# EU-SILC 2005 in Estonia: Intermediate Quality Report

Tallinn 2006

# CONTENTS

# 1. COMMON CROSS-SECTIONAL EUROPEAN UNION INDICATORS

*Table 1.1.* Laeken indicators and other indicators

| Indicator | Value | Standard error |
|---|---|---|
| **PRIMARY LAEKEN INDICATORS OF SOCIAL COHESION** | | |
| **At-risk-of-poverty rate after social transfers: Total** | 18.3 | 0.64 |
| At-risk-of-poverty rate after social transfers: Male | 17.4 | 0.72 |
| At-risk-of-poverty rate after social transfers: Female | 19.1 | 0.77 |
| At-risk-of-poverty rate after social transfers: 0-15 total | 21.4 | 1.37 |
| At-risk-of-poverty rate after social transfers: 0-15 male | 22.5 | 1.78 |
| At-risk-of-poverty rate after social transfers: 0-15 female | 20.4 | 1.74 |
| At-risk-of-poverty rate after social transfers: 0-64 total | 18.0 | 0.66 |
| At-risk-of-poverty rate after social transfers: 0-64 male | 18.4 | 0.78 |
| At-risk-of-poverty rate after social transfers: 0-64 female | 17.5 | 0.77 |
| At-risk-of-poverty rate after social transfers: 16+ total | 17.7 | 0.65 |
| At-risk-of-poverty rate after social transfers: 16+ male | 16.2 | 0.74 |
| At-risk-of-poverty rate after social transfers: 16+ female | 18.9 | 0.79 |
| At-risk-of-poverty rate after social transfers: 16-64 total | 17.1 | 0.64 |
| At-risk-of-poverty rate after social transfers: 16-64 male | 17.3 | 0.80 |
| At-risk-of-poverty rate after social transfers: 16-64 female | 16.8 | 0.73 |
| At-risk-of-poverty rate after social transfers: 16-24 total | 18.2 | 1.22 |
| At-risk-of-poverty rate after social transfers: 16-24 male | 18.7 | 1.67 |
| At-risk-of-poverty rate after social transfers: 16-24 female | 17.6 | 1.59 |
| At-risk-of-poverty rate after social transfers: 25-49 total | 16.2 | 0.80 |
| At-risk-of-poverty rate after social transfers: 25-49 male | 15.5 | 1.06 |
| At-risk-of-poverty rate after social transfers: 25-49 female | 16.8 | 0.93 |
| At-risk-of-poverty rate after social transfers: 50-64 total | 18.0 | 1.17 |
| At-risk-of-poverty rate after social transfers: 50-64 male | 19.9 | 1.62 |
| At-risk-of-poverty rate after social transfers: 50-64 female | 16.5 | 1.34 |
| At-risk-of-poverty rate after social transfers: 65+ total | 20.3 | 1.66 |
| At-risk-of-poverty rate after social transfers: 65+ male | 9.6 | 1.69 |
| At-risk-of-poverty rate after social transfers: 65+ female | 25.5 | 2.09 |
| At-risk-of-poverty rate after social transfers: At work total | 7.5 | 0.46 |
| At-risk-of-poverty rate after social transfers: At work male | 6.9 | 0.56 |
| At-risk-of-poverty rate after social transfers: At work female | 8.1 | 0.64 |
| At-risk-of-poverty rate after social transfers: Not at work total | 29.7 | 1.18 |
| At-risk-of-poverty rate after social transfers: Not at work male | 29.4 | 1.49 |
| At-risk-of-poverty rate after social transfers: Not at work female | 29.9 | 1.38 |
| At-risk-of-poverty rate after social transfers: Unemployed total | 60.0 | 2.81 |
| At-risk-of-poverty rate after social transfers: Unemployed male | 61.4 | 3.41 |
| At-risk-of-poverty rate after social transfers: Unemployed female | 57.5 | 4.53 |
| At-risk-of-poverty rate after social transfers: Retired total | 22.8 | 1.72 |
| At-risk-of-poverty rate after social transfers: Retired male | 10.8 | 1.77 |
| At-risk-of-poverty rate after social transfers: Retired female | 28.0 | 2.09 |
| At-risk-of-poverty rate after social transfers: Other inactive total | 29.2 | 1.42 |
| At-risk-of-poverty rate after social transfers: Other inactive male | 31.6 | 2.30 |
| At-risk-of-poverty rate after social transfers: Other inactive female | 27.6 | 1.59 |
| At-risk-of-poverty rate after social transfers: No dependent children | 19.0 | 1.01 |
| At-risk-of-poverty rate after social transfers: Single total | 36.4 | 2.45 |
| At-risk-of-poverty rate after social transfers: Single male | 34.8 | 3.83 |
| At-risk-of-poverty rate after social transfers: Single female | 37.2 | 2.96 |
| At-risk-of-poverty rate after social transfers: Single <65 years | 32.5 | 2.90 |
| At-risk-of-poverty rate after social transfers: Single 65+ | 40.9 | 3.73 |
| At-risk-of-poverty rate after social transfers: 2 adults no children, <65 years | 10.6 | 1.74 |
| At-risk-of-poverty rate after social transfers: 2 adults no children, 65+ | 8.8 | 1.66 |
| At-risk-of-poverty rate after social transfers: Other, no dependent children | 17.7 | 1.48 |
| At-risk-of-poverty rate after social transfers: All households with dep. children | 40.3 | 0.90 |
| At-risk-of-poverty rate after social transfers: Single parent | 13.0 | 3.83 |

| | | |
|---|---|---|
| At-risk-of-poverty rate after social transfers: 2 adults 1 dependent child | 12.4 | 1.64 |
| At-risk-of-poverty rate after social transfers: 2 adults 2 dependent children | 25.0 | 1.66 |
| At-risk-of-poverty rate after social transfers: 2 adults 3+ dependent children | 13.0 | 3.24 |
| At-risk-of-poverty rate after social transfers: Other households with children | 21.4 | 1.65 |
| At-risk-of-poverty rate after social transfers: Other households, not possible to determine type | 19.0 | 9.57 |
| At-risk-of-poverty rate after social transfers: Owner or rent-free | 17.5 | 0.65 |
| At-risk-of-poverty rate after social transfers: Tenant | 33.6 | 4.07 |
| **S80/S20 income quintile share ratio** | 5.9 | 0.20 |
| **Relative median at-risk-of-poverty gap: Total** | 24.0 | 1.30 |
| Relative median at-risk-of-poverty gap: Male | 28.6 | 1.85 |
| Relative median at-risk-of-poverty gap: Female | 20.7 | 1.23 |
| Relative median at-risk-of-poverty gap: 0-15 | 30.5 | 2.67 |
| Relative median at-risk-of-poverty gap: 16+ total | 22.4 | 1.20 |
| Relative median at-risk-of-poverty gap: 16+ male | 28.2 | 1.94 |
| Relative median at-risk-of-poverty gap: 16+ female | 19.2 | 1.08 |
| Relative median at-risk-of-poverty gap: 16-64 total | 28.5 | 1.66 |
| Relative median at-risk-of-poverty gap: 16-64 male | 30.0 | 2.11 |
| Relative median at-risk-of-poverty gap: 16-64 female | 27.4 | 1.80 |
| Relative median at-risk-of-poverty gap: 65+ total | 11.0 | 1.02 |
| Relative median at-risk-of-poverty gap: 65+ male | 13.2 | 2.66 |
| Relative median at-risk-of-poverty gap: 65+ female | 10.7 | 1.12 |
| SECONDARY LAEKEN INDICATORS OF SOCIAL COHESION | | |
| **Dispersion around the risk-of-poverty threshold: 40%** | 7.1 | 0.44 |
| **Dispersion around the risk-of-poverty threshold: 50%** | 11.3 | 0.54 |
| **Dispersion around the risk-of-poverty threshold: 70%** | 26.2 | 0.67 |
| **At-risk-of-poverty rate before all transfers** | 38.9 | 0.63 |
| **At-risk-of-poverty rate before transfers including old-age and survivors' benefits** | 24.2 | 0.59 |
| **Gini coefficient** | 34.1 | 0.64 |
| OTHER INDICATORS | | |
| **Equivalised disposable income** | 56790 | 783 |
| **Risk-of-poverty threshold – one person household** | 27981 | 355 |
| **Risk-of-poverty threshold – household with 2 adults and 2 children** | 58760 | 745 |

# 2. ACCURACY

## 2.1. Sample design

### 2.1.1. Type of sampling design

As 2005 operation was the second round of EU-SILC in Estonia, the sample comprised of two parts. The first part consisted of households that were selected for the survey in 2004 and were followed up in 2005 operation. Initially this part consisted of 4 rotational groups, one of which was to be dropped after 2004 operation and not to be included to the survey of 2005 according to the original integrated design. However, due to the smaller than expected response rate in 2004, it was decided to keep all rotational groups in the sample. In addition to the repeated part, a new sample of 997 households was selected to the sample. For this new part stratified unequal probability sampling of households was used, with a household selected with probability proportional to the number of persons aged 14+ in it. It is because a sample of persons aged 14+ (so called address-persons) is selected first with equal probabilities within strata, and then the household of the selected person is identified, and all eligible persons in the household are interviewed.

Unequal probability design is likely to have 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.

### 2.1.2. Sampling units

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

## 2.1.3. Stratification and sub-stratification criteria

While selecting the new part of the sample the register of persons aged 14 or over was sorted and stratified on county level (i.e. geographically). The counties (and capital Tallinn) were grouped into three strata by the population size. Hiiu County formed a separate stratum as the smallest county with the population size times smaller of the next smallest. The rest of the regions were divided into two strata – big counties (Tallinn, Harju, Ida-Viru, Lääne-Viru, Pärnu, Tartu) and small counties (Jõgeva, Järva, Lääne, Põlva, Rapla, Saare, Valga, Viljandi, Võru). Unproportional distribution of old (repeated) part of the sample between counties resulted from the stratification in 2004.

## 2.1.4. Sample size and allocation criteria

Repeated part of the sample comprised of 4,226 households. These are households selected to the survey in 2004, which responded to the survey and split-off households. Due to financial constrains new part of the sample consisted of only 997 households. While selecting the new sample, persons from different strata were included with different probabilities as indicated in Table 2.1. Specified proportions between strata were similar to those of the sample of 2004 operation. $R_h$ stands for the number of persons aged 14 and over living in stratum $h$ as at 01.01.2005, $n_g$ 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 2005

| Stratum h | Counties | $R_g$ | $n_g$ | $n_g/R_g$ % |
|---|---|---|---|---|
| Large | Tallinn, Harju, Ida-Viru, Lääne-Viru, Pärnu, Tartu | 856,437 | 598 | 0.07 |
| Small | Jõgeva, Järva, Lääne, Põlva, Rapla, Saare, Valga, Viljandi, Võru | 288,668 | 366 | 0.13 |
| Hiiu | Hiiu | 9,117 | 33 | 3.62 |

## 2.1.5. Sample selection schemes

The old (repeated) part of the sample resulted from indirect selection of households through the panel persons, i.e. all households that responded in 2004 and those resulting from split-off of these households were included to sample of 2005. The new part of the sample was formed by systematic sampling procedure of address-persons with foregoing sample sizes in each stratum.

## 2.1.6. Sample distribution over time

Fixed income reference period was used and therefore the sample was not principally divided into months or weeks. Although for the convenience of fieldwork administration, the sample was equally distributed along the fieldwork period (April-July 2005). Due to difficulties in getting contact, ca 4% of households were interviewed after the official end of fieldwork period in August and September 2005.

## 2.1.7. Renewal of sample: Rotational groups

The sample of 2004 was randomly divided into 4 rotational groups. For a split-off household the rotational group was set the same as one of original household. As was mentioned before, none of these rotational groups were dropped out of the survey in 2005. The new part of the sample formed new, 5th, rotational group, so in total there are 5 rotational groups in the survey in 2005.

## 2.1.8. Weightings

*2.1.8.1. Detailed description of weighting*

Weighting procedure of Estonian EU-SILC 2005 operation was in line with the standard weighting procedure proposed by Eurostat (document 157/05). Some peculiarities are due to modified rotation scheme. Weighting procedure was as follows.

Step 1. Computation of the panel person design weights (only for panel of wave 2, i.e. repeated part of the sample).

As every rotational group is to be weighted independently, the panel person design weight is the design weight DB080 of the corresponding household in 2004 multiplied by 4.

Step 2. Correction for non-response due to attrition (only for panel of wave 2, i.e. repeated part of the sample).

To correct for non-response due to attrition, panel person design weights of step 1 were divided by the empirical response rate calculated as

$$r_p = \frac{n_2}{n_1 - M} ,$$

where $n_2$ is the number of panel person in the sample of 2005, $n_1$ is the number of panel persons in the sample of 2004 and $M$ is the number of panel persons out of scope for 2005 operation. A person is considered as out of scope if corresponding address cannot be located (DB120=21) or address does not exist, is non-residential or is unoccupied or not principal residence (DB120=23).

Empirical response rate to correct for attrition was calculated by county to take regional differences into account.

Step 3. Computation of the sub-sample household weights.

Case 1. Panel of wave 1, i.e. new part of the sample.

Sub-sample household weight of a household from the new part of the sample is the inverse of its inclusion probability. The latter is proportional to the inclusion probability of the address person, whereas the additional factor depends on the number of household members aged 14 or over. The household having $p$ members aged 14 and over has the inclusion probability $p$ times higher than the household having only one person in that age group. The inclusion probability of the household $h$ in stratum $g$ is

$$\pi_h = \frac{n_g p_h}{R_g} ,$$

Where $p_h$ is the number of persons aged 14 and over in this household, given by the household, strata sample sizes $n_{ag}$ and population counts $R_{ag}$ are given in Table 2.1. For non-responders the number of persons aged 14+ registered to the household's address is taken as $p_h$.

Thus, $w_h = 1 / \pi_h$ .

Case 2. Panel of wave 2, i.e. repeated part of the sample

Sub-sample household weights for repeated part of the sample were calculated using the Generalized Weight Share Method as follows.

$$w_h = \frac{\sum_{j \in h^o} d_j}{N_h} ,$$

where $d_j$ is attrition-corrected design weight of person $j$ (result of step 2), $h^0$ is a set panel persons of household $h$, and $N_h$ is the total number of members of $h$ who have belonged to the target population at the time of selection of the panel.

Step 4. Computation of the sample household design weights

No households made up only of immigrants were found in the sample, so household design weights (DB080) were calculated in the same way for all households:

$$DB080_h = w_h / 5.$$

Division by five is caused by the fact that due to modified rotational scheme, the population of private households in 2005 is represented 5 times (i.e. every rotational group represents it).

Step 5. Correction for non-response at the household level

Non-response was compensated using the post-stratification technique. For post-stratification the counties were divided into groups that are expected to be homogeneous in respect to the measured variables of the population, i.e. in any group (post-stratum) the households are relatively similar by the measured variables and the response probability.

*Table 2.2.* Post-strata, Estonian EU-SILC 2005

| Group k (post-stratum) | Stratum | Counties |
|---|---|---|
| 1 | Large | Tallinn |
| 2 | Large | Ida-Viru, Lääne-Viru |
| 3 | Large | Harju, Pärnu, Tartu |
| 4 | Small | Jõgeva, Põlva, Valga, Viljandi, Võru |
| 5 | Small | Järva, Lääne, Rapla, Saare |
| 6 | Hiiu | Hiiu |

The non-response adjustment factors ($\rho_k$) were calculated as the ratio of the number of sampled households in group $k$ to the number of responding households in group $k$. And finally, the household non-response adjusted weights were calculated as the product of the design weight and the adjustment factor:

$$DB080_h^{corrected} = DB080_h \cdot \rho_k$$

Step 6. Calibration of the household weights to external data source

Weights $DB080_h^{corrected}$ were corrected in the way that the population size estimates of the survey coincide with the official population numbers (less the persons in institutions). The weights were corrected for each household using the calibration method with integrated groups defined by sex, age, residential area and level of urbanization (urban/rural area).

In conclusion the following variables were used in calibration:

- sex (female, male) $\times$ age group (0-12, 13-14, 15, 16-19, 20-24, …, 70-74, 75+)
- county (+Tallinn)
- rural / urban area

The program Bascula exploiting the linear consistent weighting method was used for calibration.

The household weights, which result from the procedure of calibration, are the household cross-sectional weights (DB090).

Step 7. Computation of the personal cross-sectional weights for all household members

Personal cross-sectional weight of a person (RB050) is equal to the cross-sectional weight DB090 of its household.

Step 8. Computation of the personal cross-sectional weights for all household members aged 16 and over

To correct for within-household non-response, the weights RB050 were divided by the within-household empirical response rate of household members aged 16 and over. Next, these adjusted personal weights were calibrated to the same totals as for step 6.

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

| Stratum | Household Size | Mean | Standard Deviation | CV |
|---|---|---|---|---|
| 1 | 1 | 323.3 | 88.2 | 0.27 |
| 1 | 2 | 199.5 | 88.1 | 0.44 |
| 1 | 3 | 172.4 | 58.5 | 0.34 |
| 1 | 4 | 146.9 | 60.1 | 0.41 |
| 1 | 5 | 122.1 | 37.5 | 0.31 |
| 1 | 6 | 95.3 | 46.9 | 0.49 |
| 1 | 7 | 72.6 | 15.0 | 0.21 |
| 1 | 12 | 52.7 | . | . |
| 2 | 1 | 173.5 | 58.4 | 0.34 |
| 2 | 2 | 97.1 | 32.2 | 0.33 |
| 2 | 3 | 81.6 | 36.2 | 0.44 |
| 2 | 4 | 78.4 | 24.6 | 0.31 |
| 2 | 5 | 50.0 | 17.6 | 0.35 |
| 2 | 6 | 49.0 | 17.5 | 0.36 |
| 2 | 7 | 34.4 | 5.0 | 0.15 |
| 2 | 9 | 42.6 | 11.3 | 0.27 |
| 3 | 1 | 59.9 | 24.8 | 0.41 |
| 3 | 2 | 37.4 | 20.7 | 0.55 |
| 3 | 3 | 21.2 | 4.5 | 0.21 |
| 3 | 4 | 26.6 | 13.7 | 0.51 |
| 3 | 5 | 17.5 | 2.3 | 0.13 |
| Total | | 133.6 | 99.8 | 0.75 |

Procedure of calibration results in personal cross sectional weights for all household members aged 16 and over (PB040).

Weights were not scaled at any step; trimming was done only at the stage of calibration. Bounds for correction factor at the calibration stage were from 0.3 to 1.8.

Distributional characteristics of final household cross-sectional weights (DB090) by stratum and household size are presented in Table 2.3.

*2.1.8.2. Substitution*
No substitution was used.

# 2.2. Sampling errors

## 2.2.1. Standard error and effective sample size

To calculate the estimates for standard errors and design effects of the common cross-sectional EU indicators, these indicators were first linearized using the Deville's linearization method and 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. Design effects were estimated with a technique similar to one used in POULPE. Corresponding variances under simple random sampling were obtained with procedure SURVEYMEANS of SAS.

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

*Table 2.4.* Estimates, their standard error, design effect, achieved and effective sample size for the common cross-sectional EU indicators, 2005

| Indicator | Value | Standard error | Achieved sample size (no of households) | Design effect | Effective sample size (no of households) |
|---|---|---|---|---|---|
| At-risk-of-poverty rate after social transfers | 18.3 | 0.64 | 4,151 | 1.05 | 3,953 |
| Inequality of income distribution S80/S20 income quintile share ratio | 17.4 | 0.20 | 4,151 | 1.32 | 3,145 |
| Relative median at-risk-of-poverty gap | 19.1 | 1.30 | 4,151 | 0.97 | 4,279 |
| Dispersion around the risk-of-poverty threshold - 40% | 7.1 | 0.44 | 4,151 | 1.02 | 4,070 |
| Dispersion around the risk-of-poverty threshold - 50% | 11.3 | 0.54 | 4,151 | 1.04 | 3,991 |
| Dispersion around the risk-of-poverty threshold - 70% | 26.2 | 0.67 | 4,151 | 1.08 | 3,844 |
| At-risk-of-poverty rate before social transfers except old-age and survivors' benefits | 38.9 | 0.63 | 4,151 | 1.05 | 3,953 |
| At-risk-of-poverty rate before social transfers including old-age and survivors' benefits - total | 24.2 | 0.59 | 4,151 | 1.10 | 3,774 |
| Gini coefficient | 34.1 | 0.64 | 4,151 | 1.52 | 2,731 |
| Mean equivalised disposable income | 56,790 | 783 | 4,151 | 1.27 | 3,277 |

# 2.3. Non-sampling errors

## 2.3.1. Sampling frame and coverage errors

The basis for 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 having living permission. 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, Citizenship and Migration Board and other governmental organisations.

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

- was dead;
- had moved to another county;
- stayed in an institution permanently (had been there over half a year);
- and none of his/her household was of age 16 or older;
- was surveyed through one of his/her household members;

The amount of the frame errors due to over-coverage in the new part of the sample in 2005 was 49 households, which makes the proportion of the over-coverage in the new part of the sample 4.9% and of the whole sample 0.9% (Table 2.3).

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 incomplete address were dropped out of the register before selecting the sample.

In the new part of the sample of 2005 there were 76 address-persons those address in the population register was definitely wrong and no information on 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 target population nor constitute frame over-coverage. Above mentioned reasons for non-contact are classified under DB120=21.

Due to absence of registration law in Estonia, there is also some under-coverage of persons and addresses 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%.

*Table 2.5.* Reasons for over-coverage in the new part of the sample, 2005

| Frame error | Number of households | Proportion in the frame error (%) |
|---|---|---|
| Total, of which | 49 | 100.0 |
| Address person was dead | 7 | 14.3 |
| Address person has left Estonia | 32 | 65.3 |
| Address person was staying in an institution | 10 | 20.4 |

## 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. The experience from the first wave of the survey was further used to improve the questionnaire for the 2005 operation. The main modifications concerned self-employment income, child-care, 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 callbacks declined and the quality of the information obtained this way was increased due to a remarkably faster pace of the whole cycle.

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 the training, 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. A separate session was held on tracing and specifics of assigning household and person numbers in the longitudinal survey. Interviewers were tested, 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 90 households per interviewer.

*2.3.2.2. Processing errors*
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.

Information on occupations and fields of activity was recorded by open-ended questions at the field and later coded centrally. Altogether almost 400 codes could not be coded based on the recorded information. Where possible, these codes were assigned based on the information from the previous

wave. However, in most cases (78% of errors) a callback was necessary. More than half of all coding mistakes concerned the occupation of parents.

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 and information on parents (module 2005). 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

Data for 4,169 households were accepted for database and used in analysis. This makes the share of complete household interviews accepted for the database 86.8% (of contacted households). The contact rate was 91.9% and the household response rate was 87.64%. On personal level, the share of complete personal interviews within the households accepted for the database was 98.1 – 9,643 interviews (of possible 9,827).

For rotational group breakdown see 2.3.3.3.

### 2.3.3.2. Unit non-response

The final response rates for the total sample were as follows:

Household non-response rate $NRh$ = 18.84

Individual non-response rate $NRp$ = 1.85

Overall non-response rate $*NRp$ = 20.35

Response rates for the new part of the sample were:

Household non-response rate $NRh$ = 34.39

Individual non-response rate $NRp$ = 1.94

Overall non-response rate $*NRp$ = 35.66

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 |
|---|---|
| <ul><li>Address-person does not live at given address and no information is available on new address</li><li>Address-person has moved to another address, no information on new address available</li><li>Given address does not exist</li><li>Address can be located, but no contact can be made since nobody is at home</li></ul> | <ul><li>The house given is located but given address can not be accessed (due to locked doors or gates, etc)</li><li>Address of address-person can not be accesses due to poor weather conditions etc</li></ul> |

### 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, 2005

| | Rotational group 1 | | Rotational group 2 | | Rotational group 3 | | Rotational group 4 | | Rotational group 5 | | Total | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | % | N | % | N | % | N | % | N | % | N | % |
| **Total** (DB120=11 to 23) | 1,063 | 100 | 1,078 | 100 | 1,022 | 100 | 1,063 | 100 | 997 | 100 | 5,223 | 100 |
| Address contacted (DB120=11) | 1,007 | 95 | 1,015 | 94 | 975 | 95 | 1,001 | 94 | 802 | 80 | 4,800 | 92 |
| Address non-contacted (DB120=21 to 23) | 56 | 5 | 63 | 6 | 47 | 5 | 62 | 6 | 195 | 20 | 423 | 8 |
| **Total address non-contacted** (DB120=21 to 23) | 56 | 100 | 63 | 100 | 47 | 100 | 62 | 100 | 195 | 100 | 423 | 100 |
| Address cannot be located (DB120=21) | 41 | 73 | 54 | 86 | 40 | 85 | 49 | 79 | 144 | 74 | 328 | 78 |
| Address unable to access (DB120=22) | 1 | 2 | 2 | 3 | 2 | 4 | 2 | 3 | 2 | 1 | 9 | 2 |
| Address does not exist or is non-residential address or is unoccupied or not principal residence (DB120=23) | 14 | 25 | 7 | 11 | 5 | 11 | 11 | 18 | 49 | 25 | 86 | 20 |

*Table 2.7.* Distribution of addresses contacted by 'household questionnaire result' and by household interview acceptance, 2005

| | Rotational group 1 | | Rotational group 2 | | Rotational group 3 | | Rotational group 4 | | Rotational group 5 | | Total | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | % | N | % | N | % | N | % | N | % | N | % |
| **Total** (DB130=11 to 24) | 1,007 | 100 | 1,015 | 100 | 975 | 100 | 1,001 | 100 | 802 | 100 | 4,800 | 100 |
| Household questionnaire completed (DB130=11) | 896 | 89 | 910 | 90 | 870 | 89 | 906 | 91 | 626 | 78 | 4,208 | 88 |
| Interview not completed (DB130= 21 to 24) | 111 | 11 | 105 | 10 | 105 | 11 | 95 | 9 | 176 | 22 | 592 | 12 |
| **Total interview not completed** (DB130=21 to 24) | 111 | 100 | 105 | 100 | 105 | 100 | 95 | 100 | 176 | 100 | 592 | 100 |
| Refusal to co-operate (DB130=21) | 82 | 74 | 72 | 69 | 71 | 68 | 74 | 78 | 125 | 71 | 424 | 72 |
| Entire household temporarily away for duration of fieldwork (DB130=22) | 16 | 14 | 18 | 17 | 18 | 17 | 13 | 14 | 35 | 20 | 100 | 17 |
| Household unable to respond (illness, incapacity, etc) (DB130=23) | 12 | 11 | 15 | 14 | 15 | 14 | 7 | 7 | 15 | 9 | 64 | 11 |
| Other (DB130=24) | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 4 | 1 |
| **Household questionnaire completed** (DB135=1 to 2) | 896 | 100 | 910 | 100 | 870 | 100 | 906 | 100 | 626 | 100 | 4,208 | 100 |
| Interview accepted to database (DB135=1) | 884 | 99 | 905 | 99 | 862 | 99 | 896 | 99 | 622 | 99 | 4,169 | 99 |
| Interview rejected (DB135=2) | 12 | 1 | 5 | 1 | 8 | 1 | 10 | 1 | 4 | 1 | 39 | 1 |

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

| | | Total (RB245=1 to 4) | Current household members aged 16 and over (RB245 = 1) | Selected respondent (RB245=2) | Not selected respondent (RB245=3) | Not eligible person (RB245=4) |
|---|---|---|---|---|---|---|
| **Rotational group 1** | N | 2,489 | 2,026 | 0 | 0 | 463 |
| | % | 100 | 81.4 | 0 | 0 | 18.6 |
| **Rotational group 2** | N | 2,660 | 2,201 | 0 | 0 | 459 |
| | % | 100 | 82.7 | 0 | 0 | 17.3 |
| **Rotational group 3** | N | 2,539 | 2,102 | 0 | 0 | 437 |
| | % | 100 | 82.8 | 0 | 0 | 17.2 |
| **Rotational group 4** | N | 2,546 | 2,105 | 0 | 0 | 441 |
| | % | 100 | 82.7 | 0 | 0 | 17.3 |
| **Rotational group 5** | N | 1,714 | 1,393 | 0 | 0 | 321 |
| | % | 100 | 81.3 | 0 | 0 | 18.7 |
| **Total** | N | 11,948 | 9,827 | 0 | 0 | 2,121 |
| | % | 100 | 82.2 | 0 | 0 | 17.8 |

## 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.

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

*Table 2.9.* 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) | 4,153 | 99.6 | 28 | 0.7 | 468 | 11.2 |
| Total disposable household income (HY020) | 4,157 | 99.7 | 14 | 0.3 | 586 | 14.1 |
| Total disposable household income before social transfer other than old-age and survivors' benefits (HY022) | 4,098 | 98.3 | 29 | 0.7 | 516 | 12.6 |
| Total disposable household income before social transfers including old-age and survivors' benefits (HY023) | 3,763 | 90.3 | 38 | 1.0 | 494 | 13.1 |
| **Net income components at household level** | | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Income from rental of a property or land (HY040N) | 49 | 1.2 | 0 | 0.0 | 0 | 0.0 |
| Family/ children related allowances (HY050N) | 1,763 | 42.3 | 0 | 0.0 | 0 | 0.0 |
| Social inclusion not elsewhere classified (HY060N) | 25 | 0.6 | 0 | 0.0 | 0 | 0.0 |
| Housing allowances (HY070N) | 142 | 3.4 | 5 | 3.5 | 0 | 0.0 |
| Regular inter-household cash transfers received (HY080N) | 165 | 4.0 | 1 | 0.6 | 0 | 0.0 |
| Interest, dividends, profit from capital investments in incorporated business (HY090N) | 164 | 3.9 | 42 | 25.6 | 5 | 3.0 |
| Income received by people aged under 16 (HY110N) | 92 | 2.2 | 0 | 0.0 | 0 | 0.0 |
| Regular taxes on wealth (HY120N) | 2,650 | 63.6 | 41 | 1.5 | 0 | 0.0 |
| Regular inter-household cash transfers paid (HY130N) | 247 | 5.9 | 4 | 1.6 | 0 | 0.0 |
| Tax on income and social contributions, net (HY140N) | 0 | 0.0 | 0 | . | . | . |
| Repayments/ receipts for tax adjustment (HY145N) | 1,363 | 32.7 | 88 | 6.5 | 25 | 1.8 |
| **Gross income components at household level** | | | | | | |
| Income from rental of a property or land (HY040G) | 49 | 1.2 | 0 | 0.0 | 0 | 0.0 |
| Family/ children related allowances (HY050G) | 1,763 | 42.3 | 0 | 0.0 | 0 | 0.0 |
| Social inclusion not elsewhere classified (HY060G) | 25 | 0.6 | 0 | 0.0 | 0 | 0.0 |
| Housing allowances (HY070G) | 142 | 3.4 | 5 | 3.5 | 0 | 0.0 |
| Regular inter-household cash transfers received (HY080G) | 165 | 4.0 | 1 | 0.6 | 0 | 0.0 |
| Interest, dividends, profit from capital investments in incorporated business (HY090G) | 164 | 3.9 | 42 | 25.6 | 5 | 3.0 |
| Income received by people aged under 16 (HY110G) | 92 | 2.2 | 0 | 0.0 | 0 | 0.0 |
| Regular taxes on wealth (HY120G) | 2,650 | 63.6 | 41 | 1.5 | 0 | 0.0 |
| Regular inter-household cash transfers paid (HY130G) | 247 | 5.9 | 4 | 1.6 | 0 | 0.0 |
| Tax on income and social contributions, gross (HY140G) | 3,032 | 72.7 | 3,032 | 100.0[1] | 0 | 0 |

---

[1] Actual amounts are not collected from respondents. The value is calculated based on whether different taxes were paid or not.

*Table 2.10.* 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 | % |
| **Net income components at personal level** | | | | | | |
| Employee cash or near cash income (PY010N) | 5,095 | 52.8 | 248 | 4.9 | 28 | 0.5 |
| Non-cash employee income (PY020N) | 162 | 1.7 | 0 | 0 | 159 | 98.1 |
| Contributions to individual private pension plans (PY035N) | 437 | 4.5 | 81 | 18.5 | 2 | 0.5 |
| Cash benefits or losses from self employment (PY050N) | 687 | 7.1 | 68 | 9.9 | 16 | 2.3 |
| Pension from individual private plans (PY080N) | 8 | 0.1 | 1 | 12.5 | 0 | 0.0 |
| Unemployment benefits (PY090N) | 172 | 1.8 | 5 | 2.9 | 0 | 0.0 |
| Old-age benefits (PY100N) | 2,409 | 25.0 | 14 | 0.6 | 1 | 0.04 |
| Survivor's benefits (PY110N) | 95 | 1.0 | 2 | 2.1 | 0 | 0.0 |
| Sickness benefits (PY120N) | 498 | 5.2 | 48 | 9.6 | 0 | 0.0 |
| Disability benefits (PY130N) | 639 | 6.6 | 2 | 0.3 | 0 | 0.0 |
| Education-related benefits (PY140N) | 179 | 1.9 | 1 | 0.6 | 0 | 0.0 |
| **Gross income components at personal level** | | | | | | |
| Employee cash or near cash income (PY010G) | 5,095 | 52.8 | 248 | 4.9 | 30 | 0.6 |
| Non-cash employee income (PY020G) | 162 | 1.7 | 0 | 0 | 159 | 98.1 |
| Employer's social insurance contributions (PY030G) | 4,900 | 50.8 | 4,900 | 100.0[2] | 0 | 0.0 |
| Contributions to individual private pension plans (PY035G) | 437 | 4.5 | 81 | 18.5 | 2 | 0.5 |
| Cash benefits or losses from self employment (PY050G) | 714 | 7.4 | 95 | 13.3 | 16 | 2.3 |
| Pension from individual private plans (PY080G) | 8 | 0.1 | 1 | 12.5 | 0 | 0.0 |
| Unemployment benefits (PY090G) | 172 | 1.8 | 5 | 2.9 | 0 | 0.0 |
| Old-age benefits (PY100G) | 2,409 | 25.0 | 14 | 0.6 | 1 | 0.04 |
| Survivor's benefits (PY110G) | 95 | 1.0 | 2 | 2.1 | 0 | 0.0 |
| Sickness benefits (PY120G) | 498 | 5.2 | 48 | 9.6 | 0 | 0.0 |
| Disability benefits (PY130G) | 639 | 6.6 | 2 | 0.3 | 0 | 0.0 |
| Education-related benefits (PY140G) | 179 | 1.9 | 1 | 0.6 | 0 | 0.0 |

---

[2] Actual amounts are not collected from respondents. The value is calculated based on whether different taxes were paid or not.

*Table 2.11.* 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, 2005

| | 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 | 11,884 | 64 | NA | 968 |
| *By age and gender* | | | | |
| men total | 5,554 | 64 | NA | 968 |
| women total | 6,330 | 64 | NA | 968 |
| 0-15 years[3] | 2,083 | 64 | NA | 968 |
| 16-24 years | 2,104 | 64 | NA | 968 |
| 25-49 years | 3,753 | 64 | NA | 968 |
| 50-64 years | 2,135 | 64 | NA | 968 |
| 65+ years | 1,774 | 64 | NA | 968 |
| 16+ years | 9,766 | 64 | NA | 968 |
| 16-64 years | 7,992 | 64 | NA | 968 |
| 0-64 years | 10,075 | 64 | NA | 968 |
| men 16-24 years | 1,049 | 64 | NA | 968 |
| men 25-49 years | 1,788 | 64 | NA | 968 |
| men 50-64 years | 970 | 64 | NA | 968 |
| men 65+ years | 666 | 64 | NA | 968 |
| men 16+ years | 4,473 | 64 | NA | 968 |
| men 16-64 years | 3,807 | 64 | NA | 968 |
| men 0-64 years | 4,868 | 64 | NA | 968 |
| women 16-24 years | 1,055 | 64 | NA | 968 |
| women 25-49 years | 1,965 | 64 | NA | 968 |
| women 50-64 years | 1,165 | 64 | NA | 968 |
| women 65+ years | 1,108 | 64 | NA | 968 |
| women 16+ years | 5,293 | 64 | NA | 968 |
| women 16-64 years | 4,185 | 64 | NA | 968 |
| women 0-64 years | 5,207 | 64 | NA | 968 |
| **By most frequent activity status and gender** | | | | |
| employed | 4,854 | 203 | 146 | 968 |
| unemployed | 467 | 203 | 146 | 968 |
| retired | 1,974 | 203 | 146 | 968 |
| other inactive | 2,186 | 203 | 146 | 968 |
| men, employed | 2,444 | 203 | 146 | 968 |
| men, unemployed | 299 | 203 | 146 | 968 |
| men, retired | 676 | 203 | 146 | 968 |
| men, other inactive | 903 | 203 | 146 | 968 |

---

[3] Children born in 2005 are excluded (35 persons).

| | | | | |
|---|---|---|---|---|
| women, employed | 2,410 | 203 | 146 | 968 |
| women, unemployed | 168 | 203 | 146 | 968 |
| women, retired | 1,298 | 203 | 146 | 968 |
| women, other inactive | 1,283 | 203 | 146 | 968 |
| **By household type[4]** | | | | |
| single, < 65 years | 375 | 64 | NA | 968 |
| single, 65+ years | 397 | 64 | NA | 968 |
| single, male | 227 | 64 | NA | 968 |
| single, female | 545 | 64 | NA | 968 |
| single, total | 772 | 64 | NA | 968 |
| 2 adults, no children, both < 65 | 1,066 | 64 | NA | 968 |
| 2 adults, no children, at least one 65+ | 1,058 | 64 | NA | 968 |
| other households without children | 1,309 | 64 | NA | 968 |
| single parent, at least one child | 665 | 64 | NA | 968 |
| 2 adults, 1 child | 1,509 | 64 | NA | 968 |
| 2 adults, 2 children | 1,748 | 64 | NA | 968 |
| 2 adults, 3+ children | 1,122 | 64 | NA | 968 |
| other households with children | 2,550 | 64 | NA | 968 |
| households without children | 4,205 | 64 | NA | 968 |
| households with children | 7,594 | 64 | NA | 968 |
| By accommodation tenure status | | 64 | | |
| owner or rent-free | 11,386 | 65 | NA | 968 |
| tenant | 497 | 65 | NA | 968 |
| **At-risk-of-poverty threshold** | | | | |
| Median of the equivalised disposable household income | 11,884 | 64 | NA | 968 |
| At-risk-of-poverty threshold - total | 11,884 | 64 | NA | 968 |
| **Inequality of income distribution S80/S20 income quintile share ratio** | 11,884 | 64 | NA | 968 |
| **Relative median at-risk-of-poverty gap** | | 64 | | |
| Total | 2,326 | 64 | NA | 968 |
| *By age and gender* | | | | |
| men total | 1,058 | 64 | NA | 968 |
| women total | 1,268 | 64 | NA | 968 |
| 0-15 years | 531 | 64 | NA | 968 |
| 16-64 years | 1,504 | 64 | NA | 968 |
| 65+ years | 291 | 64 | NA | 968 |
| 16+ years | 1,795 | 64 | NA | 968 |
| men, 16-64 years | 713 | 64 | NA | 968 |
| men, 65+ years | 62 | 64 | NA | 968 |
| men, 16+ years | 775 | 64 | NA | 968 |
| women, 16-64 years | 791 | 64 | NA | 968 |
| women, 65+ years | 229 | 64 | NA | 968 |

---

[4] Persons in households where it was impossible to determine household type are excluded (85 persons).

| | | | | |
|---|---|---|---|---|
| women, 16+ years | 1,020 | 64 | NA | 968 |
| **Dispersion around the risk-of-poverty threshold** | | | | |
| 40% | 11,884 | 64 | NA | 968 |
| 50% | 11,884 | 64 | NA | 968 |
| 70% | 11,884 | 64 | NA | 968 |
| **At-risk-of-poverty rate before social transfers except old-age and survivors' benefits –total** | 11,884 | 64 | NA | 968 |
| **At-risk-of-poverty rate before social transfers including old-age and survivors' benefits – total** | 11,884 | 64 | NA | 968 |
| **Gini coefficient** | 11,884 | 64 | NA | 968 |
| **Mean equivalised disposable income** | 11,884 | 64 | NA | 968 |

Notes:

Item non-response: number of eligible persons in households with missing HY025 (64) + number of questionnaires with no information on most frequent activity status, when applicable to indicator (139) + number of questionnaires with no information on tenure status, when applicable to indicator (1).

Non-response on individual level: individual questionnaire missing (in households with non-missing HY025), when applicable to indicator (146).

Non-response on household level: interview not competed, DB130=21 to 24 (592) + interview rejected, DB135=2 (39) + address cannot be located, DB120=21 (328) + address unable to access, DB120=22 (9).

## 2.4. Mode of data collection

*Table 2.12.* Distribution of household members aged 16 and over by Data Status and rotational group (RB250)
HOUSEHOLD MEMBERS 16+ (RB245= 1 to 3)

| | Total | RB250=11 | =12 | =13 | =21 | =22 | =23 | =31 | =32 | =33 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Total** | 9,827 | 9,643 | 0 | 0 | 15 | 118 | 0 | 27 | 24 | 0 |
| **%** | 100.0 | 98.1 | 0.0 | 0.0 | 0.2 | 1.2 | 0.0 | 0.3 | 0.2 | 0.0 |
| **Rotation group 1** | 2,026 | 1,995 | 0 | 0 | 1 | 16 | 0 | 6 | 8 | 0 |
| **%** | 100.0 | 98.5 | 0.0 | 0.0 | 0.1 | 0.8 | 0.0 | 0.3 | 0.4 | 0.0 |
| **Rotation group 2** | 2,201 | 2,156 | 0 | 0 | 9 | 24 | 0 | 5 | 7 | 0 |
| **%** | 100.0 | 98.0 | 0.0 | 0.0 | 0.4 | 1.1 | 0.0 | 0.2 | 0.3 | 0.0 |
| **Rotation group 3** | 2,102 | 2,067 | 0 | 0 | 0 | 31 | 0 | 4 | 0 | 0 |
| **%** | 100.0 | 98.3 | 0.0 | 0.0 | 0.0 | 1.5 | 0.0 | 0.2 | 0.0 | 0.0 |
| **Rotation group 4** | 2,105 | 2,061 | 0 | 0 | 4 | 25 | 0 | 8 | 7 | 0 |
| **%** | 100.0 | 97.9 | 0.0 | 0.0 | 0.2 | 1.2 | 0.0 | 0.4 | 0.3 | 0.0 |
| **Rotation group 5** | 1,393 | 1,364 | 0 | 0 | 1 | 22 | 0 | 4 | 2 | 0 |
| **%** | 100.0 | 98.0 | 0.0 | 0.0 | 0.1 | 1.5 | 0.0 | 0.3 | 0.2 | 0.0 |

*Table 2.13.* Distribution of household members aged 16 and over by 'Type of Interview and rotational group' (RB260)
HOUSEHOLD MEMBERS 16+ (RB245= 1 to 3) and RB250= 11 or 13

|  | Total | RB260=1 | =2 | =3 | =4 | =5 | Missing |
|---|---|---|---|---|---|---|---|
| **Total** | 825 | 8,300 | 25 | 9 | 479 | 5 | 9,643 |
| **%** | 8.6 | 86.1 | 0.3 | 0.1 | 5.0 | 0.1 | 100.0 |
| **Rotation 1** | 155 | 1,737 | 6 | 4 | 92 | 1 | 1,995 |
| **%** | 7.8 | 87.1 | 0.3 | 0.2 | 4.6 | 0.1 | 100.0 |
| **Rotation 2** | 190 | 1,835 | 7 | 0 | 123 | 1 | 2,156 |
| **%** | 8.8 | 85.1 | 0.3 | 0 | 5.7 | 0.1 | 100.0 |
| **Rotation 3** | 191 | 1,781 | 2 | 0 | 92 | 1 | 2,067 |
| **%** | 9.2 | 86.2 | 0.1 | 0 | 4.5 | 0.1 | 100.0 |
| **Rotation 4** | 175 | 1,790 | 4 | 3 | 89 | 0 | 2,061 |
| **%** | 8.5 | 86.9 | 0.2 | 0.2 | 4.3 | 0 | 100.0 |
| **Rotation 5** | 114 | 1,157 | 6 | 2 | 83 | 2 | 1,364 |
| **%** | 8.4 | 84.8 | 0.4 | 0.2 | 6.1 | 0.2 | 100.0 |

## 2.5. Interview duration

Mean duration of household interview: 16 minutes and 25 seconds
Mean interview duration per household: 57 minutes and 57 seconds
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.

### 3.1.4. The income reference period used

There were no divergences from the common definition. The income reference period was last calendar year (2004).

### 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 reference period (last 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 (last 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 9 months. Nevertheless, in the case of 96% of the households, the difference is 7 months and less.

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

Due to the launch of CAPI as the new method for data collection, the start of the fieldwork was postponed by one month – from March to April. In March, interviews with 27 households were carried out in the course of the testing of CAPI and data-entry program. The official fieldwork period lasted from April 2005 to July 2005. In a few interviewing areas where some organisational problems appeared and the fieldwork period was further extended until early September. The prolonged fieldwork period concerned only a minor fraction of households – 96% of households were interviewed during the official fieldwork period and only 7 households were visited in September.

### 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*
Income received by people under 16 other than wages and salaries and survivors' benefits is not recorded. Nevertheless, it is extremely unlikely that this omission affects the estimation of total household income in any meaningful way (See 3.2.1.8.). Also, some interest payments on current accounts are in all likelihood not reported. Yet the amounts involved are very small in relation to total income. Thus, the effect on total income is again likely to be minor (See 3.2.1.6.).

*3.2.1.2. Total disposable household income*
See 3.2.1.1.

*3.2.1.3. Total disposable household income, before social transfers other than old-age and survivors' benefits*
See 3.2.1.1.

*3.2.1.4. Total disposable household income, before social transfers including old-age and survivors' benefits*
See 3.2.1.1.

*3.2.1.5. Imputed rent*
Variable was not recorded.

*3.2.1.6. Interest, dividends, profit from capital investments in unincorporated business*
It is likely that at least some of the interest paid on current accounts have gone unreported. Yet these sums are, as a rule, very small – usually less than 4 EUR a year. It is thus likely to have only negligible impact on total income.

*3.2.1.7. Interest paid on mortgages*
Variable was not recorded.

*3.2.1.8. Income received by people aged under 16*
Only wages and salaries received by people under 16 were recorded, other types of income received by these persons were not collected. Yet as the vast majority of income received by people under 16 is likely to be in the form of wage and salaries, the effect of this omission on the total income is likely to be small[5]. Also, 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*
Compensation paid to the unemployed participating in community placement could not be separated from other unemployment benefits and is recorded under PY090 rather than under variable PY010.

---

[5] In 2006 operation where this error was corrected, no household member aged 15 or less reported receiving any income other than wages and survivors' benefits.

The share of people of total population receiving this type of benefit in 2004 was less than 0.03%. This misclassification is thus unlikely to affect the comparability of results.

### 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
Variable was not recorded.

### 3.2.1.12. Unemployment benefits
See cash or near cash employee income.

### 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 calculated gender pay gap.

### 3.1.1.15. All other variables not listed above
There were no divergences from common definitions.


## 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. summarises 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, HY130, 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, PY020 and PY050 were collected as both net and gross or if interviewee knew only one of them as either net or gross. The remaining variables were collected only in net.

*Table 3.1.* Mode of collection for gross income variables in Estonian EU-SILC 2004 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 | | | |
| HY110 | | | X |
| HY130 | X | | |
| HY140 | | X | |
| PY010 | | | X |
| PY020 | | 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 bases of recorded values. Conversion algorithms were created on the bases 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 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.

*The HBS* is a continuous survey of households, which has been carried since 1996. The yearly sample size is approximately 4,500 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

Table 4.1 compares the aggregate amounts of income components in EU-SILC and National Accounts. The total wages and salaries are by 12% lower in EU-SILC compared to NA. That is to be expected, given that NA figure also includes non-cash employee income. Income from rental of property or land is substantially higher in EU-SILC, whereas the property income is considerably underestimated. The total transfers are 13% higher in EU-SILC. Total transfers as calculated from HBS are very similar to EU-SILC: 11,701.0 millions of kroons in 2004.

*Table 4.1.* Total income in EU-SILC and NA in millions of kroons, 2004

| Type of income | Income component in EU-SILC | Total in EU-SILC | Total in NA | Distributive transaction in NA |
|---|---|---|---|---|
| Wages and salaries | PY010G | 47,711.1 | 53,919.6 | D11 |
| Income from rental | HY040G | 73.9 | 22.2 | D45 |
| Property income | HY090G | 108.0 | 4,424.8 | D41+D42 |
| Transfers | PY090G + PY100 + PY110G + PY120G + PY130G + PY140G + HY050N | 11,728.3 | 10,387.8 | D621 |

Next, EU-SILC income data is compared component by component to income data from HBS and administrative sources for income year 2004. Table 4.2 presents the comparisons by average amounts and Table 4.3 by number of recipients. Only the income components where definitions are similar enough to warrant comparisons are presented here.

Turning to the cash employee income first, the average amount is by 3% higher in EU-SILC than compared to 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).

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 and the number of recipients is somewhat lower in EU-SILC than in two other sources. 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 number of recipients also agrees well with the administrative data, whereas the amounts seem to be somewhat overestimated in EU-SILC.

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 higher in EU-SILC.

*Table 4.2.* Average amounts of income components by source of information, 2004

| Income component | EU-SILC | HBS | Other sources* |
|---|---|---|---|
| *Person-level components* | | | |
| Net cash or near-cash employee income (PY010N) | 60,088 | 58,231 | 68,100 |
| Net non-cash employee income (PY020N) | 17,165 | 10,092 | |
| Net old-age benefits (PY100N) | 27,771 | 27,155 | 26,431 |
| Net survivor's benefits (PY110N) | 11,747 | 12,533 | 8,484 |
| Net sickness benefits (monthly) | 3,703 | 1,542 | |
| Net disability benefits (PY130N) | 12,267 | 13,281 | |
| *Household-level components* | | | |
| Contributions to individual private pension plans (PY035N) | 4,959 | 4,583 | |
| Net cash profits or losses from self-employment (PY050N) | 17,791 | 41,722 | |
| Net income from rental of property or land (HY040N) | 11,758 | 24,721 | |
| Family/children related allowances (HY050N) | 10,424 | 9,570 | 11,936 |
| Housing allowances (HY070N) | 5,783 | 10,829 | |
| Regular inter-household cash transfer received (HY080N) | 13,827 | 14,014 | |
| Net property income (HY090N) | 4,989 | 6,774 | |
| Regular inter-household transfers paid (HY130N) | 13,278 | 13,654 | |
| Total disposable household income (HY020) | 91,362 | 80,988 | |
| Total disposable household income, before social transfers other than old-age and survivor's benefits (HY022) | 84,756 | 74,872 | |
| Total disposable household income, before social transfers including old-age and survivor's benefits (HY023) | 69,877 | 60,254 | |
| Total disposable per capita income | 39,636 | 35,152 | |

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

Although contributions to individual private pension plans as well as self-employment income are person-level variables in EU-SILC, the comparison with HBS is only possible on household level. Tables 4.2 and 4.3 shows that while the average amounts of payments to private pension plans are similar in both surveys, the number households making payments is twice as high in EU-SILC. In the

case of self-employment income the opposite is true, however. The number of households in recipience is nearly identical, whereas the average amounts are substantially higher in HBS. This is again related to irregular nature of this income component as well as different ways of collecting self-employment income.

The number of households receiving rental income is very similar in two surveys, but the amounts are substantially higher in HBS. In the case of property income, however, the number of recipients is twice as high in EU-SILC. The cash transfers paid to and received from other households inhibit a different pattern: the average amounts are nearly identical, whereas number of households in recipience is considerably higher in EU-SILC as compared to HBS.

The figures for family benefits should be compared with some caution, because HBS data is missing maternity benefits and administrative sources are also in part short of the same benefit as far as the number of recipients is concerned. Given these qualifications, it appears that data from all three sources is reasonably similar. The same cannot be said about the housing benefits, however. The differences that emerge are probably related to the irregular nature of this component.

These component level differences between EU-SILC and HBS also translate into differences in total income. As can be seen from Table 4.2, the average disposable income is by 11% higher in EU-SILC as compared to the respective figure from HBS. It must be taken into account in this juncture, that well-off households are more likely to non-respond to diaries in HBS and thus to give detailed information about their income. In 2004, 31% of households in the highest as opposed to 25% in the lowest quintile refused to fill out the diaries.

In conclusion, the coherence between EU-SILC and HBS is not easy to assess due to different definitions and extrapolation of monthly income to yearly income. When an omission is made to these instances, it appears, that the number of people receiving a benefit is in most cases higher in EU-SILC, whereas the amounts received are either the same or moderately lower in HBS. Sickness and unemployment benefits, as small and non-salient types of income seem to suffer from a recall bias in EU-SILC, while the other income components are reasonably coherent with HBS administrative data.

*Table 4.3.* Number of recipients of income components by source of information, 2004

| Income component | EU-SILC | HBS | Other sources* |
|---|---|---|---|
| *Person-level components* | | | |
| Net cash or near-cash employee income (PY010N) | 624,902 | 539,411 | 474,061 |
| Net non-cash employee income (PY020N) | 22,319 | 7,829 | |
| Net unemployment benefits (PY090N) | 12,826 | 13,942 | 39,338 |
| Net old-age benefits (PY100N) | 293,390 | 296,346 | 301,658 |
| Net survivor's benefits (PY110N) | 10,406 | 10,964 | 11,012 |
| Net sickness benefits (PY120N) | 62,161 | 12,420 | |
| Net disability benefits (PY130N) | 92,518 | 86,068 | |
| *Household-level components* | | | |
| Contributions to individual private pension plans (PY035N) | 55,398 | 26,899 | |
| Net cash profits or losses from self-employment (PY050N) | 48,370 | 49,037 | |
| Net income from rental of property or land (HY040N) | 4,653 | 4,026 | |
| Family/children related allowances (HY050N) | 190,368 | 176,341 | 189,007 |
| Housing allowances (HY070N) | 15,870 | 11,080 | |
| Regular inter-household cash transfer received (HY080N) | 27,402 | 10,151 | |
| Net property income (HY090N) | 21,650 | 10,012 | |
| Regular inter-household transfers paid (HY130N) | 34,867 | 8,974 | |

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

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

The differences in the share of household possessing various consumer durables are negligible. In most cases, the difference is less than 1% and within the standard error of estimate. The only durable where the differences are more noticeable are computers. This is not unexpected, as computer ownership is steadily increasing among the households and EU-SILC data reflects the situation in the first half of the year, whereas HBS estimate is an average for a whole year.

*Table 4.4.* Share of households in possession of various consumer durables based on EU-SILC and the HBS, 2005

| Consumer durable | EU-SILC | HBS |
|---|---|---|
| Telephone, including mobile phone | 91.7 | 92.9 |
| TV | 97.3 | 96.5 |
| Internet | 34.2 | 34.5 |
| Washing machine | 82.4 | 83.1 |
| Car | 45.7 | 45.2 |
| Personal computer | 40.4 | 42.1 |

Table 4.5 presents the distribution of households by dwelling type. In general, the differences between data from two surveys are small, the greatest divergence appearing in the case of detached houses.

*Table 4.5.* Households by the type of dwelling based on EU-SILC and the HBS, 2005

| Type of dwelling | EU-SILC | HBS |
|---|---|---|
| Detached house | 27.2 | 25.4 |
| Semi-detached or terraced house | 3.9 | 4.1 |
| Apartment or flat | 67.8 | 68.2 |
| Some other kind of accommodation | (1.1)* | 2.3 |
| Total | 100.0 | 100.0 |

* Unreliable estimate, based on 20-39 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 less people with post-secondary non-tertiary education according to EU-SILC and more 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, 2005

| ISCED level | EU-SILC | LFS |
|---|---|---|
| 0 Pre-primary education | 0.5 | 0.4 |
| 1 Primary education | 2.4 | 3.1 |
| 2 Lower secondary education | 18.2 | 17.5 |
| 3 (Upper) secondary education | 45.5 | 44.8 |
| 4 Post-secondary non tertiary education | 5.8 | 8.7 |
| 5 First stage of tertiary education | 27.0 | 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 4.7 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 4.7.* Distribution of population aged 16 and over by self-defined activity status based on EU-SILC and the HBS, 2005

| Activity status | EU-SILC | HBS |
|---|---|---|
| Working full-time | 50.8 | 51.1 |
| Working part-time | 4.0 | 3.2 |
| Unemployed | 5.6 | 5.2 |
| Pupil, student | 9.4 | 9.7 |
| In retirement | 22.4 | 21.5 |
| Permanently disabled | 3.7 | 4.3 |
| Fulfilling domestic tasks and care responsibilities | 4.1 | 3.4 |
| Other inactive | …* | 1.6 |
| Total | 100.0 | 100.0 |

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