EU-SILC 2007 in Estonia: Intermediate Quality Report

Tallinn 2008

CONTENTS

1. COMMON CROSS-SECTIONAL EUROPEAN UNION INDICATORS

Table 1.1. Common cross-sectional European Union indicators and their standard errors, 2007

Indicator	Value	Standard error
Portfolio of overarching indicators		
At-risk-of-poverty threshold – one person household	41749	488,7
At-risk-of-poverty threshold – household with 2 adults and 2 children	87 672	1026,4
At-risk-of-poverty rate after social transfers: Total	19,4	0,6
At-risk-of-poverty rate after social transfers by Age and Gender		
At-risk-of-poverty rate after social transfers: Male	16,7	0,7
At-risk-of-poverty rate after social transfers: Female	21,6	0,7
At-risk-of-poverty rate after social transfers: 0-17 total	18	1,1
At-risk-of-poverty rate after social transfers: 18-64 total	16	0,6
At-risk-of-poverty rate after social transfers: 18-64 male	15	0,7
At-risk-of-poverty rate after social transfers: 18-64 female	17	0,7
At-risk-of-poverty rate after social transfers: 65+ total	33,1	1,7
At-risk-of-poverty rate after social transfers: 65+ male	20,7	2,1
At-risk-of-poverty rate after social transfers: 65+ female	39,2	1,9
At-risk-of-poverty rate after social transfers by most frequent activity status and gender		
At-risk-of-poverty rate after social transfers: At work total	7.7	0,4
At-risk-of-poverty rate after social transfers: At work male	6,4	0,5
At-risk-of-poverty rate after social transfers: At work female	9,0	0,6
At-risk-of-poverty rate after social transfers: Not at work total	35,9	1,2
At-risk-of-poverty rate after social transfers: Not at work male	35,0	1,5
At-risk-of-poverty rate after social transfers: Not at work female	36,4	1,3
At-risk-of-poverty rate after social transfers: Unemployed total	61,7	3,4
At-risk-of-poverty rate after social transfers: Unemployed male	64,4	4,0
At-risk-of-poverty rate after social transfers: Unemployed female	56,0	6,0
At-risk-of-poverty rate after social transfers: Retired total	36,9	1,8
At-risk-of-poverty rate after social transfers: Retired male	25,2	2,3
At-risk-of-poverty rate after social transfers: Retired female	42,1	2,0
At-risk-of-poverty rate after social transfers: Other inactive total	30,1	1,3
At-risk-of-poverty rate after social transfers: Other inactive male	35,0	2,0
At-risk-of-poverty rate after social transfers: Other inactive female	27,0	1,5
At-risk-of-poverty rate after social transfers by household type		
At-risk-of-poverty rate after social transfers: Household without children	23.1	0,9
At-risk-of-poverty rate after social transfers: Single male	41,6	3,3
At-risk-of-poverty rate after social transfers: Single female	52,9	2,4
At-risk-of-poverty rate after social transfers: Single person aged under 65	32.8	2,4
At-risk-of-poverty rate after social transfers: Single person aged 65 and over	68.5	2,8
At-risk-of-poverty rate after social transfers: Two adults younger than 65 years	14.4	1,5
At-risk-of-poverty rate after social transfers: Two adults, at least one aged 65 and over	10.9	1,5
At-risk-of-poverty rate after social transfers: Three or more adults	7.5	1,2
At-risk-of-poverty rate after social transfers: Households with dependent children	16.2	0,8
At-risk-of-poverty rate after social transfers: Single parent with dependent children	43.9	3,8
At-risk-of-poverty rate after social transfers: Two adults with one dependent child	11.5	1,4
At-risk-of-poverty rate after social transfers: Two adults with two dependent children	11.9	1,5
At-risk-of-poverty rate after social transfers: Two adults with three or more dependent children	21.4	2,8
At-risk-of-poverty rate after social transfers: Three or more adults with dependent children	10,5	1,4
At-risk-of-poverty rate after social transfers by accommodation tenure status		
At-risk-of-poverty rate after social transfers: Owner	18,9	0,6
At-risk-of-poverty rate after social transfers: Rent	27,2	3,5

Relative median at-risk-of-poverty gap: Total	20,2	0,9
Relative median at-risk-of-poverty gap: Male	24,2	1,4
Relative median at-risk-of-poverty gap: Female	18,5	0,9
Relative median at-risk-of-poverty gap: 0-17	26	2,2
Relative median at-risk-of-poverty gap: 18-64 total	26	1,4
Relative median at-risk-of-poverty gap: 18-64 male	29	1,9
Relative median at-risk-of-poverty gap: 18-64 female	23	1,4
Relative median at-risk-of-poverty gap: 65+ total	14,2	1,0
Relative median at-risk-of-poverty gap: 65+ male	14,3	1,8
Relative median at-risk-of-poverty gap: 65+ female	14,2	1,1
Dispersion around the at-risk-of-poverty threshold:		
Dispersion around the at-risk-of-poverty threshold: 40%	6	0,4
Dispersion around the at-risk-of-poverty threshold: 50%	11	0,5
Dispersion around the at-risk-of-poverty threshold: 60%	27	0,6
At-risk-of-poverty rate before social transfers, excl old-age and survivors' benefits: Total	25	0,6
At-risk-of-poverty rate before social transfers (excl. pensions): Male	23	0,7
At-risk-of-poverty rate before social transfers (excl. pensions): Female	27	0,7
At-risk-of-poverty rate before social transfers (excl. pensions): 0-17	28	1,2
At-risk-of-poverty rate before social transfers (excl. pensions): 18-64 total	21	0,6
At-risk-of-poverty rate before social transfers (excl. pensions): 18-64 male	21	0,7
At-risk-of-poverty rate before social transfers (excl. pensions): 18-64 female	22	0,7
At-risk-of-poverty rate before social transfers (excl. pensions): 65+ total	36	1,7
At-risk-of-poverty rate before social transfers (excl. pensions): 65+ male	23	2,1
At-risk-of-poverty rate before social transfers (excl. pensions): 65+ female	42	1,9
In-work at-risk-of-poverty rate: Total	8	0,4
In-work at-risk-of-poverty rate: Male	6	0,5
In-work at-risk-of-poverty rate: Female	9	0,6
S80/S20 income quintile share ratio	5,5	0,3
Inequality of income distribution: Gini coefficient	33,4	1,5
Relative median income ratio	0,66	-
Aggregate replacement ratio: Total	0.47	-
Aggregate replacement ratio: Male	0.4	-
Aggregate replacement ratio: Female	0.56	-
At-risk-of-poverty rate before social transfers (incl. pensions): Total	37,4	0,5
At-risk-of-poverty rate before social transfers (incl. pensions): Male	33,5	0,6
At-risk-of-poverty rate before social transfers (incl. pensions): Female	40,7	0,6
At-risk-of-poverty rate before social transfers (incl. pensions): 0-17	31	1,2
At-risk-of-poverty rate before social transfers (incl. pensions): 18-64 total	27	0,6
At-risk-of-poverty rate before social transfers (incl. pensions): 18-64 male	26	0,8
At-risk-of-poverty rate before social transfers (incl. pensions): 18-64 female	29	0,7
At-risk-of-poverty rate before social transfers (incl. pensions): 65+ total	83,1	1,5
At-risk-of-poverty rate before social transfers (incl. pensions): 65+ male	81,7	1,1
At-risk-of-poverty rate before social transfers (incl. pensions): 65+ female	83,9	1,2

Table 1.2. Equivalised disposable income, 2007

Indicator	Value (kroons)
Median equalised yearly disposable income	69 581
Mean equalised yearly disposable income	82 988

2. ACCURACY

2.1. Sample design

As 2007 operation was the forth round of EU-SILC in Estonia, the sample comprised of four parts:

1. The first part consists of households selected for the survey in 2004 and followed up in 2005, 2006 and 2007. Of the overall 4 rotational groups selected in 2004, two were dropped after the 2006 survey and thus the survey of 2007 included the remaining two rotational groups (in total 1779 households including fresh split-off households).

2. The second part consisted of households selected for the survey in 2005 and followed-up in 2006 and 2007 (in total 584 households including fresh split-off households).

3. The third part consisted of households selected for the survey in 2006 and followed-up in 2007 (in total 1709 households including fresh split-off households).

4. The sample of 2444 households introduced into the survey in 2007.

In what follows we call parts 1, 2 and 3 together a repeated or old part of the sample and part 4 a new part of the sample.

Sub-sections 2.1.1 – 2.1.5 describe the design of new sub-sample. The sampling designs of 2004, 2005 and 2006 sub-samples can be found in quality reports of the respective years. They were also originally selected by unequal probability design, similar to one used for the selection of the new sub-sample.

Unequal probability design is likely to have a negative effect on sample efficiency, and research on the possibilities of improving the design has been carried out and will continue in the future. So far, however, no suitable frame for selecting addresses has been found.

2.1.1. Type of sampling design

The design used is a one-stage stratified unequal probability sampling of households, where a household's selection probability is proportional to the number of persons aged 14+ in it. This is so because a sample of persons aged 14+ (so called address-persons) is selected first with equal probabilities within their strata, after which the household of the selected person is identified, and all eligible persons in the household are interviewed. Stratification is done by geographical region (see 2.1.3).

2.1.2. Sampling units

One stage sampling design was used. Households were regarded as sampling units although selection was made using the sample of address-persons.

2.1.3. Stratification and sub-stratification criteria

Geographical stratification was used. The counties (and capital Tallinn) were grouped into three strata by the population size:
1. big counties: Tallinn, Harju (excluding Tallinn), Ida-Viru, Lääne-Viru, Pärnu, Tartu;
2. small counties: Jõgeva, Järva, Lääne, Põlva, Rapla, Saare, Valga, Viljandi, Võru;
3. Hiiu County formed a separate stratum as the smallest county with their population being a few times smaller than the next county in size.

2.1.4. Sample size and allocation criteria

Inclusion probabilities of address-persons in different strata are shown in Table 2.1. R_g stands for the number of persons aged 14 and older living in stratum g as at 01.01.2007, n_g is the sample size of the stratum g and n_g/R_g (%) is the sampling fraction in the corresponding stratum.

Table 2.1. Stratification of the new part of the sample by counties, Estonian EU-SILC 2007

Stratum h	Counties	R_g	n_g	n_g/R_g %
Large	Tallinn, Harju, Ida-Viru, Lääne-Viru, Pärnu, Tartu	859,764	1,464	0.17
Small	Jõgeva, Järva, Lääne, Põlva, Rapla, Saare, Valga, Viljandi, Võru	288,078	902	0.31
Hiiu	Hiiu	9,226	78	0.85

The following table shows the sample size by rotational group: the initial sample size, the number of split-off households and the final sample size.

Table 2.2. Sample size by rotational group, 2007

Year a rotational group started	Rotational group	Initial sample size in 2007	Nr of split-off households	Final sample size
2004	3	847	32	879
	4	874	26	900
2005	5	568	16	584
2006	6	1646	63	1709
2007	7	2444	-	2444
Total		6379	137	6516

2.1.5. Sample selection schemes

Systematic sampling of address-persons with foregoing sample sizes in each stratum. For households this procedure results in unequal probability sampling with inclusion probabilities proportional to household size (number of persons aged 14+ in it).

2.1.6. Sample distribution over time

A fixed income reference period was used and therefore the sample was not principally divided into months or weeks. The fieldwork period was from March to June 2007. For the convenience of fieldwork administration, the old part of the sample was equally allocated into the whole fieldwork period (with a slightly smaller sample size in June), while the new part was allocated into the first three months (March-May) only. When allocating households into the months of the fieldwork period, a uniform workload for interviewers was the target. Due to a lack of interviewers in some areas, ca 5% of households were interviewed after the official end of the fieldwork period in July 2007.

2.1.7. Renewal of sample: Rotational groups

The sample consists of 5 rotational groups:
1. 2 rotational groups from 2004 (groups 3 and 4);
2. 1 rotational group from 2005 (group 5);
3. 1 rotational group from 2006 (group 6)
4. New sub-sample (group 7).

2.1.8. Weightings

2.1.8.1 Introduction

The weighting scheme was generally in line with 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 the modified rotational scheme. This section will describe in detail the actual algorithm used. The description is not following the structure proposed in the Eurostat technical document for quality reports, but rather mentioned documents by V. Verma and J.-M. Museux

The sample of year 2007 consists of four sub-samples to be weighted independently and combined thereafter:

s_4	households started in 2004 and their split-offs, participating for the fourth time (= two rotational groups)
s_3	households started in 2005 and their split-offs, participating for the third time (= one rotational group)
s_2	households started in 2006 and their split-offs, participating for the second time (= one rotational group)
s_1	households started in 2007, participating for the first time (= one rotational group)

First, we need to calculate base weights for the year 2007 for each sub-sample. For sub-sample s_1 they are calculated from the beginning, for the other sub-samples base weights of 2006 should be corrected for attrition between 2006 and 2007.

In what follows we describe the procedure of obtaining base weights of 2006 for each sub-sample independently.

2.1.8.2 Calculation of base weights

The sub-sample s_1 is a usual random sample from the population and it does not depend on other sub-samples. The 2007 base weights for this sub-sample are usual first-year cross-sectional weights. With these weights s_1 represents the cross-sectional population of 2007. Calculation of first-year cross-sectional weights is done according to same procedure as in 2006. For details, see 2006 Intermediate Quality Report, section 2.1.8.2.1

Base weights for sub-samples s_2, s_3 and s_4 are obtained after correcting the base weights of 2006 for attrition. Prior to that we exclude from consideration persons no longer in scope in 2007 (see 2006 Intermediate Quality Report, section 2.1.8.2.2 for details). Correction for attrition is done with the help of a logistic model with tenure status, household equivalised income, urbanization status and county of place of residence, social status, age, sex, ethnic nationality, household's assessment to its ability to make ends meet and number of children in household as auxiliary variables. Details on the procedure of this correction can be found in 2006 Intermediate Quality Report.
Base weights were extended to other individuals in the household as follows:
  − children born to sample, women get the base weight of the mother;
  − persons moving into a 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 a base weight of zero.
Correction factors at this step were checked to be in limits from 1/3 to 3.

For sub-samples s_3 and s_4 a correction for returnees was also made. Details can be found in 2006 Intermediate Quality Report.

2.1.8.3. Calculation of cross-sectional weights
The same procedure as in 2006.
Personal cross-sectional weights for all household members aged 16 and over (PB040) coincide in 2007 with weights RB050, as within-household non-response is imputed.

Table 2.3. Distributional characteristics of final household cross-sectional weights by stratum and household size, 2006

Stratum	Household size	Mean	Std	CV (%)
1	1	259.4	133.1	51.3
	2	155.8	78.9	50.6
	3	134.3	75.4	56.1
	4	118.9	59.7	50.2
	5	94.5	53.0	56.1
	6	79.6	32.5	40.8
	7	62.3	24.4	39.1
	8	57.4	31.7	55.2
	9	56.2	.	.
	10	44.2	3.6	8.2
	11	45.1	12.4	27.4
2	1	110.6	61.1	55.2
	2	61.8	26.1	42.3
	3	51.9	25.5	49.0
	4	50.3	24.3	48.3
	5	37.1	17.1	46.1
	6	30.2	13.6	45.0
	7	23.1	8.7	37.5
	8	24.4	11.1	45.4

	9	21.6	8.5	39.2
	10	20.6	0.3	1.7
	11	19.5	5.1	26.0
3	1	47.0	23.8	50.7
	2	22.9	11.1	48.5
	3	17.0	8.4	49.2
	4	14.4	5.7	39.7
	5	11.9	5.4	45.4
	6	11.0	4.3	39.0
	7	9.3	3.7	40.2
	8	4.0	.	.
All				

2.1.8.2. Substitution
No substitution was used.

2.2. Sampling errors

2.2.1. Standard error and effective sample size

Variance estimation of the common cross-sectional EU indicators was done using the Deville linearization approach, with the help of linearization macros provided by Eurostat. After linearization the variance estimates were computed using the Bascula module of Blaise. All sub-samples were treated as if they were freshly selected. Variances were computed at the final stage of weighting procedure (2.1.8.3) together with final calibration.

Standard errors of the common cross-sectional indicators broken down by background variables are shown in Table 1.1.

A recent simulation study showed that the design effect of the at-risk-of-poverty rate was 1.2 and for the mean equivalized income it was 1.25. In this simulation study a self-weighting sub-sample was sub-selected from the new part of the sample of 2006. To compare variances, a simple random sample of the same size was also selected from the new part (thus reproducing the sampling design used in the survey). Variances of the at-risk-of-poverty rate and the mean equavalised income were calculated from both samples. The design effect was calculated as the ratio of these two variances.

Of the total of 6,516 households in the sample, data from 5,146 households were accepted into the final database. In these households 11,875 persons aged 16+ were interviewed. The effective sample size is thus 4288 households and 9,896 persons (according to Commission regulation we use the design effect of the at-risk-of-poverty rate). The minimum requirements are thus satisfied (3500 households and 7750 persons).

2.3. Non-sampling errors

2.3.1. Sampling frame and coverage errors

The sampling frame for selection of the new part of the sample was the Population Register of Estonia. This is the document-based register of Estonian citizens and those who have a residence permit. Records of the register are updated both in real-time and regularly from administrative sources. The register data originates from local governments, civilian registry offices, county councils, courts, the Citizenship and Migration Board and other governmental organisations.

The frame error is considered to be an over-coverage error if an address-person did not actually belong to target population, i.e.

- was dead;
- had moved to another county;
- was staying in an institution permanently (had been there over half a year);
- was surveyed through one of his/her own household members;

All households classified under DB120=23 are considered to constitute over coverage error. The amount of this error in the new part of the sample in 2007 was 96 households, which makes the

proportion of the over-coverage in the new part of the sample 3.9% and of the whole sample 1.5% (Table 2.4).

Since there is no registration law in Estonia, people do not need to show their actual addresses in the Population Register. For that reason the register contains some amount of records without any address and for some part of records the address shown is not correct. Records without an address or with an incomplete address were dropped out of the register before selecting the sample (for example, in 2007 2.9% of all records referring to persons aged 14+ were dropped before selecting the sample).

In the new part of the sample of 2007 there were 161 address-persons whose address in the population register was definitely wrong and no information on the new address could be obtained from neighbours. According to national classification, this includes the following reasons for non-contact:

- Address-person does not live at given address, no information on new address available;

- Address-person has moved to another address, no information on new address available;

- Given address does not exist.

It does not seem reasonable to assume that these persons do not belong to the target population or constitute frame over-coverage. The abovementioned reasons for non-contact are currently classified under DB120=21.

Due to the absence of a registration law in Estonia, there is also some under-coverage of persons and households present in the population register. Investigations made by the Sampling Working Group of HBS in 1999 showed that on average under-coverage of addresses in the population register may reach 5-6%. Degree of under coverage of households is much more difficult to asses, since even if a person is missing from the Population Register or his/her address is incorrect or not precise enough, a household could be reached through another household member. Assuming that all persons living permanently in Estonia are registered in the Population Register and considering the amount of imprecise addresses in PR, under coverage of households may be at most 1-1.5%.

Table 2.4. Reasons for over-coverage in the new part of the sample, 2007

Frame error	Number of households	Proportion in the frame error (%)
Total, of which	96	100
Address person was dead	9	9.4
Address person has left Estonia	70	72.9
Address person was staying in an institution	17	17.7
Address person was surveyed through one of his/her household members	0	0.0

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 practices 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. The experience from the first three waves of the survey was further used to improve the questionnaire for the 2007 operation. The main modifications in 2007 concerned the inclusion of questions about own consumption, and those necessary for calculating imputed rent in the household questionnaire. In the personal questionnaire the most important improvements concerned the inclusion of the 'education obtained since previous interview' for the longitudinal respondents, simplifying the salary questions by giving the respondent more options for naming it in time and gross/net categories, adding questions about the use of a company car and other non-cash employee income and developing income questions for self-employed persons and entrepreneurs.

The questions on child-care, family benefits and unemployment benefits were also improved.

Other notable modifications concerned the following variables:

1) Personal ID number of household member responsible for dwelling split between owners and tenants

2) Adding in cost of utilities

3) Developing mortgage payment and interest questions

4) Simplifying for the respondent questions about income from rent of property or land

5) Updating lists of social benefits and including question about alimony paid and received

6) Adding in questions to filter out information on incomes, employer social contributions etc for temporary workers and entrepreneurs.

To reduce interviewer-induced measurement errors, the training programme was conducted in 4 smaller groups of about 15 people, with more emphasis on practical work and discussion of mistakes from previous years as in preceding interviewer trainings. All returning interviewers attended a day long training session. During the training, the EU-SILC team briefed the interviewers on all renewals in the questionnaires, discussed previous years' errors and tracing and specifics of assigning household and person numbers in the longitudinal survey. Practical work sessions were conducted in groups of five and each interviewer had to conduct a model interview regarding for a simulated situation using their laptop. At the end of the training session, each interviewer received personal feedback about their mistakes the previous year. Interviewers new to EU-SILC attended a 2 day training session, which included a thorough overview of questionnaires and practical exercises as well as all the topics covered with returning interviewers.

Overall, 58 interviewers were responsible for conducting the interviews. The household– interviewer ratio was 82 households per interviewer.

2.3.2.2. Processing errors

Checking the data was done in three stages: data-entry checks during the interview, additional in-office checks during fieldwork and lastly data cleaning.

The data for 2007 operation was collected using CAPI. The data-entry program was written in Blaise and contained most of the consistency checks. In 2007, the Blaise consistency checks underwent further extensive development, with many new logical checks creating error messages in described situations put in place. In Statistics Estonia, interviewers are required to react in some form to all error messages that occur during interviewing. The solution is either to correct an erroneous situation or if situation is unusual but correct to add a remark to the data entry-program explaining this error. In assessing the quality of interviewers work, not adding a remark to otherwise correct situation is also counted as an error. This way, most of the errors could already be corrected during an interview.

The primary data-entry consistency 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/should not have received a certain type of income reported it or not (e.g. whether households with children received family benefits, retired people (or people below retirement age) received pensions, employed persons received wages and so on);

   b. whether benefits reported to have been received were logical in the age and gender dimensions. For instance student benefits for over 50 year-olds, income taxes for under 15 year-olds, maternity leave and childbirth allowances for men etc;

   c. Whether an educational level attained was possible below a certain age, or educational levels were possible in said combinations for given years;

   d. whether answers provided to different non-monetary deprivation items agreed with each other;

   e. whether households not in an electrical grid could have electrical appliances, or households not connected to a sewerage could have a shower etc ;

   f. whether the relationships in the household matrix were consistent with each other as well as with the age and sex of the household members;

g. whether the difference between the starting and finishing time of the interview was too short or too long and so on.

h. whether reported taxes or medical benefits received were consistent with income

i. membership in pension plans checked by year of birth to see if legally bound to have joined pension pillar.

j. checks for correct survey area, interviewer code and personal numbers matching household numbers.

2) Lower and upper bounds of income variables (incl. benefits). 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 for correct survey area, interviewer code and household and personal numbers matching.

5) Checks not allowing for occupations to be written on too general a scale for coding. (e.g. salesperson, cleaner)

6) Checks for goods produced for own consumption, for instance quantities;

7) 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 the 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 scrutinized one by one.

4) All category 'other' answers were gone through to see if they could be classified under one of the given options.

5) Additionally paid income tax was checked in-household to check for double-reporting.

6) Errors in coding.

7) Study benefits were checked by possibility of obtaining them in the school the respondent attended and legally set amounts.

8) Consistency between time reported working under socio-economic status and months that salary was received. Also time spent in prison.

9) Reported amounts of family benefits were checked compared with eligibility based on the structure of the family and benefit levels set out in legislation.

10) Demographic information in the interviewers' reports was compared to the data recorded in the electronic questionnaires.

All mistakes found through the secondary in-office data editing were put up in a shared excel table, and had to be clarified with the interviewer or interviewee by the end of the fieldwork period. This was done in co-operation by the EU-SILC team and the Data Gathering department's Fieldwork Managers.

The number of primary consistency errors dropped dramatically in 2007 after a special training given to interviewers about the necessity to give an explanatory remark for every consistency check that pops up in Blaise. In 2006 there had been a total of 5654 errors, in 2007 the number had fallen to 1677. There was no such training at the end of 2007, and in 2008 there was a small increase in the number of errors, totalling at 1779. Out of all the errors in 2007, 60% (998 cases) required callback and clarification with the interviewer or interviewee.

As can be seen from table 2.5, the most common types of errors in 2007 had to do with interviewers not correcting their mistakes after an error code had prompted them to do so, and the use of category

'other', while a suitable category existed. These were the categories with the least errors in 2006, and as the error numbers have not increased, it is clear that in all other categories, error counts have decreased considerably.

Table 2.5. Interviewer errors and their processing, 2007

Type of error	Number of errors detected	Share of errors requiring a call-back
No remark explaining unusual situation	28	10
Interviewer made an error, but did not correct it	485	249
Interviewer's remark does not explain unusual situation	57	16
Data not sufficient for coding	89	30
Starting and finishing times recorded incorrectly	9	0
Use of category Other, while a suitable category exists	429	391
In-office checks	105	83
Interviewer has misunderstood a question	226	163
Data entry mistake	16	16
Not interviewers error	195	22
Total	1639	980

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. Extreme values for all income components as well as total income also were checked.

2.3.3. Non-response errors

2.3.3.1. Achieved sample size

Data for 5146 households were accepted into the database and used in analysis. This makes the overall share of complete household interviews accepted for the database 79.0%. On personal level, the share of complete personal interviews within the households accepted for the database was 99.2% – or 11,875 interviews of the possible 11,971. Income data for the remaining 96 persons who did not complete a personal interview was imputed by closest neighbour full record imputation (RB250=14).

For rotational group breakdown see 2.3.3.3.

2.3.3.2. Unit non-response

Indicator	Total sample	New part
Address contact rate (Ra)	0.919	0.843
Proportion of complete household interviews accepted for the database (Rh)	0.878	0.771
Proportion of complete personal interviews within the households accepted for the database (Rp)	0.992	0.990
Household non-response rates (NRh)	19.3	35.0
Individual non-response rates (NRp)	0.8	1.0
Overall individual non-response rates ($*NRp$)	20.0	35.7

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 2.1.1):

DB120=21	DB120=22
▪ Address-person does not live at given address no information on new	▪ The house given is located but given address can not be accessed (due to

address available	locked doors or gates, etc)
▪ Address-person has moved to another address, no information on new address available	▪ Address of address-person can not be accesses due to poor weather conditions etc
▪ Given address does not exist	
▪ Address can be located, but no contact can be made since nobody is at home	

2.3.3.3. Distribution of households by 'record of contact at address' (DB120), by 'household questionnaire result' (DB130) and by 'household interview acceptance' (DB135) for each rotational group and for the total

Table 2.6. Distribution of households by 'record of contact at address' (DB120) for each rotational group and in total, 2007

Record of contact at address	Rotational group 3		Rotational group 4		Rotational group 5		Rotational group 6		Rotational group 7		Total	
	N	%	N	%	N	%	N	%	N	%	N	%
Total (DB120=11 to 23)	879	100,0	900	100,0	584	100,0	1709	100,0	2444	100,0	6516	100,0
Address contacted (DB120=11)	847	96,4	864	96,0	551	94,3	1623	95,0	1979	81,0	5864	90,0
Address non-contacted (DB120=21 to 23)	32	3,6	36	4,0	33	5,7	86	5,0	465	19,0	652	10,0
Total address non-contacted (DB120=21 to 23)	32	100,0	36	100,0	33	100,0	86	100,0	465	100,0	652	100,0
Address cannot be located (DB120=21)	22	68,8	24	66,7	22	66,7	63	73,3	341	73,3	472	72,4
Address unable to access (DB120=22)	0	0,0	5	13,9	2	6,1	9	10,5	28	6,0	44	6,7
Address does not exist or is non-residential address or is unoccupied or not principal residence (DB120=23)	10	31,2	7	19,4	9	27,3	14	16,3	96	20,6	136	20,9

Table 2.7. Distribution of addresses contacted by 'household questionnaire result' and by household interview acceptance, 2007

Household questionnaire result	Rotational group 3		Rotational group 4		Rotational group 5		Rotational group 6		Rotational group 7		Total	
	N	%	N	%	N	%	N	%	N	%	N	%
Total (DB130=11 to 24)	847	100,0	864	100,0	551	100,0	1623	100,0	1979	100,0	5864	100,0
Household questionnaire completed (DB130=11)	801	94,6	830	96,1	513	93,1	1480	91,2	1530	77,3	5154	87,9
Interview not completed (DB130= 21 to 24)	46	5,4	34	3,9	38	6,9	143	8,8	449	22,7	710	12,1
Total interview not completed (DB130=21 to 24)	46	100,0	34	100,0	38	100,0	143	100,0	449	100,0	710	100,0
Refusal to co-operate (DB130=21)	18	39,1	9	26,5	19	50,0	89	62,2	320	71,3	455	64,1
Entire household temporarily away for duration of fieldwork (DB130=22)	4	8,7	2	5,9	5	13,2	9	6,3	49	10,9	69	9,7
Household unable to respond (illness, incapacity, etc) (DB130=23)	4	8,7	2	5,9	4	10,5	9	6,3	32	7,1	51	7,2
Other (DB130=24)	20	43,5	21	61,8	10	26,3	36	25,2	48	10,7	135	19,0
Household questionnaire completed (DB135=1 to 2)	801	100,0	830	100,0	513	100,0	1480	100,0	1530	100,0	5154	100,0
Interview accepted to database (DB135=1)	800	99,9	829	99,9	513	100,0	1478	99,9	1526	99,7	5146	99,8
Interview rejected (DB135=2)	1	0,1	1	0,1	0	0,0	2	0,1	4	0,3	8	0,2

Table 2.8. Distribution of household members by Respondent Status and rotational group (RB245), 2007

Responent Status	Rotational group 3		Rotational group 4		Rotational group 5		Rotational group 6		Rotational group 7		Total	
	N	%	N	%	N	%	N	%	N	%	N	%
Total (RB245=1 to 4)	2280	100.0	2323	100.0	1379	100.0	4159	100.0	4231	100.0	14372	100.0
Current household members aged 16 and over (RB245 = 1)	1944	85.3	1965	84.6	1165	84.5	3458	83.1	3439	81.3	11971	83.3
Selected respondent (RB245=2)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Not selected respondent (RB245=3)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Not eligible person (RB245=4)	336	14.7	358	15.4	214	15.5	701	16.9	792	18.7	2401	16.7

Table 2.9. Distribution of household members by Data Status and rotational group (RB250), 2007

Data Status	Rotational group 3		Rotational group 4		Rotational group 5		Rotational group 6		Rotational group 7		Total	
	N	%	N	%	N	%	N	%	N	%	N	%
Total (RB250=11 to 33)	1944	100.0	1965	100.0	1165	100.0	3458	100.0	3439	100.0	11971	100.0
Information completed only from interview (RB250 = 11)	1934	99.5	1956	99.5	1152	98.9	3429	99.2	3404	99.0	11875	99.2
Information completed only from registers (RB250 = 12)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Information completed from both (RB250 = 13)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Information completed from full-record imputation (RB250=14)	10	0.5	9	0.5	13	1.1	29	0.8	35	1.0	96	0.8
Individual unable to respond (RB250=21)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Individual failed to return self-completed questionnaire (RB250=22)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Refusal to cooperate (RB250=23)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Person temporarily away and no proxy available (RB250=31)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
No contact for other reasons (RB250=32)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Information not completed, reason unknown (RB250=33)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0

Table 2.10. Distribution of household members by Type of Interview and rotational group (RB260), 2007

Responent Status	Rotational group 3		Rotational group 4		Rotational group 5		Rotational group 6		Rotational group 7		Total	
	N	%	N	%	N	%	N	%	N	%	N	%
Total (RB245=1 to 5)	1934	100.0	1956	100.0	1152	100.0	3429	100.0	3404	100.0	11875	100.0
Face to face interview - PAPI (RB260 = 1)	19	1.0	21	1.1	14	1.2	70	2.0	105	3.1	229	1.9
Face to face interview - CAPI (RB260 = 2)	1689	87.3	1705	87.2	1013	87.9	2944	85.9	2936	86.3	10287	86.6
CATI, telephone interview (RB260=3)	3	0.2	2	0.1	2	0.2	7	0.2	8	0.2	22	0.2
Self-administered by respondent (RB260=4)	0	0.0	1	0.1	0	0.0	0	0.0	1	0.0	2	0.0
Proxy interview (RB260=5)	223	11.5	227	11.6	123	10.7	408	11.9	354	10.4	1335	11.2

2.3.3.4. Distribution of substituted units

Substitution was not used.

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.

Income values imputed by full-record imputation are included.

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

Table 2.11. Distribution of item non-response, household-level variables, 2007

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)	5132	99.7	46	0.9	2332	45.4
Total disposable household income (HY020)	5137	99.8	28	0.5	2444	47.6
Total disposable household income before social transfer other than old-age and survivors' benefits (HY022)	5096	99.0	93	1.8	2277	44.7
Total disposable household income before social transfers including old-age and survivors' benefits (HY023)	4823	93.7	228	4.7	2127	44.1
Net income components at household level						
Imputed rent (HY030N)	4961	96.4	4961	100.0	0	0.0
Income from rental of a property or land (HY040N)	79	1.5	1	1.3	0	0.0
Family/ children related allowances (HY050N)	2134	41.5	0	0.0	8	0.4
Social inclusion not elsewhere classified (HY060N)	32	0.6	0	0.0	0	0.0
Housing allowances (HY070N)	102	2.0	13	12.7	0	0.0
Regular inter-household cash transfers received (HY080N)	235	4.6	8	3.4	0	0.0
Alimonies received, compulsory and voluntary (HY081N)	130	2.5	4	3.1	0	0.0
Interest, dividends, profit from capital investments in incorporated business (HY090N)	2005	39.0	1877[1]	93.6	69	3.4
Interest repayments on mortgage	559	10.9	186	33.3	0	0.0

---

[1] Of which 1833 are such that the only capital income of household is dividends from Estonian banks, and these are imputed based on the interval provided by respondent.

(HY100N)						
Income received by people aged under 16 (HY110N)	93	1.8	15	16.1	3	3.2
Regular taxes on wealth (HY120N)	3539	68.8	125	3.5	0	0.0
Regular inter-household cash transfers paid (HY130N)	279	5.4	5	1.8	0	0.0
Alimonies paid, compulsory and voluntary (HY131N)	101	2.0	2	2.0	0	0.0
Tax on income and social contributions, net (HY140N)[2]	0	0.0	0	-	0	-
Repayments/ receipts for tax adjustment (HY145N)	1961	38.1	184	9.4	24	1.2
Gross income components at household level						
Imputed rent (HY030G)	4961	96.4	4961	100.0	0	0.0
Income from rental of a property or land (HY040G)	79	1.5	1	1.3	0	0.0
Family/ children related allowances (HY050G)	2134	41.5	0	0.0	8	0.4
Social inclusion not elsewhere classified (HY060G)	32	0.6	0	0.0	0	0.0
Housing allowances (HY070G)	102	2.0	13	12.7	0	0.0
Regular inter-household cash transfers received (HY080G)	235	4.6	8	3.4	0	0.0
Alimonies received, compulsory and voluntary (HY081G)	130	2.5	4	3.1	0	0.0
Interest, dividends, profit from capital investments in incorporated business (HY090G)	2005	39.0	1877	93.6	69	3.4
Interest repayments on mortgage (HY100G)	559	10.9	186	33.3	0	0.0
Income received by people aged under 16 (HY110G)	93	1.8	15	16.1	3	3.2
Regular taxes on wealth (HY120G)	3539	68.8	125	3.5	0	0.0
Regular inter-household cash transfers paid (HY130G)	279	5.4	5	1.8	0	0.0
Alimonies paid, compulsory and voluntary (HY131G)	101	2.0	2	2.0	0	0.0
Tax on income and social contributions, gross (HY140G)	3844	74.7	3844	100.0	0	0.0

---

[2] Tax on income is not collected. This variable is fully computed at Statistics Estonia based on a person's and a household's income and taxes paid. Computed values are assumed to be gross, net values are set to zeroes in the database.

Table 2.12. Distribution of item non-response, person-level variables, 2007

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	%
Net income components at personal level						
Employee cash or near cash income (PY010N)	6691	55.9	529	7.9	45	0.7
Non-cash employee income (PY020N)	1474	12.3	367	24.9	189	12.8
Company car (PY021N)[3]	114	1.0	114	100.0	0	0.0
Contributions to individual private pension plans (PY035N)	723	6.0	85	11.8	1	0.1
Cash benefits or losses from self employment (PY050N)	808	6.7	108	13.4	5	0.6
Value of goods produced by own-consumption (PY070N)	8029	67.1	8029	100.0	0	0.0
Pension from individual private plans (PY080N)	6	0.1	0	0.0	0	0.0
Unemployment benefits (PY090N)	133	1.1	11	8.3	2	1.5
Old-age benefits (PY100N)	2894	24.2	44	1.5	3	0.1
Survivor's benefits (PY110N)	124	1.0	1	0.8	0	0.0
Sickness benefits (PY120N)	972	8.1	303	31.2	0	0.0
Disability benefits (PY130N)	827	6.9	15	1.8	0	0.0
Education-related benefits (PY140N)	280	2.3	14	5.0	0	0.0
Gross income components at personal level						
Employee cash or near cash income (PY010G)	6691	55.9	529	7.9	45	0.7
Non-cash employee income (PY020G)	1474	12.3	367	24.9	189	12.8
Company car (PY021G)	114	1.0	114	100.0	0	0.0
Employer's social insurance contributions (PY030G)	6445	53.8	6445	100.0	0	0.0
Contributions to individual private pension plans (PY035G)	723	6.0	85	11.8	1	0.1
Cash benefits or losses from self employment (PY050G)	808	6.7	108	13.4	5	0.6
Value of goods produced by own-consumption (PY070G)	8029	67.1	8029	100.0	0	0.0
Pension from individual private plans (PY080G)	6	0.1	0	0.0	0	0.0
Unemployment benefits (PY090G)	133	1.1	11	8.3	2	1.5
Old-age benefits (PY100G)	2894	24.2	44	1.5	3	0.1
Survivor's benefits (PY110G)	124	1.0	1	0.8	0	0.0
Sickness benefits (PY120G)	972	8.1	303	31.2	0	0.0
Disability benefits (PY130G)	827	6.9	15	1.8	0	0.0
Education-related benefits (PY140G)	280	2.3	14	5.0	0	0.0

---

[3] Non-cash income from company car is not collected from respondent, but imputed as the number of months a company car was used multiplied by 2000.

Table 2.13. Total item non-response and number of observations in the sample at unit level of the common cross-sectional EU indicators based on the cross-sectional component of EU–SILC, for equivalised disposable income, 2007

	Number of sample observations (achieved sample size)	Number of sample observations not taken into account due to item non-response	Non-response at individual level (if applicable)	Non-response at household level (number of households)
At-risk-of-poverty rate after social transfers				
Total[4]	14322	0	NA	1234
By age and gender				
men total[4]	6752	0	NA	1234
women total[4]	7570	0	NA	1234
0-15 years[4]	2351	0	NA	1234
16-24 years	2629	0	NA	1234
25-49 years	4610	0	NA	1234
50-64 years	2491	0	NA	1234
65+ years	2241	0	NA	1234
16+ years	11971	0	NA	1234
16-64 years	9730	0	NA	1234
0-64 years[6]	12081	0	NA	1234
men 16-24 years	1369	0	NA	1234
men 25-49 years	2205	0	NA	1234
men 50-64 years	1116	0	NA	1234
men 65+ years	834	0	NA	1234
men 16+ years	5524	0	NA	1234
men 16-64 years	4690	0	NA	1234
men 0-64 years[6]	5918	0	NA	1234
women 16-24 years	1260	0	NA	1234
women 25-49 years	2405	0	NA	1234
women 50-64 years	1375	0	NA	1234
women 65+ years	1407	0	NA	1234
women 16+ years	6447	0	NA	1234
women 16-64 years	5040	0	NA	1234
women 0-64 years[6]	6163	0	NA	1234
By most frequent activity status and gender				
employed	6188	317	0	1234
unemployed	335	317	0	1234
retired	2387	317	0	1234
other inactive	2744	317	0	1234
men, employed	3171	317	0	1234
men, unemployed	223	317	0	1234
men, retired	806	317	0	1234
men, other inactive	1168	317	0	1234

---

[4] Children born in 2007 are excluded (50 persons in total).

women, employed	3017	317	0	1234
women, unemployed	112	317	0	1234
women, retired	1581	317	0	1234
women, other inactive	1576	317	0	1234
By household type[5]				
single, < 65 years	515	0	NA	1234
single, 65+ years	545	0	NA	1234
single, male	316	0	NA	1234
single, female	744	0	NA	1234
single, total	1060	0	NA	1234
2 adults, no children, both < 65	1334	0	NA	1234
2 adults, no children, at least one 65+	1342	0	NA	1234
other households without children	1718	0	NA	1234
single parent, at least one child	709	0	NA	1234
2 adults, 1 child	1758	0	NA	1234
2 adults, 2 children	2084	0	NA	1234
2 adults, 3+ children	1332	0	NA	1234
other households with children	2995	0	NA	1234
households without children	5454	0	NA	1234
households with children	8878	0	NA	1234
By accommodation tenure status				1234
owner or rent-free	13778	0	NA	1234
tenant	594	0	NA	1234
At-risk-of-poverty threshold				
Median of the equivalised disposable household income	14372	0	NA	1234
At-risk-of-poverty threshold - total	14372	0	NA	1234
Inequality of income distribution S80/S20 income quintile share ratio	14372	0	NA	1234
Relative median at-risk-of-poverty gap				
Total	2908	0	NA	1234
By age and gender				
men total	1244	0	NA	1234
women total	1664	0	NA	1234
0-15 years	513	0	NA	1234
16-64 years	1804	0	NA	1234
65+ years	591	0	NA	1234
16+ years	2395	0	NA	1234
men, 16-64 years	840	0	NA	1234
men, 65+ years	147	0	NA	1234
men, 16+ years	987	0	NA	1234
women, 16-64 years	964	0	NA	1234
women, 65+ years	444	0	NA	1234

---

[5] Persons in households where it was impossible to determine household type are excluded (40 persons).

women, 16+ years	1408	0	NA	1234
Dispersion around the risk-of-poverty threshold				
40%	14372	0	NA	1234
50%	14372	0	NA	1234
70%	14372	0	NA	1234
At-risk-of-poverty rate before social transfers except old-age and survivors' benefits –total [6]	14322	0	NA	1234
At-risk-of-poverty rate before social transfers including old-age and survivors' benefits – total [6]	14322	0	NA	1234
Gini coefficient	14372	0	NA	1234
Mean equivalised disposable income	14372	0	NA	1234

Notes:

Item non-response: number of questionnaires with no information on most frequent activity status, when applicable to indicator (317).

Non-response on individual level: not present in 2007 since income information in missing questionnaires is imputed

Non-response on household level: interview not competed, DB130=21 to 24 (710) + interview rejected, DB135=2 (8) + address cannot be located, DB120=21 (472) + address unable to access, DB120=22 (44).

2.4. Mode of data collection

Distribution of household members aged 16 and over by Data Status and by Type of Interview can be found in Tables 2.9 and 2.10 in Section 2.3.3.3.

2.5. Interview duration

Mean interview duration per household: 48 minutes and 21 seconds (interview duration for 8 households and 13 persons was unknown). Thus, mean interview duration per household is lower than the one-hour limit set in Regulation 1177/2003.

3. COMPARABILITY

3.1. Basic concepts and definitions

3.1.1. The reference population

Persons living in collective households are included in the reference population. The share of persons who are living in collective households and who are not at the same time members of some other private household is likely to be very low. Additionally, there is no feasible way to estimate their share in the total population. Thus, the exclusion of these persons is unlikely to affect the comparability and reliability of the estimates.

3.1.2. The private household definition

There were no divergences from the common definition.

3.1.3. The household membership

There were no divergences from the common definition.

---

[6] Children born in 2007 are excluded (50 persons in total).

3.1.4. The income reference period used

There were no divergences from the common definition. The income reference period was the previous calendar year (2006).

3.1.5. The period for taxes on income and social insurance contributions

There were no divergences from the common definition. Tax on income and social insurance contributions, as well as tax repayments and receipts refer to the income received during the income reference period (previous calendar year).

3.1.6. The reference period for taxes on wealth

There were no divergences from the common definition. Taxes on wealth paid during the income reference period (previous calendar year) were recorded.

3.1.7. The lag between the income reference period and current variables

The lag between the income reference period and current variables ranges from 3 to 7 months, thus not exceeding 8 months stipulated in the regulation.

3.1.8. The total duration of the data collection of the sample

Data collection was planned to last from March till June, but had to be extended by a further one month due shortage of interviewers. Thus, data was collected during a 5 month period, although the extension of the fieldwork period did not provide an improvement in the overall response rates by more than a few percentage points.

3.1.9. Basic information on activity status during the income reference period

There were no divergences from the common definition.

3.2. Components of income

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

3.2.1.1. Total household gross income
There were no divergences from the common definition.

3.2.1.2. Total disposable household income
There were no divergences from the common definition.

3.2.1.3. Total disposable household income, before social transfers other than old-age and survivors' benefits
There were no divergences from the common definition.

3.2.1.4. Total disposable household income, before social transfers including old-age and survivors' benefits
There were no divergences from the common definition.

3.2.1.5. Imputed rent
There were no divergences from the common definition. User cost method was employed, as the share of market rents is very small. External data used for modelling refers to survey year and not income year. As sale prices have been rising quickly, imputed rent value may consequently be overestimated compared to other income variables. From 2007, both gross and net values are available for imputed rent.

3.2.1.6. Interest, dividends, profit from capital investments in unincorporated business
There were no divergences from the common definition.

3.2.1.7. Interest paid on mortgages
There were no divergences from the common definition.

3.2.1.8. Income received by people aged under 16
Survivors' benefits received by people aged 15 or less are recorded under variable PY110 (see below).

3.2.1.9. Cash or near-cash employee income
There were no divergences from the common definition.

3.2.1.10. Cash profits or losses from self-employment (including royalties)
There were no divergences from the common definitions. Profits or losses reported in annual accounts for tax purposes were recorded. In the case of unregistered self-employment, the respondents were asked to estimate the income received this way.

3.2.1.11. Value of goods produced for own consumption
There were no divergences from the common definition. Most quantities were imputed from answers provided by respondents and unit costs were taken from the Household Budget Survey. Production costs were deducted from the total price thus obtained for own-consumption goods, and the profits were transferred to the personal level. The transfer was done by dividing the household aggregate characteristic by all members of the household aged 16 or over who answered the personal questionnaire. This value was added to their record as variable PY070N.

3.2.1.12. Unemployment benefits
There were no divergences from the common definition.

3.2.1.13. Survivors' benefits
If more than one household member is eligible for survivors' benefits, the individual benefits are, by default, combined and paid as a single sum to one household member. Due to infeasibility of dividing the survivors' benefit received between household members, the whole benefit is recorded only for the household member to whose account it was transferred. This can marginally affect variable HY110 (income received by those under 16), but has no effect on total household income.

3.2.1.14. Gross monthly earnings for employees
Variable was not recorded, as EU-SILC is not used to calculate the gender pay gap.

3.1.1.15. All other variables not listed above
The definition for dependent child was changed to 0-17 according to Eurostat definition change.

3.2.2. The source or procedure used for the collection of income variables
Income variables were collected via face-to-face interviews at component or where applicable at sub-component level.

3.2.3. The form in which income variables at component level have been obtained
Table 3.1 summarizes mode in which different income variables were collected. It should be noted that where collection of only gross values is indicated designate in fact income components, which are not taxable (HY060, HY070, HY080, HY090, HY100, HY120, HY130, PY035, PY130, PY140), i.e. where gross equals net. The only exception is interest, dividends and profit from capital investments in unincorporated businesses, which were collected in gross. Variables HY040, HY110, PY010 and PY050 were collected as either net or gross, depending on which was easier for the respondent to report. The remaining variables were collected only in net.

Table 3.1. Mode of collection for gross income variables in Estonian EU-SILC 2006 operation

Income component	Collected gross	Collected net of tax and social contributions	Mixed mode net/gross
HY040			X
HY050		X	
HY060	X		
HY070	X		
HY080	X		
HY090	X		
HY100	X		
HY110			X
HY120	X		
HY130	X		
HY140		X	
HY145		X	

Variable	Col 1	Col 2	Col 3
PY010			X
PY020		X	
PY035	X		
PY050			X
PY080		X	
PY090		X	
PY100		X	
PY110		X	
PY120		X	
PY130	X		
PY140	X		

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

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 recorded values. Conversion algorithms were created on the basis of the local tax system. Information as to which taxes were paid on income components were also collected and taken into account in conversions.

4. COHERENCE

This section will compare the EU-SILC 2007 data to various external sources, including EU-SILC 2006, the National Accounts (NA), the Household Budget Survey (HBS), the Labour Force Survey (LFS), wage statistics and social protection statistics.

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

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. From 2006, LFS is carried out using CAPI. 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

In Table 4.1, EU-SILC income data is compared component by component to income data from administrative sources for income year 2006. Table 4.1 presents the comparisons by total amounts and Table 4.2 by number of recipients. Only the income components where definitions are similar enough to warrant comparisons are presented here.

Table 4.1. Total amounts (in kroons) of income components by source of information, income year 2006

Income component	EU-SILC	Other sources*
Net cash or near-cash employee income (PY010N)	56,063,250,586	47,124,543,520
Net non-cash employee income (PY020N)	1,410,000,000	
Gross old-age benefits (PY100G)	10,800,000,000	10,908,510,000
Gross sickness benefits (PY120G)[7]	394,000,000	1,020,080,000
Gross disability benefits (PY130G)	1,321,145,473	1,969,140,000
Gross survivor's benefits (PY110G)	108,053,845	173,320,000

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

---

[7] Monthly in EU-SILC, per leave in administrative sources.

Turning to the cash employee income first, the figure from wage statistics is almost 9 billion kroons lower than its EU-SILC equivalent. When comparing the number of people receiving wages and salaries, it appears that there are some 150,000 persons more in EU-SILC who report this type of income than in wage statistics. 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. That is to say, EU-SILC also catches part-time employment and unofficial earnings, making the amounts received higher and the number of recipients larger. In wage statistics, PY020G is included in PY010G and could not be separated from it for individual analysis. The data concerning wage statistics comes from in-house sources, not administrative registers.

Variable PY100G in EU-SILC also includes pension benefits received from abroad, which tend to be higher than national benefits. The Estonian state at the same time pays old age benefits to its citizens residing abroad while the EU-SILC survey does not have people currently living abroad in its sample. The EU-SILC survey also includes other old-age benefits that are not taken into account in the national administrative sources (such as local benefits provided by the local government to pensioners residing in their municipality).

The old age benefits paid to the institutionalized population are not included in the administrative data sources' total amounts presented in Table 4.1 but they are included in the number of recipients in Table 4.2. The latter explains the higher number of recipients according to national records. Underestimation of number of recipients is probably also related to some below retirement age persons failing to report superannuated pensions in the case of EU-SILC.

Table 4.2. Number of recipients of income components by source of information, income year 2006

Income component	EU-SILC	Other sources*
Person-level components		
Net cash or near-cash employee income (PY010N)	666,494	518,593
Net non-cash employee income (PY020N)	162,593	
Old-age benefits (PY100)	287,568	291,580
Disability benefits (PY130)	66,379	65,477
Net survivor's benefits (PY110N)	8,482	9,537

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

The numbers of people having reported to receive disability benefits in EU-SILC and number of recipients according to national data sources are very similar. The administrative records number should be somewhat higher though as the number given includes only disability and early retirement benefits. The numbers of recipients for care allowances and economic integration of the handicapped are not included, whereas the amounts received by them are included in Table 4.1.

Survivors' benefits as best compared at household level, as they are usually paid to a household as a whole. The administrative figure is somewhat higher than the EU-SILC figure, indicating underestimation in EU-SILC.

Table 4.3 compares the mean and number of recipients of most income components in EU-SILC 2007 to the estimates from 2006 operation. Changes that emerge are, in general, in line with what could be expected. Much like the previous year, 2007 was a year of considerable economic growth and increase in real salaries in Estonia. While the average salary increased by 19%, the increase in wage receivers was more modest – 2%. Comparably, there was a decrease in people receiving unemployment benefits as many discouraged workers found employment and unemployment levels reached record lows. Administrative data confirms the survey results.

Table 4.3. Mean and number of recipients of income components in EU-SILC 2006 and 2007

	Mean		Number of recipients	
	2006	2007	2006	2007
Individual level components				
PY010N	70,768	84,116	653,971	666,494
PY020N	16,843	8,676	26,099	162,593
PY035N	5,207	5,111	78,039	84,401
PY050N	9,587	36,000	73,750	60,224
PY090N	7,693	10,566	16,004	10,836

PY100N	32,398	37,409	296,346	287,568
PY110N	12,533	12,739	10,964	8,482
PY120N	3,580	3,261	80,244	102,774
PY130N	17,362	19,903	65,049	66,379
PY140N	10,659	10,092	18,782	22,284
Household level components				
HY040N	12,067	11,290	6,321	5,726
HY050N	10,724	12,498	190,720	185,502
HY070N	6,066	4,295	12,604	11,790
HY080N	14,156	15,138	21,187	28,870
HY090N	1,420	1,519	183,962	238,307
HY110N	2,428	4,473	5,770	6,092
HY120N	485	430	308,450	356,151
HY130N	12,390	14,728	28,999	33,360
HY145N	-2,390	-2,455	201,894	213,965
HY010	128,581	155,991		
HY020	107,329	130,759		
HY022	102,080	125,883		
HY023	92,554	130,453		

The large differences in PY020N from 2006 to 2007 result from different calculation rules that came to force in 2007. In 2006 PY020N only included company car, but from 2007 onwards other kinds of non-cash income were added to the amount.

Cash benefits from self employment have grown in part due to favourable economic conditions and partly due to a decrease in income tax. The number of entrepreneurs seems to fluctuate between survey years which hints to a relatively big pool of short-lived businesses.

The increase in the mean of PY090N has to do with a new unemployment benefit attached to one's earlier salary which is considerably lower than the fixed unemployment allowance. There was also a small increase in the latter in 2007.

The amounts paid to private pension schemes and the number of people making these payments have both increased, which is to be expected given that funds are actively campaigning for more people to join up.

Old-age benefits have increased due to a small increase of 300 kroons a month to the national pension, while the number of receivers has not changed much.

The survivors' and sickness benefits continued the trend of the previous year with the number of recipients increasing, whereas the actual amounts have not changed much. This is because neither allowance was increased by the government. Average disability benefits have increased while the number of recipients has not changed much, which is in line with what could be expected. Education-related benefits have on the other hand stayed the same while the numbers of recipient have gone up. This is also because the allowance levels are centrally fixed, and allowances did not increase that year. The increase in number of recipients is likely to be caused by a large birth cohort reaching the age of 18 and becoming eligible for university and vocational school student support.

Household level variables reflect changes in line with personal level variables. The average amounts have not changed much for HY145N, which should not fluctuate much regardless of the state of the economy; all other characteristics have had some changes.

For HY040N the decrease in the average amount is probably due a change in the questionnaire making it less likely for the respondent to mix up net or gross amounts. The small decrease in the number of households receiving income from rental of a property or land probably has to do with sample fluctuations.

Family allowances have increased, in compliance with increases of national benefit levels. The average amount of housing allowances has decreased which can be explained by an overall increase in wages, leaving fewer households eligible, and for smaller sums. The number of households receiving and paying transfers from other households has increased. This might have something to do with more help from former household members living and working abroad and sending money home or just a more favourable economic setting where people have greater financial possibilities for helping their relatives.

A rapid increase in income collected through HY090 reflects a positive situation on the financial markets, with more people investing and large returns that have to do with overall economic growth.

More people also had to pay taxes on wealth, but the amounts are fairly small and have not changed substantially. The decrease has to do with more people having to pay tax but on smaller sums.

The general economic picture should also account for the increase in income received by people aged under 16—more short-term and summer jobs were available for young people because in 2006, there was a big shortage of workers.

Total household gross income and net income increased by approximately 21% each in 2006. The increases stem from the higher wages and other income components, most of which have gone up considerably. HY022 and HY023 follow the same pattern.

4.2. Comparison of other target variables with external sources

The differences in the share of household possessing consumer durables are negligible in the case of television sets, washing machines and cars(Table 4.4). The lower levels in HBS in ownership of a telephone and personal computer probably stem from more narrow response categories—only a stationary computer at one's house is included in the question, laptops are not; similarly only a cell phone is included but stationary phones at one's house are not.

Table 4.4. Share of households in possession of various consumer durables based on EU-SILC and the HBS, 2007

Consumer durable	EU-SILC	HBS
Telephone, including mobile phone	95.6	86.6
TV	97.7	97.2
Washing machine	86.2	87.5
Car	50.4	49.9
Personal computer	53.3	51.7

Table 4.5 presents the distribution of households by dwelling type. In 2006,the differences between the two surveys were non-existent regarding this variable. In 2007, EU-SILC has a much higher share of people living in apartments or flats, and an equally lower proportion of the sample living in detached houses or farms. It is especially the larger share of farm households in HBS that is behind this difference.

Table 4.5. Households by the type of dwelling based on EU-SILC and the HBS, 2007

Type of dwelling	EU-SILC	HBS
Detached house	26.5	36.6
Semi-detached or terraced house	4.1	4.6
Apartment or flat	68.7	58.1
Some other kind of accommodation	(0,8)*	(0.7)*
Total	100.0	100.0

* Unreliable estimate, based on 31 sample observations.

In Table 4.6 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 4 and 5. There are somewhat more people with post-secondary non-tertiary education according to EU-SILC and less people with first stage of tertiary education. Given that the questions used in the two surveys are identical, this must be due to sample fluctuations.

Table 4.6. Distribution of population aged 16-74 by ISCED level, based on the EU-SILC and the LFS, 2007

ISCED level	EU-SILC	LFS
0 Pre-primary education	0,6	0,5
1 Primary education	4,2	3,5
2 Lower secondary education	18,6	17,9
3 (Upper) secondary education	44,1	45,8
4 Post-secondary non tertiary education	7,1	5,2
5 First stage of tertiary education	24,9	26,8
6 Second stage of tertiary education	(0,4)*	(0,3)*
Total	99,9	100.0

* Unreliable estimate, based on 35-44 sample observations.

Finally, Table 4.7 presents the comparison of population aged 16 or over by 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 4.7. Distribution of population aged 16 and over by self-defined activity status based on EU-SILC and the HBS, 2007

Activity status	EU-SILC	HBS
Working full-time	54.2	50.8
Working part-time	3.9	3.2
Unemployed	3.2	2.8
Pupil, student	8.8	12.3
In retirement	21.7	21.5
Permanently disabled	3.7	4.1
Fulfilling domestic tasks and care responsibilities	4.4	1.5
Other inactive	(0.0)*	3.4
Total	99.9	99.6

* Unreliable estimate, based on 1 sample observation.