stem from the questionnaire (its wording, design etc), the interviewees, the interviewers and the data collection method. While it is impossible to avoid this type of errors completely, steps were taken to reduce them as much as possible.

The questionnaires were drawn up following the international practises in collecting income data. Also, where possible questions from the existing surveys carried out by the Statistics Estonia and known to be valid and reliable, were used. Pilot surveys were carried out in 2002 and 2003 with the main aim of testing the questionnaires. The results were thoroughly analysed and feedback sessions with interviewers were carried out. The questionnaires were modified accordingly for the use in the main operation in 2004. The experience from the first wave of the survey was further used to improve the

<sup>3</sup> Imputation includes both fully and partially missing values of national components of the income variable. Net/gross conversion is not considered as imputation. Households for which within-household inflation factor (HY025) could not be calculated are excluded.

questionnaire for the 2005 operation. The main modifications concerned self-employment income, change of job and different types of social insurance payments.

To reduce the measurement error stemming from the data collection method, CAPI was introduced as a data collection method from 2005 operation onwards. The main source of errors in the questionnaires in the 2004 operation resulted from routing mistakes and inconsistencies between questions. CAPI eliminates the former type of error and considerably reduces the latter, as the data-entry program includes several checks. As a result, the need to make call-backs declined and the quality of the information obtained this way was increased due to a remarkably faster pace of the whole cycle.

In 2004, all interviewers attended daylong training session in small groups. The EU-SILC team briefed the interviewers on all aspects relating to the fieldwork organisation, the questionnaire and general interviewing techniques. Special emphasis was placed on survey questions about income – types of income, their more common amounts and recipients. The first completed questionnaires gathered from field were thoroughly examined by the EU-SILC team and interviewers were notified of all of the errors made.

In 2005, all interviewers attended a two-day training session in small groups. To introduce CAPI to interviewers, the first part of the training was dedicated to general IT skills and data-entry program specific instruction. In the second half of the training, the EU-SILC team briefed the interviewers on similar income related topics as in the previous year. A separate session was held on tracing and specifics of assigning household and person numbers in the longitudinal survey. Interviewers were tested in both years, including testing of factual knowledge as well as simulated interview situations.

Overall, 58 interviewers were responsible for conducting the interviews. The household (gross sample) – interviewer ratio was 103 households per interviewer in 2004 and 90 households in 2005.

#### 2.3.2.2. *Processing errors*

In 2004, the interviews were carried out using PAPI and the data was entered centrally. The data-entry program was written in Blaise and contained most of the logical checks. The checks included, but were not limited to routing checks, consistency between different answers and upper and lower bounds for income variables. The most common mistake made by the interviewers was failing to mark an answer to one or more question or sub-questions (74% of all mistakes). Other most frequent types of error were marking the answer so that the correct answer remained unclear (for example by using wrong codes) and inconsistencies between answers provided to different questions (accounting to 7% and 6% of all mistakes respectively). The questions that were most prone to mistakes were:

- 1) Enforced lack of durables (missing answers due to unclear questionnaire layout);
- 2) Family benefits (inconsistent with household composition);
- 3) Relationship between household members (implausible relationships recorded in the household matrix);
- 4) Number of years in employment (inconsistencies with the time of taking up the first job);
- 5) Limitations in daily activities for health reasons (sometimes skipped despite there being no routing);
- 6) Calendar of activities (inconsistencies with other data).

20% of all questionnaires contained one error or more. Social Statistics Department personnel checked all errors discovered in the course of data entry. Errors that could be corrected using other data in the questionnaire or external data were corrected in office. The errors that could not be solved this way were forwarded to the interviewers' network, who consulted with the interviewer and when necessary made call-backs to the household. 20% of all errors (4% of all questionnaires) were forwarded for call-backs.

Not all checks could be feasibly implemented during the data-entry; so further data cleaning was carried out at a later stage using SAS. These checks were mainly targeted to detecting extreme income values and data-entry mistakes. Finally, the Eurostat data-checks were also implemented.

In 2005, the checking of the data consisted of 3 stages: the data-entry checks during interview, additional in-office checks during fieldwork and later data cleaning.

As mentioned above, the data for 2005 operation was collected using CAPI. The data-entry program was written in Blaise and contained most of the checks. This way, most of the errors could already be corrected during the interview. The data-entry controls were of 4 major types:



- 1) Checks of consistency between different answers. These included, but were not limited to following instances:
  - a. whether a household or a person who according to other data should have received a certain type of income reported it or not (e.g. whether households with children received family benefits, employed persons received wages and so on);
  - b. whether answers provided to different non-monetary deprivation items agreed with each other;
  - c. whether the relationships in the household matrix were consistent with each other as well as with the age and sex of the household members;
  - d. whether the difference between the starting and finishing time of the interview was too short or too long and so on.
- 2) Lower and upper bounds of income variables. These checks were developed with regard to data collected in the previous wave as well as administrative information.
- 3) Tracing checks. These controls were implemented to ensure that all split-off households and new household members were assigned correct split numbers and person numbers respectively.
- 4) Checks with information from the previous year. These controls concerned demographic data, information on educational level and labour status as well as the calendar of activities.

The in-office staff promptly checked the questionnaires that were electronically transmitted to the central office. This stage included following controls:

- 1) All the errors suppressed by interviewers were activated and checked;
- 2) All remarks made by interviewers in the data entry-program were read through and where necessary relevant corrections were made.
- 3) All split-off households as well as all households from which at least one member had left were scrutinised one by one.
- 4) Demographic information in the interviewers' reports, which were still filled out on paper, was compared to the data recorded in the electronic questionnaires.
- 5) Additionally, a few questions (child care, place of residence) had to be screened due to mistakes in the data-entry program.

The third and final stage involved later in-office data cleaning. The controls implemented at this stage involved further checks of data consistency and of extreme income values and as a final step the Eurostat data-checks. The checks of data consistency were mainly concerned with non-income variables, such as education. Also extreme values for all income components as well as total income were checked.

### 2.3.3. Non-response errors

#### *2.3.3.1. Achieved sample size*

*Table 2.3. Achieved sample size by rotational group, 2005*

	<b>Rotational group 1</b>	<b>Rotational group 2</b>	<b>Rotational group 3</b>	<b>Rotational group 4</b>	<b>Total</b>
No of households for which and interview is accepted for the database (DB135=1)	876	892	854	889	3511
No of persons 16 years or older in these households, who completed a personal interview, of which	1979	2127	2051	2047	8204
sample persons	1933	2087	2010	1998	8028
co-residents	46	40	41	49	176

### 2.3.3.2. Unit non-response

Tables 2.4-2.8 provide information used to compute response rates for households. Sample outcome in previous wave is omitted since only households for which an interview was accepted for the database in 2004 (DB135=1) were followed up in 2005.

Tables 2.4 -2.6 are based on the longitudinal database, Tables 2.7 and 2.8 on both the cross-sectional and longitudinal databases.

*Table 2.4. Number of households by household status and rotational group, 2005*

No		Rotational group 1	Rotational group 2	Rotational group 3	Rotational group 4	Total
1	<b>Total</b> (DB110=1 to 10)	1044	1057	1008	1047	4156
2	At the same address as previous interview (DB110=1)	929	942	918	923	3712
3	Entire household moved to a private household within the country (DB110=2)	46	38	38	42	164
4	Entire household moved to a collective household or institution within the country (DB110=3)	2	0	0	1	3
5	Household moved outside the country (DB110=4)	1	2	3	1	7
6	Entire household died (DB110=5)	10	5	1	9	25
7	Household does not contain sample person (DB110=6)	0	0	0	0	0
8	Address non-contacted (unable to access, lost - no information on record on what happened to the household) (DB110=7)	14	26	17	24	81
9	Split-off household (DB110=8)	41	44	31	47	163
10	New address added to the sample this wave or first wave (DB110=9)	0	0	0	0	0
11	Fusion (DB110=10)	1	0	0	0	1
12	<b>Total split-off households, households moved to a private household within the country or new households</b> (DB110=2, 8,9)	87	82	69	89	327

*Table 2.5. Number of households by contact at address (applicable to split-off households, households moved to a private household within the country and new households, DB110=2, 8, 9), 2005*

No		Rotational group 1	Rotational group 2	Rotational group 3	Rotational group 4	Total
13	<b>Total</b> (DB120=11 to 23)	87	82	69	89	327
14	Address contacted (DB120=11)	62	55	45	63	225
15	Address cannot be located (DB120=21)	25	27	24	26	102
16	Address unable to access	0	0	0	0	0

	(DB120=22)					
17	Address does not exist or is non-residential address or is unoccupied or not principal residence (DB120=23)	0	0	0	0	0

Table 2.6. Number of households by household questionnaire result, interview acceptance and rotational group (applicable to household residing at the same address as previous interview, DB110=1, and contacted, DB120=11), 2005

No		Rotational group 1	Rotational group 2	Rotational group 3	Rotational group 4	Total
18	<b>Total</b> (DB130=11 to 24)	991	997	963	986	3937
19	Household questionnaire completed (DB130=11)	885	896	862	897	3540
20	Refusal to co-operate (DB130=21)	78	69	69	69	285
21	Entire household temporarily away for duration of fieldwork (DB130=22)	16	18	17	13	64
22	Household unable to respond (illness, incapacity...) (DB130=23)	12	14	14	6	46
23	Other reasons (DB130=24)	0	0	1	1	2
24	<b>Total</b> questionnaire completed (DB135=1 and 2)	885	896	862	897	3540
25	Interview accepted for database (DB135=1)	876	892	854	889	3511
26	Interview rejected (DB135=2)	9	4	8	8	29

Table 2.7. Number of households followed up from 2004 (DB110 <> 8) by contact at address, questionnaire result and interview acceptance (cross-sectional database), 2005

No		Rotational group 1	Rotational group 2	Rotational group 3	Rotational group 4	Total
27	<b>Total</b> (DB120=11 to 23)	1003	1013	977	1000	3993
28	Address contacted (DB120=11)	971	968	948	957	3844
29	Address cannot be located (DB120=21)	18	36	23	30	107
30	Address unable to access (DB120=22)	1	2	2	2	7
31	Address does not exist or is non-residential address or is unoccupied or not principal residence (DB120=23)	13	7	4	11	35
32	<b>Total households contacted</b> (DB130=11 to 24)	971	968	948	957	3844
33	Household questionnaire completed (DB130=11)	869	869	852	870	3460
34	Refusal to co-operate (DB130=21)	73	69	65	67	274

35	Entire household temporarily away for duration of fieldwork (DB130=22)	16	16	17	13	62
36	Household unable to respond (illness, incapacity...) (DB130=23)	12	14	13	6	45
37	Other reasons (DB130=24), of which	1	0	1	1	3
38	fusion (DB110=10)	1	0	0	0	1
39	<b>Total questionnaire completed</b> (DB135=1, 2)	869	869	852	870	3460
40	Interview accepted (DB135=1)	860	865	844	862	3431
41	Interview rejected (DB135=2)	9	4	8	8	29

Table 2.8. Number of split-off households (DB110=8) by contact at address, questionnaire result and interview acceptance (cross-sectional database), 2005

No		Rotational group 1	Rotational group 2	Rotational group 3	Rotational group 4	Total
42	<b>Total</b> (DB120=11 and 23)	41	44	31	47	163
43	Address contacted (DB120=11)	21	29	15	29	94
44	Address cannot be located (DB120=21)	20	15	16	18	69
45	Address unable to access (DB120=22)	0	0	0	0	0
46	Address does not exist or is non-residential address or is unoccupied or not principal residence (DB120=23)	0	0	0	0	0
47	<b>Total households contacted</b> (DB130=11 to 24)	21	29	15	29	94
48	Household questionnaire completed (DB130=11)	16	27	10	27	80
49	Refusal to co-operate (DB130=21)	5	0	4	2	11
50	Entire household temporarily away for duration of fieldwork (DB130=22)	0	2	0	0	2
51	Household unable to respond (illness, incapacity...) (DB130=23)	0	0	1	0	1
52	Other reasons (DB130=24)	0	0	0	0	0
53	<b>Total questionnaire completed</b> (DB135=1, 2)	16	27	10	27	80
54	Interview accepted (DB135=1)	16	27	10	27	80
55	Interview rejected (DB135=2)	0	0	0	0	0

Since only households accepted to the database in 2004 (DB135=1) were followed up in 2005, response rates simplify to the following expressions.

Table 2.9. Response rates for households by rotational group, 2005

Indicator and formula	Rotational group 1	Rotational group 2	Rotational group 3	Rotational group 4	Total
Wave response rate, $R25/(R1-R4 - R5 - R6 - R7)$	85.0	85.0	85.1	85.8	85.2
Longitudinal follow-up rate, $(R27-R29-R31-R34-R38)/R27$	89.5	88.9	90.6	89.2	89.6
Follow-up ratio, $(R27-R29-R31-R34-R38+R42-R44-R46-R49)/R27$	91.1	91.8	91.7	91.9	91.7
Achieved sample size ratio, $R25/(R1-R9)$	87.3	88.1	87.4	88.9	87.9

Tables 2.10-2.13 provide information to compute response rates for persons.

Table 2.10. Number of persons by data status and rotational group, 2005

No		Rotational group 1	Rotational group 2	Rotational group 3	Rotational group 4	Total
56	<b>Total</b> (RB250=11 to 33)	2008	2167	2086	2090	8351
57	Total, sample persons (RB100=1, RB250=11 to 33)	1961	2127	2045	2041	8174
58	Total, co-residents (RB100=2, RB250=11 to 33)	47	40	41	49	177
59	Interview completed (RB250=11 to 13)	1979	2127	2051	2047	8204
60	Interview completed, co-resident (RB250=11 to 13 and RB100=2)	46	40	41	49	176
61	<b>Interview completed, sample person</b> (RB250=11 to 13 and RB100=1)	1933	2087	2010	1998	8028
62	Individual unable to respond and no proxy possible, sample person (RB250=21 and RB100=1)	1	9	0	4	14
63	Failed to return self-completed questionnaire, sample person (RB250=22 and RB100=1)	0	0	0	0	0
64	Refusal to co-operate, sample person (RB250=23 and RB100=1)	14	19	31	24	88
65	Person temporarily away and no proxy possible, sample person (RB250=31 and RB100=1)	6	5	4	8	23
66	No contact for other reasons, sample person (RB250=32 and RB100=1)	7	7	0	7	21
67	Information not completed but reason unknown, sample person (RB250=33 and RB100=1)	0	0	0	0	0

Table 2.11. Number of eligible persons (RB245=1) by data status and rotational group, 2004

No		Rotational group 1	Rotational group 2	Rotational group 3	Rotational group 4	Total
68	Sample persons total (RB100=1)	2277	2439	2349	2366	9431
69	Co-residents total (RB100=2)	0	0	0	0	0
70	Interview completed, sample persons (RB250=11 to 13, RB100=1)	2167	2289	2232	2218	8906
71	Interview not completed, sample persons (RB250=21 to 33, RB100=1)	110	150	117	148	525

Table 2.12. Number of sample persons (RB100=1) in households no longer in scope (DB110=3 to 6), 2005

No		Rotational group 1	Rotational group 2	Rotational group 3	Rotational group 4	Total
72	Sample persons (RB100=1) in households no longer in scope (DB110=3 to 6)	14	11	6	12	43

Table 2.13. Number of sample persons (RB100=1) by membership status, place moved to and rotational group, 2005

No		Rotational group 1	Rotational group 2	Rotational group 3	Rotational group 4	Total
73	<b>Total</b> (RB110=1 to 7)	2093	2258	2172	2193	8716
74	Was in this household in previous wave or current household member (RB110=1)	2008	2154	2093	2076	8331
75	Moved into this household from another sample household since previous wave (RB110=2)	17	31	14	32	94
76	Moved into this household from outside sample since previous wave (RB110=3)	0	0	0	0	0
77	Newly born into this household (RB110=4)	0	0	0	0	0
78	Moved out since previous wave (RB110=5)	52	57	52	67	228
79	Died (RB110=6)	16	16	13	18	63
80	Lived in the household at least three months during the income reference period and was not recorded in the register (RB110=7)	0	0	0	0	0
81	<b>Total</b> moved out (RB120=1 to 4)	52	57	52	67	228
82	To a private household in the country (RB120=1)	40	46	34	52	172
83	To a collective household or institution in the country (RB120=2)	3	5	2	7	17

84	Abroad (RB120=3)	9	6	16	8	39
85	Lost (RB120=4)	0	0	0	0	0

Since all and only sample persons from year 2004 living in responding households (DB135=1) are retained to year 2005, all co-residents in 2004 are 13 years old or younger and no new sample persons were added to the sample in 2005, response rates for persons simplify to the following expressions.

Table 2.14. Response rates for persons by rotational group, 2005

	Rotational group 1	Rotational group 2	Rotational group 3	Rotational group 4	Total
Wave response rate of sample persons, R61/(R68-R72-R79-R83-R84)	86.5	86.9	86.9	86.1	86.6
Wave response rate of co-residents	NA	NA	NA	NA	NA
Longitudinal follow-up rate for persons, R61/(R68-R72-R79-R83-R84)	86.5	86.9	86.9	86.1	86.6
Longitudinal non-follow-up rate for persons, individual unable to respond, R62/(R68-R72-R79-R83-R84)	0.0	0.4	0.0	0.2	0.2
Longitudinal non-follow-up rate for persons, failed to return self-completed questionnaire, R63/(R68-R72-R79-R83-R84)	0.0	0.0	0.0	0.0	0.0
Longitudinal non-follow-up rate for persons, refusal to cooperate, R64/(R68-R72-R79-R83-R84)	0.6	0.8	1.3	1.0	0.9
Longitudinal non-follow-up rate for persons, person temporarily away, R65/(R68-R72-R79-R83-R84)	0.3	0.2	0.2	0.3	0.2
Longitudinal non-follow-up rate for persons, no contact for other reasons, R66/(R68-R72-R79-R83-R84)	0.3	0.3	0.0	0.3	0.2
Longitudinal non-follow-up rate for persons, reason unknown, R67/(R68-R72-R79-R83-R84)	0.0	0.0	0.0	0.0	0.0
Achieved sample size ratio for sample persons, R61/R70	89.2	91.2	90.1	90.1	90.1
Achieved sample size ratio for sample persons and co-residents, R59/R70	91.3	92.9	91.9	92.3	92.1
Achieved sample size ratio for co-residents selected in the first wave	NA	NA	NA	NA	NA
Response rate for non-sample persons, R60/R58	97.9	100.0	100.0	100.0	99.4

In reporting these non-response rates we assume that all non-contacted households other than those coded as DB120=23 are in fact existing. This seems to be a reasonable assumption since codes DB120=21 and DB120=22 include the following non-contact reasons according to national classification (see the meaning of the term "address-person" in Intermediate Quality Report):

**DB120=21**

- Address-person does not live at given address and no information is available on new address
- Address-person has moved to another address, no information on new address available
- Given address does not exist
- Address can be located, but no contact can be made since nobody is at home

**DB120=22**

- The house given is located but given address can not be accessed (due to locked doors or gates, etc)
- Address of address-person can not be accessed due to poor weather conditions etc

2.3.3.3. *Distribution of households by household status (DB110), by record of contact at address (DB120), by household questionnaire result (DB130) and by household interview acceptance (DB135).*  
See tables 2.4-2.6.

2.3.3.4. *Distribution of persons by membership status (RB100)*

Table 2.15. Distribution of persons by membership status (RB100), 2005

	Rotational group 1	Rotational group 2	Rotational group 3	Rotational group 4	Total
<b>Total</b> (RB110=1 to 7)	2554	2706	2611	2639	10510
<i>Current household members</i>					
Was in this household in previous wave or current household member (RB110=1)	2364	2517	2439	2417	9737
Moved into this household from another sample household since previous wave (RB110=2)	21	35	17	40	113
Moved into this household from outside sample since previous wave (RB110=3)	59	49	48	60	216
Newly born into this household (RB110=4)	21	21	18	12	72
<i>Not current household members</i>					
Moved out since previous wave (RB110=5)	60	63	57	80	260
Died (RB110=6)	16	16	13	18	63
Lived in the household at least three months during the income reference period and was not recorded in the register (RB110=7)	13	5	19	12	49
<b>Total</b> moved out (RB120=1 to 4)	60	63	57	80	260
To a private household within the country (RB120=1)	48	52	39	63	202
To a collective household or institution within the country (RB120=2)	3	5	2	9	19



Abroad (RB120=3)	9	6	16	8	39
Lost (RB120=4)	0	0	0	0	0

### 2.3.3.5. Item non-response

The following table shows the amount of item non-response for income variables (among households whose interview was accepted for the database):

- percentage of persons/households having received an amount (other than 0),
- percentage of households for which no information for appropriate income variable was obtained from the questionnaire (missing values) and
- Percentage of households for which partial information (not all the questions required) for appropriate income variable was obtained from the questionnaire.

A value obtained by gross/net conversion was not considered as non-response.

Table 2.16. Distribution of item non-response, household-level variables, 2005

Income variable	% of hhs having received an amount		% of hhs with missing values (before imputation)		% of hhs with partial information (before imputation)	
	Count	%	Count	%	Count	%
Total household gross income (HY010)	3497	99.6	17	0.5	384	11.0
Total disposable household income (HY020)	3500	99.7	7	0.2	488	13.9
Total disposable household income before social transfer other than old-age and survivors' benefits (HY022)	3455	98.4	19	0.5	434	12.6
Total disposable household income before social transfers including old-age and survivors' benefits (HY023)	3173	90.4	28	0.9	415	13.1
<i>Net income components at household level</i>						
Income from rental of a property or land (HY040N)	41	1.2	0	0.0	0	0.0
Family/ children related allowances (HY050N)	1494	42.6	0	0.0	0	0.0
Social exclusion not elsewhere classified (HY060N)	23	0.7	0	0.0	0	0.0
Housing allowances (HY070N)	122	3.5	3	2.5	0	0.0
Regular inter-household cash transfers received (HY080N)	136	3.9	1	0.7	0	0.0
Interest, dividends, profit from capital investments in incorporated business (HY090N)	131	3.7	30	22.9	5	3.8
Income received by people aged under 16 (HY110N)	81	2.3	0	0.0	2	2.5
Regular taxes on wealth (HY120N)	2241	63.8	35	1.6	0	0.0
Regular inter-household cash transfers paid (HY130N)	202	5.8	3	1.5	0	0.0
Tax on income and social contributions, net (HY140N)	0	0.0	0	.	0	.

Repayments/ receipts for tax adjustment (HY145N)	1145	32.6	81	7.1	19	1.7
<i>Gross income components at household level</i>						
Income from rental of a property or land (HY040G)	41	1.2	0	0.0	0	0.0
Family/ children related allowances (HY050G)	1494	42.6	0	0.0	0	0.0
Social exclusion not elsewhere classified (HY060G)	23	0.7	0	0.0	0	0.0
Housing allowances (HY070G)	122	3.5	3	2.5	0	0.0
Regular inter-household cash transfers received (HY080G)	136	3.9	1	0.7	0	0.0
Interest, dividends, profit from capital investments in incorporated business (HY090G)	131	3.7	30	22.9	5	3.8
Income received by people aged under 16 (HY110G)	81	2.3	0	0.0	2	2.5
Regular taxes on wealth (HY120G)	2241	63.8	35	1.6	0	0.0
Regular inter-household cash transfers paid (HY130G)	202	5.8	3	1.5	0	0.0
Tax on income and social contributions, gross (HY140G)	2567	73.1	2567	100.0 <sup>4</sup>	0	0.0

Table 2.17. Distribution of item non-response, person-level variables, 2005

Income variable	% of persons 16+ having received an amount		% of persons with missing values (before imputation)		% of persons with partial information (before imputation)	
	Count	%	Count	%	Count	%
<i>Net income components at personal level</i>						
Employee cash or near cash income (PY010N)	4357	53.1	203	4.7	22	0.5
Non-cash employee income (PY020N)	130	1.6	128	98.5 <sup>5</sup>	2	1.5
Contributions to individual private pension plans (PY035N)	366	4.5	69	18.9	2	0.5
Cash benefits or losses from self employment (PY050N)	597	7.3	52	8.7	13	2.2
Pension from individual private plans (PY080N)	5	0.1	1	20.0	0	0.0
Unemployment benefits (PY090N)	149	1.8	3	2.0	1	0.7
Old-age benefits (PY100N)	2043	24.9	9	0.4	1	0.0
Survivors' benefits (PY110N)	86	1.0	2	2.3	0	0.0
Sickness benefits (PY120N)	428	5.2	41	9.6	0	0.0
Disability benefits (PY130N)	534	6.5	2	0.4	0	0.0
Education-related benefits (PY140N)	150	1.8	0	0.0	0	0.0
<i>Gross income components at personal level</i>						

<sup>4</sup> Actual amounts are not collected from respondents. The value is calculated based on whether different taxes were paid or not.

<sup>5</sup> Only use of company car is included in this variable. Benefit from company car was calculated as the number of months a company car was used multiplied by 2000, and thus considered as missing value.

Income variable	% of persons 16+ having received an amount		% of persons with missing values (before imputation)		% of persons with partial information (before imputation)	
	Count	%	Count	%	Count	%
Employee cash or near cash income (PY010G)	4357	53.1	203	4.7	22	0.5
Non-cash employee income (PY020G)	130	1.6	128	98.5 <sup>4</sup>	2	1.5
Employer's social insurance contributions (PY030G)	4202	51.2	4202	100.0 <sup>6</sup>	0	0.0
Contributions to individual private pension plans (PY035G)	366	4.5	69	18.9	2	0.5
Cash benefits or losses from self employment (PY050G)	617	7.5	72	11.7	13	2.1
Pension from individual private plans (PY080G)	5	0.1	1	20.0	0	0.0
Unemployment benefits (PY090G)	149	1.8	3	2.0	1	0.7
Old-age benefits (PY100G)	2043	24.9	9	0.4	1	0.0
Survivors' benefits (PY110G)	86	1.0	2	2.3	0	0.0
Sickness benefits (PY120G)	428	5.2	41	9.6	0	0.0
Disability benefits (PY130G)	534	6.5	2	0.4	0	0.0
Education-related benefits (PY140G)	150	1.8	0	0.0	0	0.0

## 2.4. Mode of data collection

Table 2.18. Distribution of household members aged 16 and over by Data Status and rotational group (RB250), 2005

HOUSEHOLD MEMBERS 16+ (RB245= 1 to 3)

	Total	RB250=11	=12	=13	=21	=22	=23	=31	=32	=33
<b>Total</b>	8351	8204	0	0	14	0	89	23	21	0
%	100.0	98.2	0.0	0.0	0.2	0.0	1.1	0.3	0.3	0.0
<b>Rotation group 1</b>	2008	1979	0	0	1	0	15	6	7	0
%	100.0	98.6	0.0	0.0	0.1	0.0	0.8	0.3	0.4	0.0
<b>Rotation group 2</b>	2167	2127	0	0	9	0	19	5	7	0
%	100.0	98.2	0.0	0.0	0.4	0.0	0.9	0.2	0.3	0.0
<b>Rotation group 3</b>	2086	2051	0	0	0	0	31	4	0	0
%	100.0	98.3	0.0	0.0	0.0	0.0	1.5	0.2	0.0	0.0
<b>Rotation group 4</b>	2090	2047	0	0	4	0	24	8	7	0
%	100.0	97.9	0.0	0.0	0.2	0.0	1.2	0.4	0.3	0.0

SAMPLE PERSONS 16+ (RB245= 1 to 3 and RB100=1)

	Total	RB250=11	=12	=13	=21	=22	=23	=31	=32	=33
<b>Total</b>	8174	8028	0	0	14	0	88	23	21	0
%	100.0	98.2	0.0	0.0	0.2	0.0	1.1	0.3	0.3	0.0
<b>Rotation group 1</b>	1961	1933	0	0	1	0	14	6	7	0
%	100.0	98.6	0.0	0.0	0.1	0.0	0.7	0.3	0.4	0.0
<b>Rotation group 2</b>	2127	2087	0	0	9	0	19	5	7	0
%	100.0	98.1	0.0	0.0	0.4	0.0	0.9	0.2	0.3	0.0

<sup>6</sup> Actual amounts are not collected from respondents. The value is calculated based on whether different taxes were paid or not.

<b>Rotation group 3</b>	2045	2010	0	0	0	0	31	4	0	0
%	100.0	98.3	0.0	0.0	0.0	0.0	1.5	0.2	0.0	0.0
<b>Rotation group 4</b>	2041	1998	0	0	4	0	24	8	7	0
%	100.0	97.9	0.0	0.0	0.2	0.0	1.2	0.4	0.3	0.0

## CO-RESIDENTS 16+ (RB245= 1 to 3 and RB100=2)

	Total	RB250=11	=12	=13	=21	=22	=23	=31	=32	=33
<b>Total</b>	177	176	0	0	0	0	1	0	0	0
%	100.0	99.4	0.0	0.0	0.0	0.0	0.6	0.0	0.0	0.0
<b>Rotation group 1</b>	47	46	0	0	0	0	1	0	0	0
%	100.0	97.9	0.0	0.0	0.0	0.0	2.1	0.0	0.0	0.0
<b>Rotation group 2</b>	40	40	0	0	0	0	0	0	0	0
%	100.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
<b>Rotation group 3</b>	41	41	0	0	0	0	0	0	0	0
%	100.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
<b>Rotation group 4</b>	49	49	0	0	0	0	0	0	0	0
%	100.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table 2.19. Distribution of household members aged 16 and over by Type of Interview and rotational group (RB260), 2005

## HOUSEHOLD MEMBERS 16+ (RB245= 1 to 3) and RB250= 11 or 13

	Total	RB260=1	=2	=3	=4	=5	Missing
<b>Total</b>	8204	707	7078	19	7	390	3
%	100.0	8.6	86.3	0.2	0.1	4.8	0.0
<b>Rotation 1</b>	1979	155	1722	6	4	91	1
%	100.0	7.8	87.0	0.3	0.2	4.6	0.1
<b>Rotation 2</b>	2127	188	1811	7	0	120	1
%	100.0	8.8	85.1	0.3	0.0	5.6	0.1
<b>Rotation 3</b>	2051	189	1769	2	0	90	1
%	100.0	9.2	86.3	0.1	0.0	4.4	0.1
<b>Rotation 4</b>	2047	175	1776	4	3	89	0
%	100.0	8.6	86.8	0.2	0.2	4.4	0.0

## SAMPLE PERSONS MEMBERS 16+ (RB245= 1 to 3, RB100=1) and RB250= 11 or 13

	Total	RB260=1	=2	=3	=4	=5	Missing
<b>Total</b>	8028	677	6941	19	7	381	3
%	100.0	8.4	86.5	0.2	0.1	4.8	0.0
<b>Rotation 1</b>	1933	150	1682	6	4	90	1
%	100.0	7.8	87.0	0.3	0.2	4.7	0.1
<b>Rotation 2</b>	2087	181	1780	7	0	118	1
%	100.0	8.7	85.3	0.3	0.0	5.7	0.1
<b>Rotation 3</b>	2010	181	1741	2	0	85	1
%	100.0	9.0	86.6	0.1	0.0	4.2	0.1
<b>Rotation 4</b>	1998	165	1738	4	3	88	0
%	100.0	8.3	87.0	0.2	0.2	4.4	0.0

## CO-RESIDENTS 16+ (RB245= 1 to 3, RB100=2) and RB250= 11 or 13

	Total	RB260=1	=2	=3	=4	=5	Missing
<b>Total</b>	176	30	137	0	0	9	0
%	100.0	17.1	77.8	0.0	0.0	5.1	0.0
<b>Rotation 1</b>	46	5	40	0	0	1	0
%	100.0	10.9	87.0	0.0	0.0	2.2	0.0
<b>Rotation 2</b>	40	7	31	0	0	2	0
%	100.0	17.5	77.5	0.0	0.0	5.0	0.0
<b>Rotation 3</b>	41	8	28	0	0	5	0
%	100.0	19.5	68.3	0.0	0.0	12.2	0.0

<b>Rotation 4</b>	49	10	38	0	0	1	0
<b>%</b>	100.0	20.4	77.6	0.0	0.0	2.0	0.0

## 2.5. Imputation procedure

As 2005 was the second survey year, it was possible, for some households and persons, to use values of previous year to impute missing values. Data of 2004 was used only if household or person received particular kind of income in 2004 and analysis showed that these two incomes are sufficiently closely related. If analysis indicated no correlation between the incomes of 2004 and 2005, values were not used in imputation. For some variables, values of previous year were corrected to take into account trends present in the data. Details on the number of values forwarded from 2004 to 2005 are given in Table 2.20.

If missing value could not be imputed with data from previous year, random regression method and software IVEware were used for imputation. As most of income variables have a skewed distribution, imputation was conducted on the log-scale. In general, empirical bounds of values present in the dataset were used in IVEware to bound imputed values.

If an income component was collected only net (PY020, PY035, PY080, PY090, PY100, PY110, PY120, HY050, HY140, HY145), then missing net values were imputed and then converted to gross using net/gross conversion algorithm, where necessary. Respectively, if an income component was collected only gross (HY060, HY070, HY080, HY090, PY130, PY140), then a gross value was imputed and then converted to net.

For income components, which were collected both net and gross (PY010, PY050, HY040, HY110, HY120), the procedure was as follows. If only gross value was obtained, it was first converted to net using gross/net conversion algorithm. If both net and gross value were obtained, the net value was used, since it is believed that people know this value better. Missing net values were imputed using IVEware. Gross components of EU-SILC variables were obtained with net/gross conversion algorithm. In this way, when only gross value was obtained, a value recorded in gross component was equal to the collected gross value, since net/gross and gross/net algorithm are in accordance with each other. Also, it allows basing both net and gross recorded values on the same collected value.

Net/gross and gross/net conversion algorithms were based on local tax system.

Following table provides numbers of values imputed for each income component by method of imputation. Numbers are given for the full sample of 2005, i.e. rotational group 5 is included in calculations.

Table 2.20. Percentage of imputed cases by income component in national questionnaire, 2005

Code	Description	National code	Description	Total number of persons/households having received an amount	No of values from previous year	No of values imputed randomly
<b>NET INCOME COMPONENTS ON PERSONAL LEVEL</b>						
PY010N	Employee cash or near cash income	H01N	Net annual wages	5109	157	95
		H07N	Total amount of additional payments that had not been taken into consideration in net wages	785	12	25
PY020N	Non-cash employee income	H18	Possibility to use company car	5109	0	0
		H20	Number of months a company car was used	176	0	0
PY035N	Contributions to individual private pension plans	HK1	Joining the 3rd pillar of pension insurance	9953	0	0
		HK2	Payments made into the 3rd pillar of pension insurance	356	0	73
		HK4	Joining any other collecting insurance scheme	9953	0	1
		HK5	Payments made to the collecting insurance	94	0	11
PY050N	Cash benefits or losses from self employment	H27N	Amount of loss from self-employment	84	0	16
		H28N	Net profit from self-employment	225	17	24
		H35N	Net amount of royalties, remuneration or payment under contract for creative or scientific work	43	0	3
		H46A2	Income from private provision of fee-charging services to other persons or households	235	0	9
		H46B2	Income from the sale of own-produced consumer goods (e.g. handicrafts, souvenirs, etc.)	15	0	1

Code	Description	National code	Description	Total number of persons/households having received an amount	No of values from previous year	No of values imputed randomly
		H46C2	Income from the sale of own-produced foodstuffs (e.g. pies, waffles, shashlik, etc.)	1	0	1
		H46D2	Income from intermediate commercial transactions	17	0	4
		H46E2	Income from agricultural or forestry activities	147	3	6
		H46F2	Income from other unregistered self-employment	13	0	0
PY080N	Pension from individual private plans	HK3	Whether received any payments from the 3rd pillar of pension insurance	9953	0	0
		HK3A	Payments received from the 3rd pillar of pension insurance	2	0	0
		HK6	Whether received payments from collecting insurance in the previous calendar year	9953	0	0
		HK6A	Payments from collecting insurance in the previous calendar year	6	0	1
PY090N	Unemployment benefits	H55A	Amount of unemployment benefit or any other benefits relating to unemployment	123	0	3
		HF1A	Amount of redundancy benefit received in 2003 and/or 2004	75	0	4
		HF2A	Amount of redundancy benefit used up in the previous calendar year	44	0	1
PY100N	Old-age benefits	H51A	Amount of old-age benefits	2393	3	12

Code	Description	National code	Description	Total number of persons/households having received an amount	No of values from previous year	No of values imputed randomly
PY110N	Survivors' benefits	H53A	Amount of survivors' pension or any other benefits relating to the loss of a provider	107	1	1
PY120N	Sickness benefits	H54A	Amount of sickness benefits or any other benefits relating to health	498	0	48
PY130N	Disability benefits	H52A	Amount of pension for incapacity for work or any other benefits relating to disability	886	1	1
PY140N	Education-related benefits	H57A2	Amount of state stipend	118	0	0
		H57G2	Education allowance	69	0	1
<b>NET INCOME COMPONENTS ON HOUSEHOLD LEVEL</b>						
HY040N	Income from rental of a property or land	D09N	Net income from rental of property	49	0	0
HY050N	Family/ children related allowances	D11B	Parental benefit received in the previous calendar year	129	0	0
		D11C	Total amount of other benefits received in the previous calendar year	1748	0	0
HY060N	Social inclusion not elsewhere classified	H58A2	Amount of other support/benefit/pension not mentioned above	28	0	0
HY070N	Housing allowances	D03A	Amount of subsistence benefit	142	2	3
HY080N	Regular inter-household cash transfers received	D16A	Amount of regular payments from other households	165	1	0
HY090G	Interest, dividends, profit from capital	H49A2	Interest income from deposits (deposit rate in bank)	138	3	37



Code	Description	National code	Description	Total number of persons/households having received an amount	No of values from previous year	No of values imputed randomly
	investments in incorporated business	H49B2	Interest income from securities (shares, bonds)	27	1	10
		H49C2	Dividend income from securities (i.e. shares, bonds)	59	1	15
HY110N	Income received by people aged under 16	D19N	Income received by children aged 16 or less	92	0	0
HY120N	Regular taxes on wealth	D10A	Amount of tax on land or any other property tax paid	2650	15	26
HY130N	Regular inter-household cash transfers paid	D14A	Amount of regular payments to other households	247	1	3
HY145	Repayments/receipts fro tax adjustment	H64A	Income tax return for the income received in the previous year	1582	39	65
		H63	Additional amount of tax paid on income in the previous calendar year	89	7	8

## 2.6. Imputed rent

Variable was not recoded.

## 2.7. Company cars

In the personal questionnaire, each employee was requested to report whether he or she had an option to use a company car for private ends during the previous calendar year or not. Those reporting the use were further asked to indicate the number of months the car was used, as well as the make, model and year of issue of the car. Since there is no reliable information on used car prices in Estonia, the construction of depreciation model was not possible and the conversion using tax rules was used instead. For each person reporting a benefit from the company car, the special benefit tax paid by the employer on the use of the car is recorded.

## 2.8. Within-household non-response inflation factor

Ca 3% of households, those data was accepted for the database, are affected by within household non-response. Ignoring this fact may cause substantial bias in main indicators. To correct the effect of non-responding individuals within a household, the inflation factor (variable HY025) is calculated by which it is necessary to multiply total household gross income, total household disposable income and the total household disposable income before social transfers.

Calculation procedure was as follows.

1. Missing total individual gross income was imputed with the unweighted median of imputation class. Imputation classes were cross-sections of 9 age groups (16-17, 18-19, 20-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75+), sex, employment status (employed/other), and quintile of equivalised total household gross income. Where possible, total individual income of previous wave was used for imputation (if it was greater than zero).
2. Quintile of equivalised total household gross income was recalculated using imputed values of previous step.
3. Steps 1 and 2 were repeated until there were no more shifts in the income quintile (5 repetitions were sufficient in 2005).
4. Within-household non-response inflation factor was calculated as the ratio of total household gross income including imputed individual income to initial household gross income. For households with negative or zero household gross income, inflation factor was set to 1 (i.e. no inflation was done). If resulting value was greater than 5, inflation factor was set to missing.

## 3. COMPARABILITY

### 3.1. Basic concepts and definitions

There were no changes in basic concepts and definitions from the first wave.

### 3.2. Components of income

#### 3.2.1. Differences between the national definitions and standard EU-SILC definitions

In 2004, profits and losses from self-employment (PY050N and PY050G) included on some occasions profits of unincorporated enterprises belonging to a respondent. This question was modified in 2005 to exclude profits not taken out. Total income variables (HY010, HY020, HY022 and HY023) are also affected by this change.

#### 3.2.2. The source or procedure used for the collection of income variables

There were no changes in the source or procedure used for the collection of income variables from the first wave.

#### 3.2.3. The form in which income variables at component level have been obtained

There were no changes in the form in which income variables at component level have been collected from the first wave.

### 3.2.4. The method used for obtaining income target variables in the required form

There were no changes in the source or procedure used for the collection of income variables from the first wave.

### 3.3. Tracing rules

There were no differences between the national tracing rules and the standard EU-SILC tracing rules.

## 4. COHERENCE

This section will compare the longitudinal EU-SILC data to various external sources, including the National Accounts (NA), the Household Budget Survey (HBS), the Labour Force Survey (LFS), wage statistics and social protection statistics. Estimates have been calculated using longitudinal weights and only for longitudinal part of the sample of 2005. For this reason, only individual level components are compared. Coherence analysis of 2004 data can be found in Final Quality Report for 2004 operation.

*The HBS* is a continuous survey of households, which has been carried since 1996. The yearly sample size is approximately 4500 households. The HBS is designed to collect information on income and expenditure of households. Data on income is gathered using a diary, where household records all income received during one month. The HBS was the source of Laeken indicators up until EU-SILC.

*The LFS* is a continuous survey, which is carried out according to the common EU methodology since 1995. The yearly sample size is about 12,000 working aged persons. Up until 2005, the LFS was carried out using PAPI and face-to-face interviews. The LFS is the main source for labour market information.

*Wage statistics* have in their current form been continuously calculated since 1992. All enterprises employing 50 persons or more are obliged to provide data. A sample is drawn from smaller enterprises. Wage data is used to calculate hourly and monthly wages, both gross and net, as well as labour costs. All figures have been converted into full-time units.

### 4.1. Comparison of income target variables and number of persons who receive income from each 'income component', with external sources

Total annual income from gross wages and salaries was 47,949.1 millions of kroons according to EU-SILC longitudinal components in 2004. The same figure was somewhat higher at 53,919.6 according to national accounts. This is to be expected given that national accounts estimate also includes non-cash employee income.

Next, EU-SILC income data is compared component by component to income data from HBS and administrative sources for income year 2004. Table 4.1 presents the comparisons by average amounts and Table 4.2 by number of recipients. Only the income components where definitions are similar enough to warrant comparisons on personal level are presented here.

Turning to the cash employee income first, the average amount is by 2.5% higher in EU-SILC than in HBS. The corresponding figure from wage statistics is considerably higher, however. When comparing the number of people receiving wages and salaries, it appears that there are some 100,000 persons more in EU-SILC who report this type of income than in HBS. The corresponding figure in wage statistics is lower still. The difference with wage statistics is to be expected, given that the latter refer to the full-time equivalent and the unofficial work relationships are not included. The EU-SILC – HBS difference in the number of recipients can probably be traced to the survey design. As HBS yearly figures are derived from monthly data, shorter employment spells that are concentrated to a few months in a year (mainly summer) are under reported in HBS.

The differences in non-cash employee income (i.e. company cars), however, are much more substantial. The average amount as well as the number of recipients is considerably higher in EU-SILC as opposed to HBS. This is again related to survey design. While in EU-SILC the question about the use made of company cars is posed directly to interviewees, in the case of HBS the respective

question is in the diary and may go unnoticed by the household. Also, the methods used to estimate the benefit from the car are different (taxation approach in EU-SILC vs. self-estimation in HBS).

*Table 4.1. Average amounts of income components by source of information, income year 2004*

<b>Income component</b>	<b>EU-SILC</b>	<b>HBS</b>	<b>Other sources*</b>
Net cash or near-cash employee income (PY010N)	59,695	58,231	68,100
Net non-cash employee income (PY020N)	17,668	10,092	
Net old-age benefits (PY100N)	27,616	27,155	26,431
Net survivor's benefits (PY110N)	11,441	12,533	8484
Net sickness benefits (monthly)	2750	1542	
Net disability benefits (PY130N)	14,796	13,281	

\* Wage statistics in the case of PY010 and administrative sources for other variables.

In the case of the unemployment benefits, the definitions that can be used differ between the sources and thus only the number of recipients can be tentatively compared. Despite the definition in administrative sources being more restrictive, the number of people receiving unemployment benefits in both EU-SILC and HBS is substantially lower, suggesting under coverage in the surveys.

The average amounts of old-age benefits received are somewhat higher in EU-SILC, while the number of recipients is nearly identical in two surveys. It must be taken into consideration, however, that the average amount in EU-SILC also includes benefits received from abroad, which tend to be higher than national benefits, as well as other old-age benefits that are not taken into accounts in the other sources. The number of recipients as well as the average amounts of survivor's benefits received is similar in the two surveys. The average amount also agrees well with the administrative data, whereas the number of recipients seems somewhat overestimated in both surveys.

The number of people receiving sickness benefits is, given the seasonal nature of the component, underestimated in HBS. The amounts are also higher in EU-SILC, despite the HBS figure including some lump-sum family benefits. Neither the number of recipients nor the average amounts paid is available from the administrative sources. The only information that can be used is the total amount of benefits paid, which is times higher than the respective figure from EU-SILC. This suggests that sickness benefits are under reported in EU-SILC. It is likely that interviewees do not separate smaller amounts from wages and salaries.

The average disability benefits received by people are not too different between the two surveys, although the number of people receiving is lower in EU-SILC, as disability benefits paid to retired persons have been recorded as old-age benefits.

*Table 4.2. Number of recipients of income components by source of information, income year 2004*

<b>Income component</b>	<b>EU-SILC</b>	<b>HBS</b>	<b>Other sources*</b>
Net cash or near-cash employee income (PY010N)	631,687	539,411	474,061
Net non-cash employee income (PY020N)	20,883	7,829	
Net unemployment benefits (PY090N)	18,642	13,942	39,338
Net old-age benefits (PY100N)	297,375	296,346	301,658
Net survivor's benefits (PY110N)	10,968	10,964	11,012
Net sickness benefits (PY120N)	62,165	12,420	
Net disability benefits (PY130N)	62,909	86,068	

\* Wage statistics in the case of PY010 and administrative sources for other variables.

## **4.2. Comparison of other target variables with external sources**

In Table 4.3 the distribution of population aged 16-74 derived from EU-SILC and LFS is compared. Most of the differences are minor, with the only exceptions being ISCED levels 1 and 3. There are somewhat less people with upper secondary education according to EU-SILC and more people with primary education. Given that the questions used in the two surveys are identical, this must be due to sample fluctuations.

*Table 4.3.* Distribution of population aged 16-74 by ISCED level, based on the longitudinal EU-SILC and the LFS, 2005

<b>ISCED level</b>	<b>EU-SILC</b>	<b>LFS</b>
0 Pre-primary education	0.3	0.4
1 Primary education	5.4	3.1
2 Lower secondary education	18.3	17.5
3 (Upper) secondary education	43.0	44.8
4 Post-secondary non tertiary education	7.6	8.7
5 First stage of tertiary education	24.6	25.1
6 Second stage of tertiary education	...*	0.6
Total	100.0	100.0

\* Extremely unreliable estimate, based on less than 20 sample observations

Finally, Table 4.4 presents the comparison of population aged 16 or over by most frequent current activity status. The differences that can be observed between the two data sources are relatively minor and may be due to misclassification to 'other inactive' category in HBS.

*Table 4.4.* Distribution of population aged 16 and over by self-defined activity status based on longitudinal EU-SILC and the HBS, 2005

<b>Activity status</b>	<b>EU-SILC</b>	<b>HBS</b>
Working full-time	51.3	51.1
Working part-time	3.9	3.2
Unemployed	5.8	5.2
Pupil, student	9.0	9.7
In retirement	22.6	21.5
Permanently disabled	3.5	4.3
Fulfilling domestic tasks and care responsibilities	3.9	3.4
Other inactive	...*	1.6
Total	100.0	100.0

\* Extremely unreliable estimate, based on less than 20 sample observations