



STATISTIKOS DEPARTAMENTAS
STATISTICS LITHUANIA

INTERMEDIATE QUALITY REPORT EU-SILC 2007 OPERATION

Vilnius 2008

CONTENTS

1.	Common cross-sectional European Union indicators.....	3
2.	Accuracy.....	5
2.1.	Sample design.....	5
2.1.1.	Type of sampling design.....	5
2.1.2.	Sampling units.....	5
2.1.3.	Stratification criteria.....	5
2.1.4.	Sample size.....	5
2.1.5.	Sample selection schemes.....	5
2.1.6.	Sample distribution over time.....	5
2.1.7.	Renewal of sample: Rotational groups.....	6
2.1.8.	Weightings.....	6
2.1.9.	Substitutions.....	10
2.2.	Sampling errors.....	10
2.3.	Non-sampling errors.....	12
2.3.1.	Sampling frame and coverage errors.....	12
2.3.2.	Measurement and processing errors.....	12
2.3.3.	Non-response errors.....	13
2.4.	Mode of data collection.....	21
2.5.	Interview duration.....	22
3.	Comparability.....	22
3.1.	Basic concepts and definition.....	22
3.2.	Components of income.....	23
3.2.1.	Differences between the national definitions and standard EU-SILC definitions..	23
3.2.2.	The source or procedure used for the collection of income variables.....	24
3.2.3.	The form in which income variables at component level have been obtained.....	24
3.2.4.	The method used for obtaining income target variables in the required form.....	25
4.	Coherence.....	25
4.1.	Comparison of income target variables and number of persons who received income from each ‘income component’, with external source.....	26
4.2.	Comparison of other target variables with external source.....	26

1. Common cross-sectional European Union indicators

Table 1. Laeken indicators and other indicators

Overarching indicator	Value
<i>Primary Laeken indicators of social cohesion</i>	
At-risk-of-poverty rate after social transfers - total	19.1
At-risk-of-poverty rate after social transfers - men total	16.7
At-risk-of-poverty rate after social transfers - women total	21.3
At-risk-of-poverty rate after social transfers - 0-17 years	22.1
At-risk-of-poverty rate after social transfers – 65+ years	29.8
At-risk-of-poverty rate after social transfers – 18+ years	18.4
At-risk-of-poverty rate after social transfers - 18-64 years	15.6
At-risk-of-poverty rate after social transfers - men 65+ years	15.2
At-risk-of-poverty rate after social transfers - men 18+ years	15.1
At-risk-of-poverty rate after social transfers - men 18-64 years	15.1
At-risk-of-poverty rate after social transfers - women 65+ years	37.3
At-risk-of-poverty rate after social transfers - women 18+ years	21.1
At-risk-of-poverty rate after social transfers - women 18-64 years	16.1
At-risk-of-poverty rate after social transfers - employed	8.0
At-risk-of-poverty rate after social transfers – non-employed	32.0
At-risk-of-poverty rate after social transfers - unemployed	56.9
At-risk-of-poverty rate after social transfers - retired	29.8
At-risk-of-poverty rate after social transfers - other inactive	28.9
At-risk-of-poverty rate after social transfers - men, employed	7.7
At-risk-of-poverty rate after social transfers – men, non-employed	28.1
At-risk-of-poverty rate after social transfers - men, unemployed	62.7
At-risk-of-poverty rate after social transfers - men, retired	16.2
At-risk-of-poverty rate after social transfers - men, other inactive	28.7
At-risk-of-poverty rate after social transfers - women, employed	8.4
At-risk-of-poverty rate after social transfers – women, non-employed	34.5
At-risk-of-poverty rate after social transfers - women, unemployed	48.7
At-risk-of-poverty rate after social transfers - women, retired	36.1
At-risk-of-poverty rate after social transfers - women, other inactive	29.1
Median of the equivalised disposable household income	11312.0
At-risk-of-poverty threshold – single	6787.2
At-risk-of-poverty threshold - 2 adults, 2 children	14253.1
Inequality of income distribution S80/S20 income quintile share ratio	5.9
Aggregate replacement ratio – total	0.392
Aggregate replacement ratio – men total	0.375
Aggregate replacement ratio – women total	0.440
At risk-of-poverty rate anchored at a fixed moment in time (2005) - total	7.6
At risk-of-poverty rate anchored at a fixed moment in time (2005) – men total	7.4
At risk-of-poverty rate anchored at a fixed moment in time (2005) – women total	7.7
At risk-of-poverty rate anchored at a fixed moment in time (2005) – 0-17 years	10.9
At risk-of-poverty rate anchored at a fixed moment in time (2005) – 18-64 years	7.2
At risk-of-poverty rate anchored at a fixed moment in time (2005) – 65 + years	4.8

Overarching indicator	Value
At risk-of-poverty rate anchored at a fixed moment in time (2005) – men 18-64 years	7.3
At risk-of-poverty rate anchored at a fixed moment in time (2005) – men 65+ years	1.4
At risk-of-poverty rate anchored at a fixed moment in time (2005) – women 18-64 years	7.1
At risk-of-poverty rate anchored at a fixed moment in time (2005) – women 65+ years	6.5+
Relative median at-risk-of-poverty gap - total	25.7
Relative median at-risk-of-poverty gap - men total	28.2
Relative median at-risk-of-poverty gap - women total	23.5
Relative median at-risk-of-poverty gap – 0-17 years	29.6
Relative median at-risk-of-poverty gap - 18-64 years	28.9
Relative median at-risk-of-poverty gap - 65+ years	15.1
Relative median at-risk-of-poverty gap - 18+ years	23.4
Relative median at-risk-of-poverty gap - men, 18-64 years	29.6
Relative median at-risk-of-poverty gap - men, 65+ years	12.5
Relative median at-risk-of-poverty gap - men, 18+ years	27.5
Relative median at-risk-of-poverty gap - women, 18-64 years	27.6
Relative median at-risk-of-poverty gap - women, 65+ years	15.6
Relative median at-risk-of-poverty gap - women, 18+ years	21.3
Before social transfers except old-age and survivors' benefits	
At-risk-of-poverty rate before social transfers - total	25.5
At-risk-of-poverty rate before social transfers - men total	23.5
At-risk-of-poverty rate before social transfers - women total	27.3
At-risk-of-poverty rate before social transfers - 0-17 years	29.2
At-risk-of-poverty rate before social transfers - 18-64 years	22.4
At-risk-of-poverty rate before social transfers - 65+ years	33.5
At-risk-of-poverty rate before social transfers - 18+ years	24.6
At-risk-of-poverty rate before social transfers - men, 18-64 years	22.2
At-risk-of-poverty rate before social transfers - men, 65+ years	18.1
At-risk-of-poverty rate before social transfers - men, 18+ years	21.6
At-risk-of-poverty rate before social transfers - women, 18-64 years	22.6
At-risk-of-poverty rate before social transfers - women, 65+ years	41.4
At-risk-of-poverty rate before social transfers - women, 18+ years	27.0
Before social transfers including old-age and survivors' benefits	
At-risk-of-poverty rate before social transfers - total	38.3
At-risk-of-poverty rate before social transfers - men total	35.7
At-risk-of-poverty rate before social transfers - women total	40.6
At-risk-of-poverty rate before social transfers - 0-17 years	32.9
At-risk-of-poverty rate before social transfers - 18-64 years	28.9
At-risk-of-poverty rate before social transfers - 65+ years	84.4
At-risk-of-poverty rate before social transfers - 18+ years	39.7
At-risk-of-poverty rate before social transfers - men, 18-64 years	28.1
At-risk-of-poverty rate before social transfers - men, 65+ years	82.3
At-risk-of-poverty rate before social transfers - men, 18+ years	36.1
At-risk-of-poverty rate before social transfers - women, 18-64 years	29.6
At-risk-of-poverty rate before social transfers - women, 65+ years	85.5
At-risk-of-poverty rate before social transfers - women, 18+ years	42.7

2. Accuracy

2.1. Sample design

2.1.1 Type of sampling design

2007 operation was the third wave of EU-SILC in Lithuania. Households that were selected for the survey in 2005 divided into 4 rational groups. One of these groups was dropped out after 2005 operation and not included to the survey of 2006 according to the original integrated design. A new sub-sample of households was selected to the sample of year 2006. For new sample stratified sample design was used. Population register was used as a sampling frame. Simple random sample of persons was used in each stratum. The second group was dropped out after 2006 operation and not included to the survey of year 2007. A new sub-sample of households was selected to the sample of year 2007 according the same rules as selected a new sub-sample before.

2.1.2 Sampling units

The sampling units are private households.

2.1.3 Stratification criteria

While selecting the new part of the sample the country were grouped into 7 strata: 5 largest cities, other cities and rural area. Simple random sample of non-institutional persons aged 16 and over was selected from the Population Register in each stratum. Household which lives in the selected person's address was surveyed.

2.1.4 Sample size

The sample consisted of 5983 households. This number includes 4068 households, which responded to the survey in 2006 and where followed up during 2007 operation (3 rotational groups), newly selected rotational group – 1915 households.

2.1.5 Sample selection schemes

Within each of 7 strata simple random sample was used to select the person's address.

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. Fieldwork period was from the beginning of May 2007 till the end of August.

Table 2. Distribution of households by month of interview (HB050)

Month	Per cent
May	43.44
June	42.39
July	13.89
August	0.28

2.1.7. Renewal of sample: Rotational groups

In 2005 operation the sample was randomly divided into 4 equally sized rotational groups. In 2006 operation, first of four groups was dropped out after 2005 operation and not included to the survey of 2006 according to the original integrated design. Furthermore, for a split-off household the rotational group was set the same as one of original household. New rotational group was named as 1st. In 2007 operation, second of four groups was dropped out after 2006 operation and not included to the survey of 2007 according to the original integrated design. New rotational group was named as 2nd.

2.1.8. Weightings

The following sub-samples are consisted in the sample of the year 2007:

- s₁ – sample of the person in the households enumerated in 2007, persons participate for the first time (only 2nd rotational group);
- s₂ – sample of the person in the households enumerated in 2006, persons participate for the second time (only 1st rotational group);
- s₃ – Sample of the person in the households enumerated in 2005, persons participate for the third time (only 3^d rotational group).
- s₄ – Sample of the person in the households enumerated in 2005, persons participate for the third time (only 4th rotational group).

Base weights of year 2007 are calculated independently for each sub-sample.

1. Sub-sample is selected for the first time in the survey (s₁).

1.1. Design weights

Inclusion probability of a household in each stratum of new sub-sample is equal:

$$\pi_{hk} = \frac{n_h m_{hk}}{N_h},$$

here m_{hk} – the number of persons in k th household aged 16 and over in h th stratum in Population Register; n_h – the number of households in h th stratum; N_h – the number of persons aged 16 and older in h th stratum.

Sample design weights are:

$$DB080_h = d_{hk} = \frac{1}{\pi_{hk}}.$$

1.2. Adjustments for non-response at household level

To estimate household response probability logistic regression model are used. Response propensities are estimated for responding and non-responding households. Then for the each household k define variable:

$$R_k = \begin{cases} 1, & \text{if the household } k \text{ responds;} \\ 0, & \text{otherwise.} \end{cases}$$

Let define the response propensity of each household k :

$$p_k = \Pr(R_k = 1 | V_k)$$

where V_j – auxiliary variables (county group, urbanization status, age of person belonging to

address), R_k is defined above.

Then the modified design weights are defined:

$$DB080_k^{(N)} = d_{hk}^{(N)} = \frac{d_{hk}}{p_k}.$$

1.3. Adjustment to external sources (calibration)

Modified design weights are calibrated, seeking for the weights, which would remain as close as possible to sample design weights and allow obtaining some exact demographic estimates – auxiliary variables:

- number of persons aged 0 and older (including newborn children) by different strata;
- number of persons by different age groups;
- number of males by different age groups.

The product of calibration procedure is the calibrated household weight of sub-sample s_1 ; it is equals to the household base weight w_{1k}^1 for sub-sample s_1 of year 2007. Household base weight is assigned to each of its members: $w_{1i}^1 = w_{1k}^1$, $i \in k$.

The SAS macro program CLAN is used to calculate calibrated weights.

2. Sub-sample participated for the second time in the survey (s_2).

Sub-sample s_2 participated in the survey for the second time. To construct base weights of sub-sample s_2 of year 2007, we need to have base weights of this sub-sample of year 2006.

Base weights of year 2006 are calculated according steps which use in paragraph 1 (sub-sample is selected for the first time in the survey). Let denote base personal weight of sub-sample s_2 of year 2006 by w_{1i}^2 .

To determine base weight w_{2i}^2 of year 2007 from base weight w_{1i}^2 of year 2006, we use following step:

for the each person i , who are enumerated at year 2006 and still in-scope at year 2007 define variable:

$$R_i = \begin{cases} 1, & \text{if the person successfully enumerated at year 2007} \\ 0, & \text{otherwise.} \end{cases}$$

Using logit model, define the response propensity of each person i :

$$p_i = \Pr(R_i = 1 | V_i)$$

where V_i – auxiliary variables (like strata, total disposable household income, capacity to face unexpected financial expenses, lowest monthly income to make ends meet), R_i is defined above.

Then the personal base weight of sub-sample s_2 of year 2007 is defined:

$$w_{2i}^2 = 4 * \frac{w_{1i}^2}{p_i}.$$

Additionally assign the weights for new members of households of sub-sample s_2 :

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

3. Sub-sample participated for the third time in the survey (s_3).

Sub-sample s_3 (only 3^d rotational group) participated in the survey for the third time. To construct base weights of sub-sample s_3 of year 2007, we need to have base weights of this sub-sample of year 2005 and year 2006. Base weights of year 2005 are calculated according steps which use in paragraph 1 (sub-sample is selected for the first time in the survey). Let denote base personal weight of sub-sample s_3 of year 2005 by w_{1i}^3 .

To determine base weight w_{2i}^3 of year 2006 from base weight w_{1i}^3 of year 2005, we use following step: for the each person i , who are enumerated at year 2005 and still in-scope at year 2006 define variable:

$$R_{1i} = \begin{cases} 1, & \text{if the person successfully enumerated at year 2006} \\ 0, & \text{otherwise.} \end{cases}$$

Using logit model, define the response propensity of each person i :

$$p_{1i} = \Pr(R_{1i} = 1 | V_{1i})$$

where V_{1i} – auxiliary variables (like strata, total disposable household income), R_{1i} is defined above. Then the personal base weight of sub-sample s_3 of year 2006 is defined:

$$w_{2i}^3 = \frac{w_{1i}^3}{p_{1i}}.$$

Additionally assign the weights for newborns, for persons moving into sample households from outside the survey population and for persons moving into sample households from other non-sample households in the population according to the previous paragraph.

To determine base weight w_{3i}^3 of year 2007 from base weight w_{2i}^3 of year 2006, we denote for the each person i of sub-sample s_3 , who are enumerated at year 2006 and still in-scope at year 2007 variable:

$$R_{2i} = \begin{cases} 1, & \text{if the person successfully enumerated at year 2007} \\ 0, & \text{otherwise.} \end{cases}$$

Using logit model, define the response propensity of each person i :

$$p_{2i} = \Pr(R_{2i} = 1 | V_{2i})$$

where V_{2i} – auxiliary variables (like strata, dwelling type, tenure status, total disposable household income, lowest monthly income to make ends meet). Then the personal base weight of sub-sample s_3 of year 2007 is defined:

$$w_{3i}^3 = \frac{w_{2i}^3}{p_{2i}}.$$

Additionally assign the weights for new members who come in to the households in to year 2007 of sub-sample s_3 according to the previous paragraph.

We have persons of sub-sample s_3 who participated in year 2007, not participated in year 2006 and participated in year 2005. They are returnees.

Base personal weight for returnees of sub-sample s_3 of year 2005 defined by w_{1i}^3 . Denote for the each returnee i of sub-sample s_3 , who are enumerated at year 2005 and respond at year 2007 variable:

$$R_{3i} = \begin{cases} 1, & \text{if the person enumerated at year 2005 and 2007} \\ 0, & \text{otherwise.} \end{cases}$$

Using logit model, define the response propensity of each person i :

$$p_{3i} = \Pr(R_{3i} = 1 | V_{3i})$$

where V_{3i} – auxiliary variables (total disposable household income). Then the returnees' base weight of sub-sample s_3 of year 2007 is defined:

$$w_{3i}^3 = \frac{w_{1i}^3}{p_{3i}}.$$

Then final base weight of sub-sample s_3 of year 2007 is

$$w_{3i}^{3*} = \begin{cases} \frac{t - t_r}{t} \cdot w_{3i}^3, & \text{if } i \text{ is non-returnee;} \\ w_{3i}^3, & \text{otherwise.} \end{cases}$$

here t is the sum of base weights w_{3i}^3 of non-returnees, t_r is the sum of weights w_{3i}^3 of returnees.

4. Sub-sample participated for the third time in the survey (s_4).

Sub-sample s_4 (only 4th rotational group) participated in the survey for the third time. We do the same steps which described in paragraph 3 for this sub-sample. The final base weight of sub-sample s_4 of year 2007 is w_{3i}^{4*} .

5. Final cross-sectional weights (DB080, RB060, PB040, RL070)

Each sub-sample with base weights represents the whole population. The four sub-samples are combined. Averages of person base weights (w_{1i}^1 , w_{2i}^2 , w_{3i}^{3*} , w_{3i}^{4*}) are calculated for each household. As result we have the base weights for each household: w_h^1 , w_h^2 , w_h^3 and w_h^4 .

Then calculated modified base weights

$$w_h = \begin{cases} w_h^1 \cdot n_1 / \sum_{r=1}^4 n_r, & \text{if } h \in s_1; \\ w_h^2 \cdot n_2 / \sum_{r=1}^4 n_r, & \text{if } h \in s_2; \\ w_h^3 \cdot n_3 / \sum_{r=1}^4 n_r, & \text{if } h \in s_3; \\ w_h^4 \cdot n_4 / \sum_{r=1}^4 n_r, & \text{if } h \in s_4. \end{cases}$$

here n_r is the sample size of the sub-samples, $r=1, 2, 3, 4$.

Modified base weights are calibrated, seeking for the weights, which would remain as close as possible to sample design weights and allow obtaining some exact demographic estimates – auxiliary variables:

- number of persons aged 0 and older (including newborn children) by different strata;
- number of persons by different age groups;
- number of males by different age groups.

The product of calibration procedure is the calibrated household weight DB090 of year 2007.

Household cross-sectional weight is assigned to each of its members $RB050_i = DB090_h$, $i \in h$. $RB050$ are personal cross-sectional weights.

The cross-sectional weight $PB040$ for persons aged 16 or more is equal to the $RB050$ cross-sectional weight of aged 16 or more.

The children cross-sectional weight for child care *RL070* is equal to the *RB050* cross-sectional weight of group from 0 to 12 years old.

SAS macro program CLAN is used for calculation of the calibrated weights.

2.1.9. Substitutions

No substitution was used.

2.2. Sampling errors

The variance estimates were computed using SAS macro-programme *Clan*.

Table 3. Estimates, their standard error, confidence interval and design effect for the common cross-sectional indicators

Indicator	Value	Standard error	Confidence interval at 95%		CV(%)	Deff (calibration used)
At-risk-of-poverty rate after social transfers - total	19.1	0.7	17.7	20.6	3.79	1.269
At-risk-of-poverty rate after social transfers - men total	16.7	0.9	15.0	18.4	5.14	1.406
At-risk-of-poverty rate after social transfers - women total	21.3	0.8	19.7	22.8	3.69	1.146
At-risk-of-poverty rate after social transfers - 0-17 years	22.1	1.5	19.2	25.1	6.85	1.300
At-risk-of-poverty rate after social transfers - 65+ years	29.8	1.2	27.5	32.1	3.95	1.111
At-risk-of-poverty rate after social transfers - 18+ years	18.4	0.6	17.1	19.6	3.47	1.230
At-risk-of-poverty rate after social transfers - 18-64 years	15.6	0.7	14.2	17.0	4.56	1.220
At-risk-of-poverty rate after social transfers - men 65+ years	15.2	1.5	12.3	18.1	9.71	0.938
At-risk-of-poverty rate after social transfers - men 18+ years	15.1	0.7	13.6	16.5	4.94	1.169
At-risk-of-poverty rate after social transfers - men 18-64 years	15.1	0.8	13.5	16.7	5.44	1.182
At-risk-of-poverty rate after social transfers - women 65+ years	37.3	1.5	34.4	40.2	3.96	1.142
At-risk-of-poverty rate after social transfers - women 18+ years	21.1	0.7	19.6	22.5	3.43	1.229
At-risk-of-poverty rate after social transfers - women 18-64 years	16.1	0.8	14.5	17.6	5.02	1.232
At-risk-of-poverty rate before social transfers - total	25.5	0.8	23.9	27.2	3.33	1.396
At-risk-of-poverty rate before social transfers - men total	23.5	1.0	21.6	25.5	4.25	1.488
At-risk-of-poverty rate before social transfers - women total	27.3	0.9	25.5	29.0	3.24	1.212
At-risk-of-poverty rate before social transfers - 0-17 years	29.2	1.6	26.0	32.4	5.54	1.270
At-risk-of-poverty rate before social transfers - 18-64 years	22.4	0.9	20.7	24.2	3.99	1.463
At-risk-of-poverty rate before social transfers - 65+ years	33.5	1.2	31.1	35.9	3.63	1.043
At-risk-of-poverty rate before social transfers - 18+ years	24.6	0.8	23.0	26.1	3.19	1.446
At-risk-of-poverty rate before social transfers - men, 18-64 years	22.2	1.0	20.2	24.2	4.63	1.409
At-risk-of-poverty rate before social transfers - men, 65+ years	18.1	1.6	15.0	21.2	8.72	0.861
At-risk-of-poverty rate before social transfers - men, 18+ years	21.6	0.9	19.8	23.4	4.29	1.379
At-risk-of-poverty rate before social transfers - women, 18-64 years	22.6	1.0	20.7	24.4	4.22	1.356
At-risk-of-poverty rate before social transfers - women, 65+ years	41.4	1.5	38.5	44.3	3.60	1.069
At-risk-of-poverty rate before social transfers - women, 18+ years	27.0	0.8	25.4	28.6	3.07	1.327
Mean equivalised disposable income	13599	165	13275.4	13923.3	1.2	1.078

2.3. Non-sampling errors

2.3.1. Sampling frame and coverage errors

The sampling frame of EU-SILC 2007 was the Residents' Register. Residents' Register is updated regularly. However, not all movements of population within country are reflected, whereas not all population report about changing of address to the migration office. Consequently, the households, living in selected person's address, were surveyed. The sample was extracted 2 weeks before the fieldwork.

Percentage of non-contacted addresses by the reasons: address does not exist or is non-residential address or is unoccupied (DB120=23) out of total selected addresses – 2.2.

2.3.2. Measurement and processing errors

2.3.2.1. Measurement errors

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

The questionnaires for EU-SILC 2007 were developed according to the EU-SILC regulations and EU-SILC doc 65/04. The questionnaires were tested during the first wave of pilot survey conducted in 2004. Designing questionnaires for main operation errors and interviewers feedbacks from the pilot survey were considered. Also the experience from the different waves (2005 and 2006) of the survey was used to improve the questionnaire for the operation 2007.

The interviewers' training was organized in each territorial statistical office in the period between April 23 and May 5. Interviewers' manual presented instructions on filling in the questionnaires and detailed explanation for all income components, particularly benefits, were prepared. Special emphasis was placed on tracing rules and specifics of assigning household and person numbers in the longitudinal survey. Methodical explanations were combined with practical tests. Interviewers filled in questionnaires, our specialists checked and then mistakes were discussed. Fieldwork has started immediately after interviewers training.

Fieldwork was carried out by Households' interviewers who usually work for the other household surveys carried out by Statistics Lithuania with additionally hired temporary interviewers. Temporary staff was selected from current or former employees in regional statistical offices, or persons, formerly employed as enumerators in the Population Census or Agricultural Census. In total 163 interviewers were involved into 2007 year operation. One interviewer had an average 37 selected addresses.

2.3.2.2. Processing errors

Completed questionnaires were checked by specialists of the Living Standard Statistics Division of Statistics Lithuania. Necessary call-backs were made. Data were entered centrally by data entry operators in Statistics Lithuania. Blaise software was used for data entry. The computer program included the possible logical checks between questions and questionnaires, also a package of alerts (warning and error ones) related to ranges of admissible values and logical connections between questions. Coding controls were implemented in post-data-collection. After the data entry was finished the data were checked for consistency.

2.3.3. Non-response errors

2.3.3.1. Achieved sample size

Achieved sample size: 4975 households and 10913 persons aged 16 or older.

Table4. Accepted interviews

Rotational group	Number of households for which an interview is accepted for the database (DB135 = 1)	Number of persons aged 16 or older who are members of the households for which the interview is accepted for the database (DB135 = 1) and who completed personal interview (RB205 = 11 to 14)
Total	4975	10913
1	1558	3312
2	1576	3537
3	958	2148
4	883	1916

2.3.3.2. Unit non-response

The following rates are computed according to Eurostat definitions for the total sample.

Address contact rate:

$$Ra = \frac{5969}{6127 - 137} \approx 0.996$$

The proportion of completed household interviews accepted for the database:

$$Rh = \frac{4975}{5969} \approx 0.833$$

Household non-response rates:

$$NRh = (1 - (Ra * Rh)) * 100 = (1 - (0.996 * 0.833)) * 100 = 17.03$$

The proportion of completed personal interviews within the households accepted for the database:

$$Rp = \frac{10913}{10913} \approx 1$$

Individual non-response rate:

$$NRp = (1 - (Rp)) * 100 = (1 - 1) * 100 = 0$$

Overall individual non-response rate:

$$* NRp = (1 - (Ra * Rh * Rp)) * 100 = (1 - (0.996 * 0.833 * 1)) * 100 \approx 17.03.$$

The following rates are computed according to Eurostat definitions for the new replication.

Address contact rate:

$$Ra = \frac{1681}{1712 - 29} \approx 0.999$$

The proportion of completed household interviews accepted for the database:

$$Rh = \frac{1557}{1681} \approx 0.93$$

Household non-response rates:

$$NRh = (1 - (Ra * Rh)) * 100 = (1 - (0.999 * 0.93)) * 100 = 7.093$$

The proportion of completed personal interviews within the households accepted for the database:

$$Rp = \frac{3311}{3311} \approx 1$$

Individual non-response rate:

$$NRp = (1 - (Rp)) * 100 = (1 - 1) * 100 = 0$$

Overall individual non-response rate:

$$* NRp = (1 - (Ra * Rh * Rp)) * 100 = (1 - (0.999 * 0.93 * 1)) * 100 \approx 7.093 .$$

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)

Table 5. Distribution of households by ‘record of contact at address’

	Rotational group 1		Rotational group 2		Rotational group 3		Rotational group 4		Total	
	N	%	N	%	N	%	N	%	N	%
Total (DB120=11 to 23)	1713	100	2377	100	1044	100	993	100	6127	100
Address contacted (DB120=11)	1682	98.19	2303	96.87	1020	97.70	964	97.08	5969	97.42
Address non-contacted (DB120=21 to 23)	31	1.81	74	3.13	24	2.30	29	2.92	158	2.58
Total address non-contacted (DB120=21 to 23)	31	100	74	100	24	100	29	100	158	100
Address cannot be located (DB120=21)	2	6.45	18	24.32	1	4.17	0	0	21	13.29
Address unable to access (DB120=22)	0	0	0	0	0	0	0	0	0	0
Address does not exist or is non-residential address or is unoccupied or not principal residence (DB120=23)	29	93.55	56	75.68	23	95.83	29	100	137	86.71

Table 6. Distribution of address contacted by ‘household questionnaire result‘ and by ‘household interview acceptance‘

	Rotational group 1		Rotational group 2		Rotational group 3		Rotational group 4		Total	
	N	%	N	%	N	%	N	%	N	%
Total (DB130=11 to 24)	1682	100	2303	100	1020	100	964	100	5969	100
Household questionnaire completed (DB130=11)	1558	92.63	1576	68.43	958	93.92	883	91.60	4975	83.35
Interview not completed (DB130=21 to 24)	124	7.37	727	31.57	62	6.08	81	8.40	994	16.65
Total interview not completed (DB130=21 to 24)	124	100	727	100	62	100	81	100	994	100
Refusal to co-operate (DB130=21)	85	68.55	511	70.29	46	74.19	54	66.67	696	70.02
Entire household temporarily away for duration of fieldwork (DB130=22)	22	17.74	114	15.68	9	14.52	12	14.81	157	15.79
Household unable to respond (illness, incapacity, etc) (DB130=23)	1	0.81	5	0.69	0	0	1	1.23	7	0.70
Other (DB130=24)	16	12.9	97	13.34	7	11.29	14	17.29	134	13.49
Household questionnaire completed (DB135=1 to 2)	1558	100	1576	100	1020	100	964	100	4975	100
Interview accepted to database (DB135=1)	1558	100	1576	100	1020	100	964	100	4975	100
Interview rejected (DB135=2)	0	0	0	0	0	0	0	0	0	0

2.3.3.4. Item non-response

The following tables show the share of item non-response for income variables on household and individual level.

Table 7. Distribution of item non-response, household-level variables

Income variable	% of households having received an amount	% of households with missing values (before imputation)	% of households with partial* information (before imputation)
Total household gross income (HY010)	99.7	0.0	0.1
Total disposable household income (HY020)	99.7	0.0	0.1
Total disposable household income before social transfers except old-age and survivor's benefits (HY022)	98.2	0.0	0.7
Total disposable household income before social transfers including old-age and survivor's benefits (HY023)	79.5	0.0	4.0
<i>Gross income components at household level</i>			
Income from rental of a property or land (HY040G)	6.0	0.06	0.0
Family/child related allowances (HY050G)	13.2	0.0	0.0
Social exclusion not elsewhere classified (HY060G)	2.3	0.0	0.0
Housing allowances (HY070G)	4.4	0.0	0.0
Regular inter-household cash transfer received (HY080G)	7.2	0.0	0.0
Interest, dividends, etc. (HY090G)	4.9	0.0	0.0
Income received by people aged under 16 (HY110G)	0.2	0.0	0.0
Regular taxes on wealth (HY120G)	13.1	0.0.	0.0
Regular inter-household cash transfer paid (HY130G)	7.3	0.0	0.0

Table 8. Distribution of item non-response, person-level variables

Income variable	% of persons 16+ having received an amount	% of persons with missing values (before imputation)
<i>Gross income components at personal level</i>		
Employee cash or near cash income (PY010G)	48.5	0.3
Non-cash employee income (PY020G)	3.8	0.04
Company car (PY021G)	0.9	0.0
Contributions to individual private pension plans (PY035G)	1.4	0.0
Cash benefits or losses from self-employment (PY050G)	9.4	0.1
Value of goods produced for own consumption (PY070G)	11.0	0.0
Pension from individual private plans (PY080G)	0.0	0.0
Unemployment benefits (PY090G)	1.5	0.0
Old-age benefits (PY100G)	30.4	0.2
Survivor's benefits (PY110G)	1.9	0.0
Disability benefits (PY130G)	7.2	0.1
Education-related allowances (PY140G)	3.0	0.0

2.3.3.5. Total item non-response and number of observations in the sample at unit level of the common cross-sectional European Union indicators based on the cross-sectional component of EU-SILC and for equivalised disposable income

Item non-response:

- a. Number of persons with no information on most frequent activity status, when applicable (209);
- b. Number of persons with no information on household type, when applicable to indicator (17).

Non-response at individual level, i.e. an individual questionnaire is missing (0).

Non-response at household level, i.e. interview rejected for data base DB135=2 (0), address cannot be located DB120=21 (21) or address unable to access DB120=22 (0).

Table 9. Number of observations and total item non-response

	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 ¹	12777	0	NA	21
By age and gender¹				
men total	6012	0	NA	21
women total	6765	0	NA	21
0-17 years	2290	0	NA	21
18-24 years	1233	0	NA	21
25-49 years	3957	0	NA	21
50-64 years	2666	0	NA	21
65+ years	2631	0	NA	21
18+ years	10487	0	NA	21
18-64 years	7856	0	NA	21
0-64 years	10146	0	NA	21
men 18-24 years	642	0	NA	21
men 25-49 years	1989	0	NA	21
men 50-64 years	1204	0	NA	21
men 65+ years	1034	0	NA	21
men 18+ years	4778	0	NA	21
men 18-64 years	3744	0	NA	21
men 0-64 years	4978	0	NA	21
women 18-24 years	591	0	NA	21
women 25-49 years	2059	0	NA	21
women 50-64 years	1462	0	NA	21
women 65+ years	1597	0	NA	21
women 18+ years	5709	0	NA	21
women 18-64 years	4112	0	NA	21
women 0-64 years	5168	0	NA	21
By most frequent activity status² and gender				
Total 18+ years	10278	209	NA	21
employed	5380	209	NA	21
non-employed	4898	209	NA	21
unemployed	387	209	NA	21
retired	2956	209	NA	21
other inactive	1555	209	NA	21
total men 18+ year	4664	114	NA	21
men, employed	2685	114	NA	21

	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)
men, non-employed	1979	114	NA	21
men, unemployed	218	114	NA	21
men, retired	1081	114	NA	21
men, other inactive	680	114	NA	21
total women 18+ years	5614	95	NA	21
women, employed	2695	95	NA	21
women, non-employed	2919	95	NA	21
women, unemployed	169	95	NA	21
women, retired	1875	95	NA	21
women, other inactive	875	95	NA	21
<i>By household type³</i>				21
single, < 65 years	499	0	NA	21
single, 65+ years	603	0	NA	21
single, male	310	0	NA	21
single, female	792	0	NA	21
single, total	1102	0	NA	21
2 adults, no children, both < 65	1434	17	NA	21
2 adults, no children, at least one 65+	1738	17	NA	21
other households without children	1645	17	NA	21
single parent, at least one child	512	17	NA	21
2 adults, 1 child	1650	17	NA	21
2 adults, 2 children	1836	17	NA	21
2 adults, 3+ children	756	17	NA	21
other households with children	2087	17	NA	21
households without children	5919	17	NA	21
households with children	6841	17	NA	21
<i>By accommodation tenure status</i>				21
owner or rent-free	12563	0	NA	21
tenant	214	0	NA	21
Inequality of income distribution S80/S20 income quintile share ratio	12777	0	NA	21
Relative median at-risk-of-poverty gap				21

	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)
Total	2275	0	NA	21
<i>By age and gender</i>				21
men total	941	0	NA	21
women total	1334	0	NA	21
0-17 years	502	0	NA	21
18-64 years	1211	0	NA	21
65+ years	562	0	NA	21
18+ years	1773	0	NA	21
men, 18-64 years	565	0	NA	21
men, 65+ years	120	0	NA	21
men, 18+ years	685	0	NA	21
women, 18-64 years	646	0	NA	21
women, 65+ years	442	0	NA	21
women, 18+ years	1088	0	NA	21
Dispersion around the at-risk-of-poverty threshold				21
40%	12777	0	NA	21
50%	12777	0	NA	21
70%	12777	0	NA	21
At-risk-of-poverty rate before social transfers except old-age and survivors' benefits				
Total ¹	12777	0	NA	21
<i>By age and gender¹</i>				
men total	6012	0	NA	21
women total	6765	0	NA	21
0-17 years	2290	0	NA	21
18-64 years	7856	0	NA	21
65+ years	2631	0	NA	21
18+ years	10487	0	NA	21
men, 18-64 years	3744	0	NA	21
men, 65+ years	1034	0	NA	21
men, 18+ years	4778	0	NA	21
women, 18-64 years	4112	0	NA	21
women, 65+ years	1597	0	NA	21
women, 18+ years	5709	0	NA	21
At-risk-of-poverty rate before social transfers				

	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)
including old-age and survivors' benefits				
Total ¹	12777	0	NA	21
By age and gender¹				
men total	6012	0	NA	21
women total	6765	0	NA	21
0-17 years	2290	0	NA	21
18-64 years	7856	0	NA	21
65+ years	2631	0	NA	21
18+ years	10487	0	NA	21
men, 18-64 years	3744	0	NA	21
men, 65+ years	1034	0	NA	21
men, 18+ years	4778	0	NA	21
women, 18-64 years	4112	0	NA	21
women, 65+ years	1597	0	NA	21
women, 18+ years	5709	0	NA	21
Gini coefficient	12777	0	NA	21
Mean equivalised disposable income	12777	0	NA	21

¹ children born in 2006 are included;

² the information on activity status refers to the population of individuals aged 18+

³ all persons aged less than 18 are considered as dependent children, plus those economically inactive persons aged 18-24 living with at least one of their parents.

2.4. Mode of data collection

The method for data collection was paper assisted personal interview (PAPI). If necessary, telephone interviews were allowed. Proxy interview was allowed for persons temporarily away or in incapacity. To avoid non-response within household proxy interview as an exception was allowed when it was no possibility to make personal interview and another member of household could provide the information. Some data collected by proxy interview were amended by telephone, but method of data collection was not changed in the microdata.

According to Eurostat recommendations for dealing with the individual non-response problem full/partial imputation of missing personal interviews were used (14 cases). In case of full/partial imputation the variable RB250 (data status) = 14 “information completed from record imputation” and flag of variable RB260_F (type of interview) = -2.

Table 10. Distribution of household members aged 16 and over by ‘data status’ (RB250) and rotational group

HOUSEHOLD MEMBERS 16+ (RB245=1 to 3)

	Total	RB250=11	=12	=14	=21	=22	=23	=31	=32	=33
Total	10913	10885	0	28	0	0	0	0	0	0
%	100	99.74	0	0.26	0	0	0	0	0	0
Rotation 1	3312	3310	0	2	0	0	0	0	0	0
%	100	99.94	0	0.06	0	0	0	0	0	0
Rotation 2	3537	3519	0	18	0	0	0	0	0	0
%	100	99.49	0	0.51	0	0	0	0	0	0
Rotation 3	2148	2144	0	4	0	0	0	0	0	0
%	100	99.81	0	0.19	0	0	0	0	0	0
Rotation 4	1916	1912	0	4	0	0	0	0	0	0
%	100	99.79	0	0.21	0	0	0	0	0	0

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

	Total	RB260=1	RB260=2	RB260=3	RB260=4	RB260=5	Missing
Total*	10885	8270	0	329	76	2210	0
%	100	75.98	0	3.02	0.70	20.30	0
Rotation 1	3310	2416	0	125	15	754	0
%	100	72.99	0	3.78	0.45	22.78	0
Rotation 2	3519	2770	0	70	29	650	0
%	100	78.72	0	1.99	0.82	18.47	0
Rotation 3	2144	1619	0	70	19	436	0
%	100	75.51	0	3.26	0.89	20.34	0
Rotation 4	1912	1465	0	64	13	370	0
%	100	76.62	0	3.35	0.68	19.35	0

*Full imputed not included

2.5. Interview duration

Mean duration of household interview (HB100) - 22 minutes.

Mean duration of personal interview (PB120) - 18 minutes.

Mean interview duration per household – 62 minutes.

3. Comparability

3.1. Basic concepts and definitions

The reference population

No difference to the common definition. The target population of EU-SILC is all persons living in private households within the national territory of Lithuania at the time of data collection. Collective households and institutions are excluded from the target population.

The private household definition

No difference to the common definition. The private household is defined as a person living alone or a group of people, who live together in the same private dwelling and share expenditures, including the joint provision of the essentials of living.

The household membership

No difference to the common definition.

The income reference period used

No difference to the common definition. The income reference period was a fixed twelve-month period, namely the last calendar year. In the 2007 operation income data were collected for the reference year 2006.

The period for taxes on income and social insurance contributions

No difference to the common definition. Taxes on income and social insurance contributions, as well as tax repayments and receipts refer to the income reference period (year 2006).

The reference period for taxes on wealth

No difference to the common definition. Taxes on wealth paid during the income reference period (year 2006) were recorded.

The lag between the income reference period and current variables

The lag between the end of the income reference period and current variables ranges from 4 to 8 months.

The total duration of the data collection of the sample

The fieldwork period started on 2nd of May 2007 and ended on the 30th of August (only 14 households were interviewed during August). 85.8% of households were interviewed during the first 2 months and only 14.2% were interviewed in July and August.

Basic information on activity status during the income reference period

This information was collected with the questionnaire by an activity calendar covering each month of the income reference period.

3.2. Components of income

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

Imputed rent

For estimating of Imputed rent we used two step model.

1 step. Stratification method, using data from Housing Rental Price Survey was applied.

2 step. Log-linear regression method was used to estimate the rest of the missing values.

Cash or near cash employee income

To calculate Sickness benefits (PY120) data from the State Social Insurance Fund Board and the State Tax Inspectorate were used. The algorithm based on country health insurance system was used for missing values.

No-cash employee income

All components of this variable were collected. The values related to company car were recorded under variable PY021 and were added to the calculation variables HY010, HY020, HY022 and HY023.

Cash benefits or losses from self-employment

The self-employment income was collected as the amount of money drawn out of the business for household, personal use. Income from agriculture, included in this variable, was calculated as difference of total revenue from agricultural activity and total expenditure on it.

Value of goods produced for own-consumption

Variable was collected and recorded to microdata file, but was not added to the calculation variables HY010, HY020, HY022 and HY023.

The quantities of products, used for own consumption, were collected during interview. The value of goods produced for own consumption was estimated by multiplying quantity by market prices of goods from the Household Budget Survey deducting expenses incurred in the production.

Gross monthly earnings for employees

Variable was not collected because EU-SILC is not used to calculate gender pay gap.

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

Where applicable the EU-SILC income target variables were split into sub-components. The sub-components were defined according to the Lithuanian regulations and benefit system. All data related to income variables were collected from interviews.

Administrative data were used for making the survey income data more accurate or for supplementing them. The State Social Insurance Fund Board data and the State Tax Inspectorate under the Ministry of Finance of the Republic of Lithuania data have been linked to sample data and used for checking cash or near-cash employee income (PY010, PY120), maternity and maternity/paternity allowances (component of HY050), dividends from capital investments (component of HY090), social insurance contributions and taxes on income (components of HY140).

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

Employee cash and near-cash income (PY010), self-employment income (PY050), unemployment benefits (PY090), family/children related allowances (HY050), interest, dividends, profit from capital investments (HY090), income received by people aged under 16

(HY110) were collected in net and/or gross. The remaining variables were collected only in gross.

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

The gross-net/net-gross conversion was used for either gross or net was collected. Conversion algorithms were created on the bases of country tax system. All income variables that are subjected to taxation and/or social insurance contribution were recorded gross and net in to the microdata files. Other income variables were recorded only gross.

4. Coherence

This section will compare the EU-SILC data to Household Budget Survey (HBS), wage statistics and administrative data.

The HBS is continuous survey. The survey conducted in line with the current methodology has been carried out since 1996. The HBS uses two data collection methods combined into one: the interview conducted by an interviewer and self-registration of particular household indicators. Social and economic information on household members, their living conditions and income are collected during the interview. HBS was the source of Laeken indicators until started EU-SILC survey.

4.1. Comparison of income target variables and number of persons who received income from each 'income component', with external source

There are differences between EU-SILC and HBS income components definitions. Only comparable income components are presented in Table 12.

Table 12. Comparison of income target variables and number of persons/households who received income components

Income component	EU-SILC 2007	HBS 2006	Other sources*
	Annual number of people, thousand		Average annual number of people, thousand
Cash or near cash employee income (PY010N)	1493.3	1339.4	1263.7
Old-age benefits (PY100)	666.0	676.1	599.1
Survivors benefits (PY110)	54.1	38.0	...
	Annual number of households, thousand		
Housing allowances (HY070)	70.8	46.1	...

* Wage statistics in the case of PY010 and administrative source in the case of PY100

The number of people receiving employee income is higher in SILC than in the HBS and wage statistics. In HBS, the yearly income figures are derived from monthly data. People who were employed, but did not receive income during the survey month (being on vacation, started job and so on) were not included in this category. In case of wage statistics, this figure is lower whereas the illegal work has not been taken into account.

The estimate of number of persons receiving old-age benefits is the same in SILC and in HBS. This is due to old-age pensions from foreign countries and disability benefits paid after the standard retirement age being included in SILC variable that have not been taken into account in the case of administrative source. The differences between SILC and HBS are not substantial.

The estimate of number of people receiving survivor's benefits is higher in SILC than in HBS. The reason of the difference is in assignment of survivor benefits value for eligible person. In SILC values of benefit are recorded to each person 16 years and older who receive this benefits. Whereas in HBS, values of benefit received by persons younger than 18 years old are recorded to the older persons in that household.

The number of households receiving housing allowances is lower in the HBS. This difference is related to the survey design of HBS and the seasonal aspect of housing allowances. As was noted above, the yearly income figures are derived from monthly data in HBS. The compensations to cover expenditure of the heating of dwelling are the most part of housing allowances and are paid in winter time. So, the number of households receiving them is lower in HBS data.

4.2. Comparison of other target variables with external source

Table 13. Distribution of households by type of dwelling

Dwelling type	EU-SILC 2007	HBS 2007
	%	%
Detached house	32,4	31,5
Semi-detached or terraced house	9,0	10,3
Apartment or flat	58,6	57,9
Some other kind of accommodation	0,0	0,3
Total	100	100

Table 14. Distribution of households by amenities in the dwellings

Amenities in the dwellings	EU-SILC 2007	HBS 2007
	%	%
Bath or shower	78,4	76,1
Indoor flushing toilet	77,2	74,5

The estimates of the number of household by household type and amenities in the dwellings are almost the same in EU-SILC and HBS.

Finally, in Table 15 there are reported data for the distribution of population by self-defined economic status. This variable is not absolutely the same in the SILC and HBS. The main activity status is self-defined in EU-SILC.

Table 15. Distribution of population aged 16 and over by self-defined activity status

Activity status	EU-SILC 2007	HBS 2007
	%	%
At work	61,8	57,0
Unemployed	5,2	4,1
Pupil, student	10,8	9,3
In retirement	15,8	22,2
Permanently disabled	4,7	4,0
Other inactive person	1,7	3,4
Total	100	100

Regina Deveikyte
Head, Living Standard Statistics Division
Statistics Lithuania
Phone +370 5 2364 919
e-mail: regina.deveikyte@stat.gov.lt