

**CENTRAL STATISTICAL OFFICE OF POLAND**

**INTERMEDIATE QUALITY REPORT**

**ACTION ENTITLED:**  
***EU-SILC 2011***

Warsaw, February 2013

	Page
<b>CONTENTS</b>	
<b>PREFACE</b>	<b>3</b>
<b>1. COMMON CROSS-SECTIONAL EUROPEAN UNION INDICATORS.....</b>	<b>4</b>
1.1. Common cross-sectional EU indicators based on the cross-sectional component of EU-SILC .....	4
<b>2. ACCURACY.....</b>	<b>6</b>
2.1. Sample design .....	6
2.1.1. <i>Type of sampling design</i> .....	6
2.1.2. <i>Sampling units</i> .....	6
2.1.3. <i>Stratification and substratification criteria</i> .....	6
2.1.4. <i>Sample size and allocation criteria</i> .....	6
2.1.5. <i>Sample selection schemes</i> .....	7
2.1.7. <i>Renewal of sample: rotational groups</i> .....	7
2.1.8. <i>Weightings</i> .....	7
2.1.9. <i>Substitutions</i> .....	10
2.2. Sampling errors .....	10
2.2.1. <i>Standard error and effective sample size</i> .....	10
2.3. Non-sampling errors.....	14
2.3.1. <i>Sampling frame and coverage errors</i> .....	14
2.3.2. <i>Measurement and processing errors</i> .....	14
2.3.3. <i>Non-response errors</i> .....	16
2.4. Mode of data collection.....	25
2.5. Interview duration .....	26
<b>3. COMPARABILITY .....</b>	<b>27</b>
3.1. Basic concepts and definitions .....	27
3.2. Components of income.....	29
3.2.1. <i>Differences between the national definitions and standard EU-SILC definitions,             and an assessment</i> .....	29
3.2.2. <i>The source or procedure used for the collection of income variables</i> .....	31
3.2.3. <i>The form in which income variables at component level have been obtained</i> .....	31
3.2.4. <i>The method used for obtaining income target variables in the required form</i> .....	31
<b>4. COHERENCE .....</b>	<b>32</b>
4.1. Comparison of EU-SILC and HBS results.....	32
4.2. Comparison of Laeken Indicators based on EU-SILC 2010 and EU-SILC 2011 .....	35
4.3. Comparison of 2010 results of SNA and EU-SILC 2011 (data for 2010) for Poland.....	35

## **PREFACE**

This quality report is the Intermediate Quality Report of EU-SILC 2011 in Poland. It follows the structure outlined in the Commission Regulation No. 1177/2003.

This report consists of four chapters.

The first chapter describes the common cross-sectional indicators.

The second chapter deals with accuracy i.e. discusses all the factors that affect the precision of estimations and results.

The third chapter reports on comparability and indicates all the differences between the standard EU definitions and those applied in the polish survey.

The fourth and last chapter, reporting on coherence, presents the comparison of the EU-SILC 2011 data with external sources.

## 1. COMMON CROSS-SECTIONAL EUROPEAN UNION INDICATORS

### 1.1. Common cross-sectional EU indicators based on the cross-sectional component of EU-SILC 2011

Indicator	Value
At-risk-of-poverty rate after social transfers - total	<b>17.7</b>
At-risk-of-poverty rate after social transfers - men total	<b>17.8</b>
At-risk-of-poverty rate after social transfers - women total	<b>17.6</b>
At-risk-of-poverty rate after social transfers - 0-17 years	<b>22.0</b>
At-risk-of-poverty rate after social transfers - 18-64 years	<b>17.1</b>
At-risk-of-poverty rate after social transfers - men 18-64 years	<b>17.6</b>
At-risk-of-poverty rate after social transfers - women 18-64 years	<b>16.5</b>
At-risk-of-poverty rate after social transfers - 65+ years	<b>14.7</b>
At-risk-of-poverty rate after social transfers - men 65+ years	<b>10.2</b>
At-risk-of-poverty rate after social transfers - women 65+ years	<b>17.4</b>
At-risk-of-poverty rate after social transfers - employed	<b>11.1</b>
At-risk-of-poverty rate after social transfers - men employed	<b>12.3</b>
At-risk-of-poverty rate after social transfers - women employed	<b>9.7</b>
At-risk-of-poverty rate after social transfers - unemployed	<b>43.7</b>
At-risk-of-poverty rate after social transfers - men unemployed	<b>49.7</b>
At-risk-of-poverty rate after social transfers - women unemployed	<b>38.2</b>
At-risk-of-poverty rate after social transfers - retired	<b>13.2</b>
At-risk-of-poverty rate after social transfers - men retired	<b>8.6</b>
At-risk-of-poverty rate after social transfers - women retired	<b>16.0</b>
At-risk-of-poverty rate after social transfers - other inactive	<b>26.9</b>
At-risk-of-poverty rate after social transfers - men other inactive	<b>28.6</b>
At-risk-of-poverty rate after social transfers - women other inactive	<b>25.9</b>
At-risk-of-poverty rate after social transfers - household without children	<b>14.8</b>
At-risk-of-poverty rate after social transfers - single total	<b>25.5</b>
At-risk-of-poverty rate after social transfers - single male	<b>31.6</b>
At-risk-of-poverty rate after social transfers - single female	<b>22.7</b>
At-risk-of-poverty rate after social transfers - single <65 years	<b>29.1</b>
At-risk-of-poverty rate after social transfers - single 65+ years	<b>22.0</b>
At-risk-of-poverty rate after social transfers - 2 adults, no children, both <65	<b>13.5</b>
At-risk-of-poverty rate after social transfers - 2 adults, no children, at least one 65+	<b>11.7</b>
At-risk-of-poverty rate after social transfers - other households without children	<b>10.7</b>
At-risk-of-poverty rate after social transfers - household with children	<b>19.6</b>
At-risk-of-poverty rate after social transfers - single parent, at least one child	<b>29.8</b>
At-risk-of-poverty rate after social transfers - 2 adults, 1 child	<b>11.3</b>
At-risk-of-poverty rate after social transfers - 2 adults, 2 children	<b>19.7</b>
At-risk-of-poverty rate after social transfers - 2 adults, 3+ children	<b>34.6</b>
At-risk-of-poverty rate after social transfers - other households with children	<b>19.3</b>
At-risk-of-poverty rate after social transfers - owner or rent free	<b>17.4</b>
At-risk-of-poverty rate after social transfers - tenant	<b>22.6</b>
At-risk-of-poverty threshold - single	<b>12044.9</b>
At-risk-of-poverty threshold - 2 adults, 2 children	<b>25294.2</b>

Relative median at-risk-of-poverty gap - total	<b>21.4</b>
Relative median at-risk-of-poverty gap - men total	<b>22.8</b>
Relative median at-risk-of-poverty gap - women total	<b>20.3</b>
Relative median at-risk-of-poverty gap - 0-17 years	<b>22.6</b>
Relative median at-risk-of-poverty gap - 18-64 years	<b>22.4</b>
Relative median at-risk-of-poverty gap - men, 18-64 years	<b>23.7</b>
Relative median at-risk-of-poverty gap - women, 18-64 years	<b>21.1</b>
Relative median at-risk-of-poverty gap - 65+ years	<b>15.7</b>
Relative median at-risk-of-poverty gap - men 65+ years	<b>10.9</b>
Relative median at-risk-of-poverty gap - women, 65+ years	<b>16.7</b>
Inequality of income distribution S80/S20 income quintile share ratio	<b>4.95</b>
In work at-risk-of-poverty rate - total	<b>11.1</b>
In work at-risk-of-poverty rate - men total	<b>12.3</b>
In work at-risk-of-poverty rate - women total	<b>9.7</b>
Relative median income ratio people aged 65+/0-64	<b>0.92</b>
Relative median income ratio people aged 65+/0-64 - men	<b>1.01</b>
Relative median income ratio people aged 65+/0-64 - women	<b>0.87</b>
Aggregate replacement ratio pensions 65-74/earnings 50-59	<b>0.55</b>
Aggregate replacement ratio pensions 65-74/earnings 50-59 - men	<b>0.62</b>
Aggregate replacement ratio pensions 65-74/earnings 50-59 - women	<b>0.53</b>
Dispersion around the at-risk-of-poverty threshold - 40% median	<b>5.7</b>
Dispersion around the at-risk-of-poverty threshold - 50% median	<b>10.5</b>
Dispersion around the at-risk-of-poverty threshold - 70% median	<b>25.7</b>
At-risk-of-poverty rate anchored at a moment in time (2005)	<b>6.1</b>
<b>Before social transfers except old-age and survivors' benefits</b>	
At-risk-of-poverty rate before social transfers - total	<b>24.1</b>
At-risk-of-poverty rate before social transfers - men total	<b>24.4</b>
At-risk-of-poverty rate before social transfers - women total	<b>23.9</b>
At-risk-of-poverty rate before social transfers - 0-17 years	<b>30.1</b>
At-risk-of-poverty rate before social transfers - 18-64 years	<b>23.8</b>
At-risk-of-poverty rate before social transfers - men, 18-64 years	<b>24.5</b>
At-risk-of-poverty rate before social transfers - women, 18-64 years	<b>23.2</b>
At-risk-of-poverty rate before social transfers - 65+ years	<b>17.3</b>
At-risk-of-poverty rate before social transfers - men 65+ years	<b>12.2</b>
At-risk-of-poverty rate before social transfers - women, 65+ years	<b>20.4</b>
<b>Before social transfers including old-age and survivors' benefits</b>	
At-risk-of-poverty rate before social transfers - total	<b>43.4</b>
At-risk-of-poverty rate before social transfers - men total	<b>41.4</b>
At-risk-of-poverty rate before social transfers - women total	<b>45.2</b>
At-risk-of-poverty rate before social transfers - 0-17 years	<b>36.7</b>
At-risk-of-poverty rate before social transfers - 18-64 years	<b>37.0</b>
At-risk-of-poverty rate before social transfers - men, 18-64 years	<b>36.2</b>
At-risk-of-poverty rate before social transfers - women, 18-64 years	<b>37.8</b>
At-risk-of-poverty rate before social transfers - 65+ years	<b>84.0</b>
At-risk-of-poverty rate before social transfers - men 65+ years	<b>83.7</b>
At-risk-of-poverty rate before social transfers - women, 65+ years	<b>84.2</b>
Mean equivalised disposable income	<b>23221.5</b>
Gini coefficient	<b>31.1</b>

## **2. ACCURACY**

### **2.1. Sample design**

#### ***2.1.1. Type of sampling design***

The two-stage sampling scheme with differentiated selection probabilities at the first stage was used. Prior to selection, sampling units were stratified.

#### ***2.1.2. Sampling units***

The first-stage sampling units (primary sampling units - PSUs) were enumeration census areas, while at the second stage dwellings were selected. All the households from the selected dwellings are supposed to enter the survey.

#### ***2.1.3. Stratification and substratification criteria***

The strata were the voivodships (NUTS2) and within voivodships, primary sampling units were classified by class of locality. In urban areas, census areas were grouped by size of town, but in the five largest cities, districts were treated as strata. In rural areas, strata were represented by rural gminas (NUTS5) of a subregion (NUTS3) or of a few neighbouring powiats (NUTS4). Altogether, 211 strata were distinguished.

#### ***2.1.4. Sample size and allocation criteria***

It was decided that the sample should include about 24 000 dwellings in the first year of the survey. Proportional allocation of dwellings to particular strata was applied. The number of dwellings selected from a particular stratum was in proportion to the number of dwellings in the stratum. Furthermore, the number of first-stage units selected from the strata was obtained by dividing the number of dwellings in the sample by the number of dwellings determined for a given class of locality to be selected from the first-stage unit. In towns with over 100 000 population, 3 dwellings per PSU were selected, in towns with 20-100 thousand population – 4 dwellings per PSU, in towns with less than 20 000 population – 5 dwellings per PSU, respectively. In rural areas 6 dwellings were selected from each PSU. Altogether, 5912 census areas and 24044 dwellings were selected for the sample in the first year of the survey. The subsample 5, selected for the survey in 2006 to replace the subsample 1, consisted of 1476 census areas and 6002 dwellings. Then, in 2007, the subsample 6 replaced the subsample 2 and consisted of 1478 census areas and 6008 dwellings. For the 2008 survey, the subsample 3 was replaced by the subsample 7. This new subsample consisted of 1479 census areas and 6016 dwellings. For the 2009 survey, the subsample 4 was replaced by the subsample 8, which consisted of 1479 census areas and 6017 dwellings. For the 2010 survey, the subsample 5 was replaced by the subsample 9 (coded as 1) which consisted of 1477 census areas and 6003 dwellings. For the 2011 survey, the subsample 6 was replaced by the subsample 2 (the old subsample 2 was not used in survey in 2011 year), which consisted of 1480 census areas and 6001 dwellings.

In official cross-sectional and longitudinal data for EU-SILC 2011 operation the following coding was used: variable DB075 (rotation group) equals 1 for subsample 1, 2 for subsample 6, 3 for subsample 7, and 4 for subsample 8.

### **2.1.5. Sample selection schemes**

Census areas were selected according to the Hartley-Rao scheme. Prior to selection, census areas were put in random order for each stratum separately and then the determined number of PSUs was selected with probabilities proportionate to the number of dwellings. Then, in each of the census areas belonging to the PSU sample, dwellings were selected using the simple random selection procedure.

### **2.1.7. Renewal of sample: rotational groups**

The selected sample of first-stage units was divided into four subsamples, equal in size. Starting from 2006, one of the subsamples is eliminated and replaced with a new one, selected independently as described above. For the 2006 survey, the subsample 5 was selected as a replacement of the subsample 1. Then, for the 2007 survey, the subsample 6 was selected which replaced the subsample 2. For the 2008 survey, the new subsample 7 replaced subsample 3. For the year 2009, the new subsample 8 replaced the subsample 4. For the year 2010, the new subsample 1 replaced the subsample 5. For the year 2011, the new subsample 2 replaced the subsample 6.

### **2.1.8. Weightings**

#### Design factor

Design factor – DB080 is equal to the dwelling sampling fraction reciprocal in the h-th stratum i.e.

$$f_h = \frac{n_h * m'_h}{M_h},$$

$$DB080 = \frac{1}{f_h}$$

where:

$n_h$  - number of PSUs selected from the h-th stratum,

$m'_h$  - number of dwellings selected from a PSU in the h-th stratum,

$M_h$  – number of dwellings in the h-th stratum.

#### Non-response adjustments

DB080 weights were then adjusted with the use of household non-response rates estimated for each class of locality separately:

<b>Code of class of locality (p)</b>	<b>Class of locality</b>	<b>Completeness rate (<math>cr_p = R_{ap} * Rh_p</math>)</b>
	Poland	0.649
1	Warsaw	0.411
2	Towns 500 000 – 1 000 000 inhabitants	0.473
3	Towns 100 000 – 500 000 inhabitants	0.625
4	Towns 20 000 – 100 000 inhabitants	0.669
5	Towns less than 20 000 inhabitants	0.684
6	Rural areas	0.747

The adjusted weights were calculated according to the formula:

$$DB080_p^{corrected} = \frac{DB080_p}{Ra_p * Rh_p},$$

Weights DB080 and DB080<sup>corrected</sup> were calculated for the new subsample 2. The next step consisted of calculating the weights DB090 and RB050 for the households of the subsample 2 with the use of the integrated calibration method. For the subsample 8 (surveyed for the third time), the subsample 1 (surveyed for the second time) and the subsample 7 (surveyed for the fourth time), the base weights were determined by the correction of the base weights from the previous year.

For the subsample 1, the following method was used:

The base weight of 2010 is equal to RB050 multiplied by 4. This weight was then adjusted by non-response and households' and individuals' falling out of the population surveyed. The calculations were made on the subsamples of the so called sample persons i.e. those who were in the surveyed sample at the age of 14 and over in 2010 and who should be surveyed in 2011. The modifying factor was determined according to the class of locality and took the form:

$$\frac{R(1)_p - M}{R(2)_p}$$

where:

$R(t)_p$  – estimated number of respondents belonging to the sample person group in the p-th class of locality in the subsample surveyed for the t-th time,

$M$  – estimated number of sample persons who belonged to the surveyed population in the first year and in the next year were out of the survey scope.

The base weights of 2010 were used for the calculation of numerator and denominator. The above expression is the reciprocal of the empirical estimate of probability that a given person will be interviewed again in the second year of the survey. In the second stage of the base weight calculation for the second year of the survey children of "sample persons" received the weights of mothers and "co-residents" i.e. additional persons included in the household

surveyed were ascribed zero weights. Then the respondents' base weights were averaged and all the members of a given household were ascribed such a mean weight. Then for the weights thus obtained the trimming of extreme weights was applied.

For the subsamples 7 and 8 (surveyed for the fourth and third time respectively) an algorithm based on method described for the subsample 1 was used. Additionally, the occurrence of re-entries (i.e. persons who were surveyed in 2009, not surveyed in 2010, and again surveyed in the year 2011) was taken into account. The base weights for such persons were computed by correction of base weights from year 2009 on data for years 2009 and 2011 (without information from year 2010). Inclusion of re-entries to the subsamples surveyed in 2011 caused the necessity of additional correction of the base weights for persons surveyed in the three successive years. Coefficients of these corrections were computed separately according to classes of locality as ratios: weighted number of respondents surveyed in all three years to the weighted number of respondents in the last survey year (i.e. with re-entries); weight used in these calculations was the weight RB050 for year 2009. Computed coefficients are shown in the following table:

Class of locality	Correction for subsample 7	Correction for subsample 8
1	0.981	0.972
2	0.998	0.987
3	0.984	0.994
4	0.993	0.979
5	0.994	0.973
6	0.996	0.996

The last stage of the base weight calculation for the fourth year of the survey consisted of receiving weights of mothers by children of "sample persons" and zero weights by "co-residents" i.e. additional persons included in the households. Then the respondents' base weights were averaged and all the members of a given household were ascribed such a mean weight. Then for the weights thus obtained the trimming of extreme weights was applied.

The last stage of calculations consisted of combining the four independent subsamples, applying the integrated calibration (for the sample 2 repeatedly) and trimming. As a result, DB090 and RB050 weights are obtained for households and individuals from the samples 7, 8, 1 and 2.

#### Adjustments to external data

Using the integrated calibration method (in hyperbolic sinus version) weights were calculated for individuals and for households simultaneously. To do this, the information about households was used (4 size categories: 1-person, 2-person, 3-person and 4- and more person households) and number of persons by age and gender (15 age groups: under 16, 16-19 years, then eleven 5-

year groups, 75 years and over). This information at the level of NUTS2, additionally classified by urban/rural areas, was derived from the 2002 Census and current demographic estimates.

#### Final cross-sectional weight

In EU-SILC 2011 the following cross-sectional weights were calculated:

DB090 – weight for households,

RB050 – weight for all household members,

$$RB050_{ij} = DB090_i$$

where:

i – household number,

j – person number in the i-th household.

PB040 – weight for respondents at the age of 16 and over who had individual interview. This weight equals the weight RB050.

RL070 – weight for children at the age of 0–12 years. It is obtained by the adjustment of RB050 weight in 26 groups, i.e. 13 years of birth and gender.

#### **2.1.9. Substitutions**

No substitution was applied if the household did not enter the survey.

### **2.2. Sampling errors**

#### **2.2.1. Standard error and effective sample size**

Estimation of standard errors was based on a resampling approach. We used a bootstrap method which resamples 500 times from each stratum  $n_h - 1$  PSU's (primary sampling units) with replacement (method of McCarthy and Snowden (1985)), where  $n_h$  denotes the sample size of PSU in the  $h$ -th stratum. After resampling the original weights were properly rescaled and bootstrap variance estimate of the corresponding indicator was obtained by the usual Monte Carlo approximation based on the independent bootstrap replicates. Computations were carried out using SAS software. Additionally, we implemented the linearization method of variance estimation for the main poverty indicators, and the results of comparisons with those obtained by the bootstrap method showed they were very similar.

<b>Indicator</b>	<b>Value</b>	<b>Standard error</b>	<b>Achieved sample size</b>	<b>Design effect</b>	<b>Effective sample size</b>
At-risk-of-poverty rate after social transfers - total	17.67	0.49	36720	1.51	24318
At-risk-of-poverty rate after social transfers - men total	17.80	0.53	17587	1.36	12979
At-risk-of-poverty rate after social transfers - women total	17.56	0.50	19133	1.21	15878
At-risk-of-poverty rate after social transfers - 0-17 years	22.02	0.82	7353	1.00	7353
At-risk-of-poverty rate after social transfers - 18-64 years	17.06	0.50	23680	1.53	15497
At-risk-of-poverty rate after social transfers - men 18-64 years	17.59	0.54	11620	1.24	9394
At-risk-of-poverty rate after social transfers - women 18-64 years	16.54	0.54	12060	1.15	10460
At-risk-of-poverty rate after social transfers - 65+ years	14.68	0.60	5687	1.13	5015
At-risk-of-poverty rate after social transfers - men 65+ years	10.21	0.84	2226	1.11	2011
At-risk-of-poverty rate after social transfers - women 65+ years	17.36	0.73	3461	0.98	3550
At-risk-of-poverty rate after social transfers - employed	11.08	0.40	13044	1.30	10003
At-risk-of-poverty rate after social transfers - men employed	12.30	0.51	7040	1.19	5906
At-risk-of-poverty rate after social transfers - women employed	9.68	0.45	6004	0.99	6046
At-risk-of-poverty rate after social transfers - unemployed	43.69	1.79	1542	1.19	1300
At-risk-of-poverty rate after social transfers - men unemployed	49.67	2.37	723	1.02	712
At-risk-of-poverty rate after social transfers - women unemployed	38.21	2.11	819	1.11	741
At-risk-of-poverty rate after social transfers - retired	13.16	0.54	6928	1.13	6147
At-risk-of-poverty rate after social transfers - men retired	8.60	0.65	2688	1.05	2567
At-risk-of-poverty rate after social transfers - women retired	15.97	0.67	4240	0.95	4482
At-risk-of-poverty rate after social transfers - other inactive	26.85	0.81	6299	1.04	6086
At-risk-of-poverty rate after social transfers - men other inactive	28.58	1.21	2326	1.03	2252
At-risk-of-poverty rate after social transfers - women other inactive	25.88	0.93	3973	0.99	4021
At-risk-of-poverty rate after social transfers - household without children	14.76	0.52	14763	1.53	9649
At-risk-of-poverty rate after social transfers - single total	25.53	0.95	2533	0.98	2598
At-risk-of-poverty rate after social transfers - single male	31.57	1.90	732	0.90	811
At-risk-of-poverty rate after social transfers - single female	22.72	1.06	1801	0.98	1845
At-risk-of-poverty rate after social transfers - single <65 years	29.11	1.54	1179	1.02	1154
At-risk-of-poverty rate after social transfers - single 65+ years	22.03	1.24	1354	1.05	1292

At-risk-of-poverty rate after social transfers - 2 adults, no children, both <65	13.50	0.88	3874	2.21	1750
At-risk-of-poverty rate after social transfers - 2 adults, no children, at least one 65+	11.71	0.88	3312	1.92	1726
At-risk-of-poverty rate after social transfers - other households without children	10.71	1.02	5044	1.56	3235
At-risk-of-poverty rate after social transfers - household with children	19.57	0.75	20901	1.50	13943
At-risk-of-poverty rate after social transfers - single parent, at least one child	29.77	2.58	1043	1.58	661
At-risk-of-poverty rate after social transfers - 2 adults, 1 child	11.27	1.03	3771	1.87	2015
At-risk-of-poverty rate after social transfers - 2 adults, 2 children	19.73	1.33	4832	1.58	3058
At-risk-of-poverty rate after social transfers - 2 adults, 3+ children	34.63	2.63	2608	1.46	1786
At-risk-of-poverty rate after social transfers - other households with children	19.26	1.21	8647	1.39	6239
At-risk-of-poverty rate after social transfers - owner or rent free	17.41	0.50	34960	1.51	23214
At-risk-of-poverty rate after social transfers - tenant	22.62	2.12	1760	1.51	1166
At-risk-of-poverty threshold - single	12044.88	85.91	36720	1.51	24318
At-risk-of-poverty threshold - 2 adults, 2 children	25294.25	180.41	36720	1.51	24318
Relative median at-risk-of-poverty gap - total	21.41	0.96	36720	1.52	24206
Relative median at-risk-of-poverty gap - men total	22.81	1.14	17587	1.53	11525
Relative median at-risk-of-poverty gap - women total	20.30	0.83	19133	1.25	15319
Relative median at-risk-of-poverty gap - 0-17 years	22.57	1.24	7353	1.00	7353
Relative median at-risk-of-poverty gap - 18-64 years	22.37	1.25	23680	1.59	14884
Relative median at-risk-of-poverty gap - men, 18-64 years	23.70	1.47	11620	1.37	8457
Relative median at-risk-of-poverty gap - women, 18-64 years	21.13	1.06	12060	1.13	10654
Relative median at-risk-of-poverty gap - 65+ years	15.69	0.99	5687	0.91	6270
Relative median at-risk-of-poverty gap - men, 65+ years	10.89	1.87	2226	0.79	2804
Relative median at-risk-of-poverty gap - women, 65+ years	16.69	1.02	3461	0.83	4175
Inequality of income distribution S80/S20 income quintile share ratio	4.95	0.09	36720	1.56	23493
Dispersion around the at-risk-of-poverty threshold - 40% median	11.08	0.40	13044	1.30	10003
Dispersion around the at-risk-of-poverty threshold - 50% median	12.30	0.51	7040	1.19	5906
Dispersion around the at-risk-of-poverty threshold - 70% median	9.68	0.45	6004	0.99	6046
At-risk-of-poverty rate anchored at a moment in time (2005)	0.92	0.01	36720	1.35	27281
<b>Before social transfers except old-age and survivors' benefits</b>					
At-risk-of-poverty rate before social transfers - total	24.14	0.56	36720	1.54	23922
At-risk-of-poverty rate before social transfers - men total	24.38	0.60	17587	1.33	13223
At-risk-of-poverty rate before social transfers - women total	23.92	0.58	19133	1.19	16038

At-risk-of-poverty rate before social transfers - 0-17 years	30.14	0.91	7353	1.00	7353
At-risk-of-poverty rate before social transfers - 18-64 years	23.84	0.59	23680	1.60	14809
At-risk-of-poverty rate before social transfers - men, 18-64 years	24.49	0.63	11620	1.22	9501
At-risk-of-poverty rate before social transfers - women, 18-64 years	23.20	0.62	12060	1.17	10290
At-risk-of-poverty rate before social transfers - 65+ years	17.31	0.64	5687	1.07	5325
At-risk-of-poverty rate before social transfers - men, 65+ years	12.17	0.90	2226	1.11	2007
At-risk-of-poverty rate before social transfers - women, 65+ years	20.39	0.77	3461	0.89	3911
<b>Before social transfers including old-age and survivors' benefits</b>					
At-risk-of-poverty rate before social transfers - total	43.36	0.60	36720	1.55	23706
At-risk-of-poverty rate before social transfers - men total	41.37	0.67	17587	1.33	13194
At-risk-of-poverty rate before social transfers - women total	45.21	0.62	19133	1.19	16119
At-risk-of-poverty rate before social transfers - 0-17 years	36.72	0.94	7353	1.00	7353
At-risk-of-poverty rate before social transfers - 18-64 years	37.00	0.66	23680	1.59	14931
At-risk-of-poverty rate before social transfers - men, 18-64 years	36.21	0.70	11620	1.19	9765
At-risk-of-poverty rate before social transfers - women, 18-64 years	37.78	0.69	12060	1.19	10126
At-risk-of-poverty rate before social transfers - 65+ years	84.02	0.74	5687	1.22	4646
At-risk-of-poverty rate before social transfers - men, 65+ years	83.70	1.05	2226	1.16	1922
At-risk-of-poverty rate before social transfers - women, 65+ years	84.21	0.77	3461	0.95	3651
Mean equivalised disposable income	23221.46	180.79	36720	1.53	24047
Gini coefficient	31.06	0.38	36720	1.63	22583

## **2.3. Non-sampling errors**

### **2.3.1. Sampling frame and coverage errors**

The samples for EU-SILC were selected from the sampling frame based on the TERYT system, i.e. the *Domestic Territorial Division Register*. Two kinds of primary sampling units (PSUs) were distinguished in the sampling frame:

- about 178 000 *CEA – census enumeration areas* with about 68 dwellings each,
- about 33 000 *ESD – enumeration statistical districts*, with about 377 dwellings each.

The whole territory of Poland is divided into enumeration statistical districts and census enumeration areas. In EU-SILC census enumeration areas are used as primary sampling units. The secondary sampling units are dwellings. For each census enumeration area a list of dwellings was made up to form the secondary sampling frame. All the households from the selected dwellings are supposed to enter the survey.

The TERYT system is updated annually with respect to the territorial division into statistical districts and census enumeration areas. The lists of dwellings, names of towns, villages and streets are updated. Other changes due to new construction, dismantle of buildings and administrative division modifications are also introduced.

The sample for EU-SILC 2005 was selected in September 2004 from the sampling frame updated as for January 1, 2004. In the sample selected some 6.8% of dwellings were found to be non-existing (cancelled, changed for non-residential units), uninhabited or temporarily inhabited, while in the sample 5 selected in 2005 for the 2006 survey about 6.2% of such dwellings were recorded. In the sample 6 selected for the 2007 survey there were about 7% of such dwellings, and in the sample 7 selected for the 2008 survey there were about 6.3% of such dwellings. In the new subsample 8 selected for the 2009 survey 7.5% of dwellings were found to be non-existing (cancelled, changed for non-residential units) as well as uninhabited or temporarily inhabited; 1% of selected dwellings had incorrect addresses.

In the subsample 1 selected for the 2010 survey 8.9% of dwellings were found to be non-existing (cancelled, changed for non-residential units) as well as uninhabited or temporarily inhabited; 1.4% of selected dwellings had incorrect addresses.

In the new subsample 2 selected for the 2011 survey 8.3% of dwellings were found to be non-existing (cancelled, changed for non-residential units) as well as uninhabited or temporarily inhabited; 1.7% of selected dwellings had incorrect addresses.

### **2.3.2. Measurement and processing errors**

As with any other statistical survey, EU-SILC may be burdened with non-sampling errors which occur at various stages of the survey and which cannot be eliminated completely. This mainly applies to interviewers' errors at the stage of collecting the information, errors due to the respondents' misunderstanding of questions and inaccurate or sometimes even false answers as well as the errors taking place at the stage of data recording.

After the household and individual interview completion the respondents were obliged to answer a few questions concerning interview performance. On the basis of this material it is possible to state that about three quarters of respondents (82.5% of those filling in the household questionnaire and 80.4% of those filling in the individual questionnaire) showed a favourable attitude towards the survey, while about 2.0% (both in the case of the household and individual interview) were unwilling towards it. In the interviewers' opinion, in about 84% of household questionnaires and in 79% individual ones the quality of non-income data collected could be recognised as good or very good, while in about 2% of both household and individual questionnaires - as doubtful. The quality of income data was evaluated as slightly worse, mainly because of item non-response. It should also be pointed out that, in our opinion, the quality of data concerning net income categories is much higher than in the case of gross income. The reason is that non-response to the highest degree affected the information on taxes and social and health insurance contributions.

In Poland, the EU-SILC survey in 2011 was carried out in May/July.

EU-SILC 2011, as in previous years was a non-obligatory, representative survey of individual households, performed by a face-to-face interview technique with the use of paper form questionnaires (the so called PAPI method). Two types of questionnaire: individual and household questionnaires were applicable.

The organisation and performance of the survey in the field was within the responsibility of regional statistical offices. Most of the interviewers were regular employees of the statistical offices having experience in other social surveys. Conducting surveys was preceded by series of trainings. Firstly regional survey coordinators were instructed by CSO Labour and Living Conditions Division staff members and then the regional survey coordinators trained interviewers at the regional statistical offices. The interviewers received written instructions concerning the survey performance.

Interviewers' visits to households were preceded by the introductory letter of the CSO President.

Small gifts were given to the families participating in the survey. Each statistical office chose the type of gift for its respondents.

Data recording from the questionnaire forms was carried out with the use of Microsoft Visual FoxPro version 9 operating under the WINDOWS system. The following two applications were designed:

- The so called interviewer's application – to be used by the interviewers to record and check the data from their areas with the use of Laptops and PCs. The data were recorded on the local disk in the VFP database. After the work was completed, the data were transmitted using Web services to the MS SQL server for the national database;
- The so called server application – to be used by the staff of Statistical Offices recording the data directly for the national database and for those supervising the regional data preparation; this application was published in the CITRIX server and made accessible with the customer's software.

Both applications shared a number of modules.

The server application had a module which allowed for works (such as checking, viewing, making statements) on the national data (from all the voivodships). The national file completeness was also checked with the use of Microsoft Visual FoxPro. Additional check-ups were made with SAS checking programmes.

Tables of EU-SILC results were compiled with the use of: SAS, SPSS, Microsoft Visual FoxPro.

### **2.3.3. Non-response errors**

#### Achieved sample size

Sample size	Rotational group				
	7	8	1	2	Total
A	2972	2949	3281	3669	12871
B	6611	6512	7299	7883	28305
C	8455	8458	9458	10349	36720

A - number of households for which an interview is accepted for the database

B - number of persons at the age of 16 years or more who are members of the households for which the interview is accepted for the database, and who completed an individual interview.

C - number of persons who are members of the households for which the interview is accepted for the database.

#### Unit non-response

- Household non-response rates  $NRh = [1 - (Ra * Rh)] * 100$ ,

$Ra = 0.993$

$Rh = 0.857$

$Ra$  – the address contact rate

$Rh$  – the proportion of complete household interviews accepted for the database

$NRh = 14.9$

- Individual non-response rates  $NRp = (1 - Rp) * 100$ ,

$Rp = 0.930$

$NRp = 7.0$

$Rp$  – the proportion of complete personal interviews within the households accepted for the database

- Overall individual non-response rates  $*NRp = [1 - (Ra * Rh * Rp)] * 100$ ,

$*NRp = 20.9$

Information on non-response	Rotational group				
	7	8	1	2	Total
Ra	1.000	1.000	1.000	0.981	0.993
Rh	0.943	0.932	0.934	0.707	0.857
NRh	5.7	6.8	6.6	30.6	14.9
Rp	0.936	0.927	0.937	0.923	0.930
NRp	6.4	7.3	6.3	7.7	7.0
*NRp	11.7	13.6	12.5	36.0	20.9

Distribution of households

- DB120 - Contact at address

DB120	Rotational group				
	7	8	1	2	Total
Address contacted (11)	3153	3164	3513	5191	15021
Address cannot be located (21)	0	0	0	103	103
Address impossible to access (22)	0	0	0	0	0
Address does not exist or is non-residential or is unoccupied or not the principal residence (23)	26	28	40	847	941
<b>Total</b>	<b>3179</b>	<b>3192</b>	<b>3553</b>	<b>6141</b>	<b>16065</b>

- DB130 - Household questionnaire result

DB130	Rotational group				
	7	8	1	2	Total
Household questionnaire completed (11)	2975	2950	3284	3676	12885
Refusal to co-operate (21)	92	116	140	1180	1528
Entire household temporarily away for duration of fieldwork (22)	53	65	67	174	359
Household unable to respond (illness, incapacity,...) (23)	27	27	18	141	213
Other reasons (24)	6	6	4	20	36
<b>Total</b>	<b>3153</b>	<b>3164</b>	<b>3513</b>	<b>5191</b>	<b>15021</b>

- DB135 - Household interview acceptance

DB135	Rotational group				
	7	8	1	2	Total
Interview accepted for database (1)	2972	2949	3281	3669	12871
Interview rejected (2)	3	1	3	7	14
<b>Total</b>	<b>2975</b>	<b>2950</b>	<b>3284</b>	<b>3676</b>	<b>12885</b>

*Item non-response (income variables)*

<b>Item non-response</b>	(A)	(B)	(C)
	% of households having received an amount	% of households with missing values	% of households with partial information
Total household gross income	27.72	9.04	63.17
Total disposable household income	64.19	7.85	27.92
Total disposable household income before social transfers other than old-age and survivor's benefits	64.16	9.83	25.09
Total disposable household income before social transfers, including old-age and survivor's benefits	57.29	14.15	19.22
<b>Net income components at household level</b>			
HY040N	1.02	0.37	0.39
HY050N	14.16	0.48	0.64
HY060N	3.70	0.27	0.04
HY070N	2.28	0.12	0.02
HY080N	5.38	0.87	0.02
HY081N	2.70	0.33	0.00
HY090N	1.55	1.27	0.00
HY100N	2.39	3.67	0.00
HY110N	2.61	0.15	0.02
HY120N	52.66	11.07	0.00
HY130N	4.63	0.37	0.01
HY131N	1.19	0.10	0.00
HY140N	26.99	50.08	21.36
HY145N	40.27	6.35	0.07
HY170N	16.52	3.50	0.00
<b>Gross income components at household level</b>			
HY040G	1.41	0.37	0.00
HY050G	13.32	0.48	1.47
HY060G	3.70	0.27	0.04
HY070G	2.28	0.12	0.02
HY080G	5.38	0.87	0.02
HY081G	2.70	0.33	0.00
HY090G	0.54	1.25	1.03
HY100G	2.39	3.67	0.00
HY110G	2.29	0.07	0.41
HY120G	52.66	11.07	0.00
HY130G	4.63	0.37	0.01
HY131G	1.19	0.10	0.00
HY140G	26.69	50.31	21.62
HY170G	16.52	3.50	0.00

	<b>% of persons 16+ having received an amount</b>	<b>% of persons 16+ with missing values</b>	<b>% of persons 16+ with partial information</b>
<b>Net income components at personal level</b>			
PY010N	31.99	9.00	0.09
PY020N	8.19	3.47	1.48
PY021N	0.30	0.21	0.00
PY035N	2.42	0.80	0.00
PY050N	6.33	3.21	0.50
PY080N	0.00	0.01	0.00
PY090N	2.12	0.25	0.02
PY100N	23.23	2.88	0.49
PY110N	2.71	0.38	0.04
PY120N	0.49	0.07	0.01
PY130N	4.61	0.76	0.07
PY140N	0.82	0.04	0.00
<b>Gross income components at personal level</b>			
PY010G	13.34	9.00	18.74
PY020G	8.19	3.47	1.48
PY021G	0.30	0.21	0.00
PY030G	2.05	29.12	0.35
PY031G	0.61	4.39	0.00
PY035G	2.42	0.80	0.00
PY050G	5.05	2.07	3.80
PY080G	0.00	0.01	0.00
PY090G	0.71	0.25	1.43
PY100G	10.88	2.88	12.85
PY110G	1.10	0.38	1.65
PY120G	0.43	0.07	0.07
PY130G	2.26	0.76	2.41
PY140G	0.82	0.04	0.00
PY200G	27.32	9.11	0.00

### ***Adopted methods of income variable imputation***

Imputation is aimed at obtaining complete records at the level of target variables. Target variables do not simply reflect questionnaire variables and their calculation algorithm is often complicated, although it principally consists in aggregation. So it is necessary to decide what aggregation level the imputation should take place at. There are three possible options:

- the level of questionnaire variables,
- the level of partly aggregated components,
- the level of ready-calculated target variables.

Since the only formal requirement is to obtain imputed target variables, all the above options are permissible and practicable, depending on the specific character of variables. However, the most frequent practice is the imputation at the level of questionnaire variables. There are certain arguments for this approach, on condition that the quantity of data and calculation algorithm details allow for it without much complication.

First of all, imputation at the lowest aggregation level can be desirable for the principal reasons related to the quality of imputation when:

- a target variable implies components of different character (i.e. taking different but rather predictable values, e.g. various social benefits, or dependent on a number of explanatory variables and thus easier to be modelled separately);
- target variables include many components and it is often the case that some of them have the missing items, while others – the correct ones which would be missed during the imputation of an aggregated variable.

Secondly, there are practical arguments for the imputation of disaggregated variables, as the same data serve as a basis for calculating national variables differing from the Eurostat's target variables. Thus the imputation of disaggregated components may be required so as to ensure the imputed data needed for other calculations.

The imputation at the target variable level is carried out only when the above circumstances do not occur or when overcoming the practical difficulties is easier than the imputation of disaggregated data.

There are several methods of component imputation. They can be classified as deterministic and stochastic methods. In case of deterministic methods the selected method and the set of explanatory variables (algorithm) clearly determine the imputation values for each record. In stochastic methods the imputation value is determined with the use of a random component. That is why it may happen that with the same algorithm and the same data file each algorithm realisation will give slightly different imputation values. Although the stochastic methods slightly increase estimator variance (introducing an additional random error component), they do not distort variance or original data distribution characteristics and allow for the correct estimation of random error. Deterministic imputation brings about variable variance reduction in the file and random error underestimation; it also distorts to a greater extent the correlation structure (increasing correlations with explanatory variables). According to item 2.7 of Regulation 1981/2003 it is recommended that for EU-SILC imputation the methods retaining distribution characteristics should be applied, which means the preference for the stochastic methods.

Out of the stochastic methods the following were used in the task presented here:

- Hot-deck method

Random selection of a representative (donor) out of the correct records.

If auxiliary categorizing variables are used in the hot-deck method, a random representative is selected out of the records showing adequate values of auxiliary variables. If it is not possible to find a donor with the equivalent values for all the auxiliary variables, the so called sequence approach is applied. The categorising variables were ranked from the most to the least significant ones. If there are no donors available, categorization is carried out with the subsequent explanatory variables being left out, starting from the least significant ones so as to obtain a subset containing donors.

- Stochastic regression imputation

Auxiliary variables are the explanatory variables of the regression model. The model takes the linear form or the logarithmic transformation is used. It is fitted on the basis of the correct records. The imputed value (or its logarithm in the case of transformed models) is a sum of the theoretical value derived from the model and a randomly selected model residual. The set of records of which the residual is selected is restricted to those which are nearest to the record imputed for the theoretical value derived from the model.

Out of the deterministic methods the following are applied:

- Regression deterministic imputation

The theoretical value from the model is adopted as the imputation value.

- Deduction imputation

The imputation value is directly determined on the basis of the relationships between variables.

In the case of imputation at the target variable level or imputation of the most significant components of target variables, stochastic imputation is applied in order to retain the variable properties distribution as required by Regulation 1981/2003.

The application of stochastic regression imputation requires a model which describes well the formation of a variable with relatively small variance of an error term and good statistical qualities. With high variance of an error term, there is a danger of getting accidental values which are not typical of the correct part of the dataset. That is why in the cases where, in accordance with the assumption referred to above, stochastic imputation is required, the hot-deck method is applied in preference to regression imputation. This is particularly justified when the number of records for imputation is rather low, or when the number of correct records is too small for a suitable model fitting.

Stochastic regression imputation is most widely used for incomes from hired employment, as:

- it is an important category of income, declared by a significant rate of respondents which, if present, has a significant share in the total household's income;
- this category can be successfully modelled with the use of the variables included in the questionnaire;
- there is a large (absolute) number of missing data, the percentage, however, being rather small; a large number of correct records make it possible to design a well-fitted model.

In case of incomes from hired employment stochastic regression imputation is applied to the majority of records with missing items, both those for which observations from the previous year are available (panel sample) and the new ones in the sample. In case of other income categories stochastic regression imputation is used as the basic imputation method when incomes of the same type for a given person/household are known from the previous year. If such income data from the previous year are not available, the hot-deck method is applied. The hot-deck method is also applied when the income data are known from the previous year

but a suitable model fitting is difficult. In such a case the income from the previous year is used as a grouping variable. If the quantitative categorizing variable is applied in the hot-deck method, the categorization criterion is a break-down into deciles.

Considering a relatively wide application of the stochastic regression imputation, supplementary protection against the effects of potential insufficient model adequacy was introduced. The residuals are not generated from the distribution of residuals for the whole sample but they are selected from a restricted subset. Although in an ideal model residuals should be in the form of white noise, showing no trend whatsoever, in reality some trends can be observed in the distribution of residuals which are not detected by the model (like those related to non-linearity of relationships which cannot be removed by known transformations).

In such a case, if we used residuals from the whole range, we could combine a particular theoretical value obtained from the model with the residual which occurs in the whole distribution but is quite improbable in combination with this particular theoretical value. So we could generate values significantly diverging from the real variable distribution. The use of residuals from the restricted range only reduces that risk.

Deterministic imputation is applied where missing data concern less significant components of target variables (taxes, burdens to the main component, additions, etc.) in the situation when the main component is known. In such cases deterministic regression imputation is usually applied. Gross/net conversion is carried out with the use of the deterministic regression method. Deduction imputation is employed in rare cases of obvious relationships and can be treated as a supplementary stage of data editing.

The explanatory variables in the models and the grouping ones in the case of hot-deck method have been selected so as to represent the relationships which, according to logics and knowledge about the phenomena studied, should occur in the data set, taking into account accessibility of the potential variables in the questionnaire. The relationships have been tested on the file of correct data and in the majority of cases they proved to be significant. Some of the explanatory variables have been retained, even if their impact on the imputed variable has not been statistically confirmed, if they expressed an economically important relationship or provided a grouping condition (interpretation criterion) in the calculation algorithm.

For the persons and households not surveyed in the previous year (a new sample, new household members, persons who could not be interviewed) or for those who did not gain a particular type of income in the previous year, explanatory variables derived from the current data file are applied. Wherever the same type of income is found in the data for the previous year, its value is treated as the main explanatory (categorizing) variable, both in the case of variables subjected to regression imputation and the hot-deck method. The current variables can be treated as additional explanatory variables.

### ***Imputation of the missing individual questionnaires***

The imputation of the missing individual questionnaires is carried out with the use of the hot-deck method. A wide set of variables providing household's characteristics (main source of maintenance) and variables from R set determining the person's position in the household and on the labour market is used as the categorization criterion. All the primary target variables related to the donor are transferred to the taker's record and then they are used for the calculation of household's total income. The records obtained as a result of imputing the

missing questionnaires are attached to the individual income data files, while the income data are included in the total income indicated in the household data file. this makes the files coherent.

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

Indicator	Achieved sample size	Total item non-response
At-risk-of-poverty rate after social transfers - total	36720	15236
At-risk-of-poverty rate after social transfers - men total	17587	7519
At-risk-of-poverty rate after social transfers - women total	19133	7717
At-risk-of-poverty rate after social transfers - 0-17 years	7353	3190
At-risk-of-poverty rate after social transfers - 18-64 years	23680	10488
At-risk-of-poverty rate after social transfers - men 18-64 years	11620	5282
At-risk-of-poverty rate after social transfers - women 18-64 years	12060	5206
At-risk-of-poverty rate after social transfers - 65+ years	5687	1558
At-risk-of-poverty rate after social transfers - men 65+ years	2226	609
At-risk-of-poverty rate after social transfers - women 65+ years	3461	949
At-risk-of-poverty rate after social transfers - employed	13044	5559
At-risk-of-poverty rate after social transfers - men employed	7040	2936
At-risk-of-poverty rate after social transfers - women employed	6004	2623
At-risk-of-poverty rate after social transfers - unemployed	1542	560
At-risk-of-poverty rate after social transfers - men unemployed	723	261
At-risk-of-poverty rate after social transfers - women unemployed	819	299
At-risk-of-poverty rate after social transfers - retired	6928	1836
At-risk-of-poverty rate after social transfers - men retired	2688	693
At-risk-of-poverty rate after social transfers - women retired	4240	1143
At-risk-of-poverty rate after social transfers - other inactive	6299	2340
At-risk-of-poverty rate after social transfers - men other inactive	2326	853
At-risk-of-poverty rate after social transfers - women other inactive	3973	1487
At-risk-of-poverty rate after social transfers - household without children	14763	5027
At-risk-of-poverty rate after social transfers - single total	2533	483
At-risk-of-poverty rate after social transfers - single male	732	160
At-risk-of-poverty rate after social transfers - single female	1801	323
At-risk-of-poverty rate after social transfers - single <65 years	1179	299
At-risk-of-poverty rate after social transfers - single 65+ years	1354	184
At-risk-of-poverty rate after social transfers - 2 adults, no children, both <65	3874	1334
At-risk-of-poverty rate after social transfers - 2 adults, no children, at least one 65+	3312	750
At-risk-of-poverty rate after social transfers - other households without children	5044	2460
At-risk-of-poverty rate after social transfers - household with children	20901	9271
At-risk-of-poverty rate after social transfers - single parent, at least one child	1043	333
At-risk-of-poverty rate after social transfers - 2 adults, 1 child	3771	1458
At-risk-of-poverty rate after social transfers - 2 adults, 2 children	4832	1996
At-risk-of-poverty rate after social transfers - 2 adults, 3+ children	2608	952
At-risk-of-poverty rate after social transfers - other households with children	8647	4532
At-risk-of-poverty rate after social transfers - owner or rent free	34960	14667

At-risk-of-poverty rate after social transfers - tenant	1760	569
At-risk-of-poverty threshold - single	36720	15236
At-risk-of-poverty threshold - 2 adults, 2 children	36720	15236
Relative median at-risk-of-poverty gap - total	36720	15236
Relative median at-risk-of-poverty gap - men total	17587	7519
Relative median at-risk-of-poverty gap - women total	19133	7717
Relative median at-risk-of-poverty gap - 0-17 years	7353	3190
Relative median at-risk-of-poverty gap - 18-64 years	23680	10488
Relative median at-risk-of-poverty gap - men, 18-64 years	11620	5282
Relative median at-risk-of-poverty gap - women, 18-64 years	12060	5206
Relative median at-risk-of-poverty gap - 65+ years	5687	1558
Relative median at-risk-of-poverty gap - men, 65+ years	2226	609
Relative median at-risk-of-poverty gap - women, 65+ years	3461	949
Inequality of income distribution S80/S20 income quintile share ratio	36720	15236
Dispersion around the at-risk-of-poverty threshold - 40% median	13044	5559
Dispersion around the at-risk-of-poverty threshold - 50% median	7040	2936
Dispersion around the at-risk-of-poverty threshold - 70% median	6004	2623
At-risk-of-poverty rate anchored at a moment in time (2005)	36720	15236
<b>Before social transfers except old-age and survivors' benefits</b>		
At-risk-of-poverty rate before social transfers - total	36720	14868
At-risk-of-poverty rate before social transfers - men total	17587	7334
At-risk-of-poverty rate before social transfers - women total	19133	7534
At-risk-of-poverty rate before social transfers - 0-17 years	7353	3092
At-risk-of-poverty rate before social transfers - 18-64 years	23680	10237
At-risk-of-poverty rate before social transfers - men, 18-64 years	11620	5161
At-risk-of-poverty rate before social transfers - women, 18-64 years	12060	5076
At-risk-of-poverty rate before social transfers - 65+ years	5687	1539
At-risk-of-poverty rate before social transfers - men, 65+ years	2226	600
At-risk-of-poverty rate before social transfers - women, 65+ years	3461	939
<b>Before social transfers including old-age and survivors' benefits</b>		
At-risk-of-poverty rate before social transfers - total	36720	14301
At-risk-of-poverty rate before social transfers - men total	17587	7068
At-risk-of-poverty rate before social transfers - women total	19133	7233
At-risk-of-poverty rate before social transfers - 0-17 years	7353	3027
At-risk-of-poverty rate before social transfers - 18-64 years	23680	9934
At-risk-of-poverty rate before social transfers - men, 18-64 years	11620	5026
At-risk-of-poverty rate before social transfers - women, 18-64 years	12060	4908
At-risk-of-poverty rate before social transfers - 65+ years	5687	1340
At-risk-of-poverty rate before social transfers - men, 65+ years	2226	505
At-risk-of-poverty rate before social transfers - women, 65+ years	3461	835
Mean equivalised disposable income	17587	7068
Gini coefficient	19133	7233

## 2.4. Mode of data collection

EU-SILC is a non-obligatory, representative survey of individual households, performed by a face-to-face interview technique with the use of paper form questionnaires (the so called PAPI method). Two types of questionnaire: individual and household questionnaire are applicable.

### Distribution of RB250 and RB260

- RB250 – Data status

RB250	Rotational group				
	7	8	1	2	Total
Information completed only from interview (11)	6611	6512	7299	7883	28305
Information not completed: imputed (14)	453	513	491	659	2116
<b>Total</b>	<b>7064</b>	<b>7025</b>	<b>7790</b>	<b>8542</b>	<b>30421</b>

- RB260 – Type of interview

RB260	Rotational group				
	7	8	1	2	Total
Face to face (1)	5316	5217	5852	6376	22761
Proxy interview (5)	1295	1295	1447	1507	5544
<b>Total</b>	<b>6611</b>	<b>6512</b>	<b>7299</b>	<b>7883</b>	<b>28305</b>

As for individual interviews, in 2011 a relatively high share (19.6 %) of proxy interviews was noted. This was thoroughly discussed with the survey coordinators in the field.

The interviewers decided on proxy interviews only if the substitute respondents were well informed about the situation in the household and there was no other possibility to get the information. Proxy interviews were performed in the following situations:

- no contact with the respondent because of long-term absence (e.g. work in another town or abroad);
- respondent's disability, illness or pathology (such as alcoholism);
- according to other members of the household, the respondent was only available late at night and was not willing to participate in such a long interview, while at the same time the proxy could provide detailed information, even based on the documents, such as tax statements.

### **2.5. Interview duration**

The average household interview duration was about 33 minutes, while the average individual interview duration was about 22 minutes. In total the average time needed to carry out a household interview and individual interviews with persons at the age of 16 years and over was 82 minutes.

This value exceeded significantly that assumed in the regulation, which results from the fact that in the Polish SILC all the information is collected during the interview. The questionnaire parts covering social benefits and self-employment (in and outside farming) have been expanded by many auxiliary questions which help to answer but, on the other hand, prolong the interview. Problem of the interview duration was already pointed out in the Intermediate Quality Report for EU-SILC 2005 - 2010.

### **3. Comparability**

#### **3.1. Basic concepts and definitions**

##### *Reference population*

No difference to the common definition.

The survey unit was a household and all the household members who would have reached the age of 16 years by December 31, 2010.

The survey did not cover collective accommodation households (such as boarding houses, workers' hostels, pensioners' houses or monasteries), except for the households of the staff members of these institutions living in these buildings in order to do their job (e.g. hotel manager, tender etc.).

The households of foreign citizens should participate in the survey.

##### *Private household definition*

No difference to the common definition.

Household is a group of persons related to each other by kinship or not, living together and sharing their income and expenditure (multi-person household) or a single person, not sharing his/her income or expenditure with any other person, whether living alone or with other persons (one-person household).

Family members living together but not sharing their income and expenditure with other family members make up separate households.

The household size is determined by the number of persons comprised by the household.

##### *Household membership*

No difference to the common definition.

The household composition accounted for:

- persons living together and sharing their income and expenditure who have been in the household for at least 6 months (either the real or the intended time of staying in the household should be considered),
- persons absent from the household because of their occupation, if their earnings are allocated to the household's expenditure,
- persons at the age of up to 15 years (inclusive), absent from the household for education purposes, living in boarding houses or private dwellings,
- persons absent from the household at the time of the survey, staying at education centres, welfare houses or hospitals, if their real or intended stay outside the household is less than 6 months.

The household composition did not account for:

- persons at the age of over 15 years, absent from the household for education purposes, living in boarding houses, students' hostels or private dwellings,
- men in military service (those performing substitute military service working in companies and living at home are included in the household),
- persons in prison,

- persons absent from the household at the time of the survey, staying at education centres, welfare houses or hospitals, if their real or intended stay outside the household is more than 6 months,
- persons (household's guests) staying in the household at the time of the survey who have been or intended to be there for less than 6 months,
- persons renting a room, including students (unless they are treated as household members),
- persons renting a room or bed for the time of work in a given place (including such works as land melioration, geodetic measurements, forest cut-down or building constructions),
- persons living in the household and employed as au pairs, helping personnel on the farm, craft apprentices or trainees.

*Income reference period(s)*

No difference to the common definition.

The income reference period was the previous calendar year (2010).

*Reference period for taxes on income and social insurance contributions*

The reference period for income tax prepayment and compulsory social insurance contributions is the year 2010. The account clearance with the Treasury Office (including payments and returns) effected in 2010 refers to the income for 2009.

*Reference period for taxes on wealth*

No difference to the common definition.

Taxes on wealth paid during the income reference period (2010) were recorded.

*Lag between the income reference period and current variables*

The lag between the income reference period and current variables is about 5 months.

*Total duration of data collection*

EU-SILC was performed on the territory of the whole country between May 4 and June 28 2011.

*Basic information on activity status during the income reference period*

No difference to the common definition.

## **3.2. Components of income**

### ***3.2.1. Differences between national definitions and standard EU-SILC definitions, and assessment:***

#### Total gross household income HY010

No difference to the common definition.

#### Total disposable household income HY020

No difference to the common definition.

#### Total disposable household income before social transfers except old-age and survivors` benefits HY022

No difference to the common definition.

#### Total disposable household income before social transfers including old-age and survivors` benefits HY023

No difference to the common definition.

In accordance with EU-SILC 065 (2011 operation) the new income components, mandatory from 2007 operation onwards, including:

- PY020G – NON-CASH EMPLOYEE'S INCOME;
- PY030G – EMPLOYER'S SOCIAL INSURANCE CONTRIBUTION;
- PY070G/HY170 – VALUE OF GOODS PRODUCED FOR OWN CONSUMPTION;
- PY080G – PENSION FROM INDIVIDUAL PRIVATE PLANS;
- HY030G – IMPUTED RENT;
- HY100G – INTEREST REPAYMENTS ON MORTGAGE

have been recorded at a component level only and are not included in the household's total income (variables: HY010G; HY020G; HY22G; HY023G).

#### Imputed rent HY030

For the purposes of imputed rent estimation, regression analysis has been used. It was decided to use econometric methods, and especially regression analysis, based on the representative method, the weighted least square method (WLSM) was applied. It was assumed that the estimated function of rentals is an exponential function which means that in the estimation form the dependent variable is a logarithm of variable.

#### Value of goods produced for own consumption HY170

This variable has been introduced starting from the 2011 operation (In accordance with EU-SILC 065). It corresponds to the former variable PY070.

### Employee's cash or near cash income PY010

This variable does not account for:

- assistance for foster families; since granting the benefit is not connected with quitting the job, this benefit has been qualified to the category of "Family related allowances" (HY050)

### Non-cash employee income's PY020

Company car (PY021) – the information on the private use of the company car is collected in the individual questionnaire. Here belongs the respondent's estimated amount he/she has gained by using the company car for private purposes. In the case of the missing value (the respondent was using the company car but did not estimate the amount gained), imputation is applied with the use of hot-deck and regression imputation with simulated residuals methods;

### Cash benefits or losses from self-employment PY050

The data on income from self-employment were collected in two different ways: the respondents were asked about the company's costs and profits and also about the amount of money gained from self-employment which was allocated to the household's expenditure. After a detailed analysis of data it was decided that the income from self-employment would be equal to the amount allocated to the household's needs.

### Survivors` benefits PY110

Death grants are not included in the income because the whole sum is used to cover the cost of the funeral.

### Sickness benefits PY120

Sickness and childcare benefits are not included (a childcare benefit is granted to the working parent of a sick child), because they are paid by the employer and cannot be detached from the income from hired employment. Therefore, they are accounted for in the income from hired employment.

*All the other variables not listed above*

### *Dwelling conditions and material deprivation items*

The analysis of questions and explanatory notes from the guideline for interviewers concerning dwelling conditions and material deprivation items showed that some records differed from those included in document 065/11.

*Ability to make ends meet* – the question does not specify that "net monthly income" is meant. However, a detailed explanation is given in the guidelines for the interviewer in accordance with the recommendations of Eurostat. In the Polish conditions it is not necessary to use the term "monthly net income". If we ask about the income, it is clear for the respondent that we mean monthly net income.

There were no other major divergences from common definitions.

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

The income data were collected during the interviews with respondents. The target income variables were split into components corresponding to particular benefits applicable in the Polish conditions.

### ***3.2.3. The forms in which income variables at a component level have been obtained***

The respondents were asked to give the net incomes and contributions (income tax prepayments and compulsory social insurance). Only in the case of income from rental of a property (HY040) the respondents were asked to give the gross income and the amount of tax paid.

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

The gross income was obtained by summing up net values, income tax prepayments and compulsory social insurance contributions. If the information on tax and insurance contributions was missing, the amounts were imputed on the basis of the results obtained. Only in the case of income from rental of property, the tax paid was subtracted from the gross income.

## **4. COHERENCE**

### **4.1. Comparison of EU-SILC and HBS results**

The objective of this section is to compare HBS (Household Budget Survey) and EU-SILC results.

Up to 2004 the HBS provided the main source of data on the living conditions of the Polish population, among others on incomes, dwelling conditions and households' equipment.

The HBS has been regularly conducted every year since 1993 up to now with the use of the rotational method. The households are surveyed in the two year panel.

In HBS the main source of data on income and expenditure is provided by the diaries, while

that concerning dwelling-related expenditure and utilities – by BR-01a questionnaire.

In addition, three other questionnaires are filled in.

When comparing these two sources we must take into account the discrepancies. The differences are to great extent brought about by the methodological diversity. Here are the main diverging points:

- Different reference periods for income variables – in HBS the reference period is 1 month and, following Eurostat's recommendation, the annual income is the monthly income multiplied by 12, which in the case of irregular income, like that from farming, can bring about considerable distortions. In EU-SILC the reference period is a calendar year preceding the survey;
- Different types of income are taken into account i.e. in HBS the information is collected both about the income in cash and in kind, while in EU-SILC – only about the income in cash (with a few exceptions), which may be important for the income from farming and social benefits other than retirement pay and pension. Moreover, EU-SILC does not take into account the so called lump sums which is the case in HBS;
- Different way of data collection – in HBS the respondents make records in the so called diary. They have to determine the data sources themselves and do not have them listed in the diary. This may cause omissions. In EU-SILC each respondent is asked detailed questions. In EU-SILC all the income missing data are imputed, while there is no imputation in HBS;
- Different way of sample selection – in HBS dwellings in which all the households refused to participate in the survey are replaced with new ones from the so called reserve list;
- Slightly different weighting of results.

In some tables given below the data are presented in the breakdown by socioeconomic group and household size. The household survey results are usually released by CSO in the breakdown by socioeconomic group and household size.

The main criterion for socio-economic group classification is the prevailing source of income. In tables below only weighted data are presented.

**Tab. 1. Structure of population by age**

Specification	EU-SILC 2011	HBS 2011
	in %	
Total	100.0	100.0
0-14	15.1	18.1
15-24	13.0	14.4
25-54	43.9	41.1
55-64	13.9	13.9
65+	14.1	12.5

**Tab. 2. Structure of population by level of education**

Specification	EU-SILC 2011	HBS 2011
	in %	
Total	100.0	100.0
No school education	1.4	0.5
Completed primary	15.3	15.0
Lower secondary	4.8	6.5
Elementary vocational	26.6	27.4
Secondary	34.2	33.0
Higher	17.6	17.6

**Tab. 3. Structure of households and persons in households by socio-economic group**

Households	Households		Persons in households	
	EU-SILC 2011	HBS 2011	EU-SILC 2011	HBS 2011
Total	13200884	13332336	37473013	37723167
<b>Total = 100</b>				
Employees	52.9	49.5	63.0	58.5
Farmers	2.7	4.0	3.6	6.0
Self-employed	5.4	6.9	6.1	8.2
Retirees	28.0	27.9	19.3	18.8
Pensioners	7.0	7.2	4.2	4.6
Maintained from non-earned sources	4.0	4.5	3.7	3.9

**Tab. 4. Average yearly equivalent income in PLN by socio-economic group**

Households	Disposable income 2010		Income from hired work 2010	
	EU-SILC	HBS	EU-SILC 2010	HBS 2009
Total	23187	21227	14108	11705
Employees	25165	22324	20923	18366
Farmers	15616	20520	1636	2054
Self-employed	28891	27600	4846	4830
Retirees	19788	18614	2298	2043
Pensioners	14058	13600	1106	1373
Maintained from non-earned sources	15668	13330	2090	1089

**Tab. 5. Average yearly equivalent income in PLN by number of persons**

Households	Disposable income 2010		Income from hired work 2010	
	EU-SILC	HBS	EU-SILC	HBS
Total	23187	21227	14108	11705
1-person	20588	20196	7099	6621
2-persons	26258	24452	12349	10647
3-persons	26595	23934	18410	15258
4-persons	23360	21769	16748	14283
5-persons	20942	18307	13252	10437
6-persons and more	18269	15971	10390	7337

**Tab. 6. Households provided with selected durables**

Specification	EU-SILC 2011	HBS 2011
	in %	
Fixed telephone	57.8	51.2
Mobile telephone	87.7	90.3
Television set	97.8	98.4
Computer	65.5	66.7
Printer	46.8	40.2
Internet connection	59.8	62.3
Microwave oven	52.7	53.3
Dishwasher	18.4	17.7
Refrigerator	98.5	98.6
Washing machine	97.8	98.0
Passenger car	62.7	60.2

## **4.2. Comparison of Laeken Indicators based on EU-SILC 2010 and EU-SILC 2011**

In 2011 further improvement in the income position of the Polish household, observed since 2004, was maintained.

In Poland the average yearly equivalent disposable income in households amounted to 5813 EUR (23 221 PLN) and increased by 4.6 p.p when compared to previous year.

In general, no significant changes were noted in 2011 as regards income differentiation. Gini coefficient was at the same level (31.4 in 2009 as compared with 31.1 in 2010), but was significantly lower than in 2005. Income inequality in the period from 2005 to 2011 decreased to a level closer to the EU27 average. However, the pace of decline of income inequality in Poland decreased.

Poland is among EU27 countries with highest at risk of poverty or social exclusion indicator (27.2% of population) when compared to the EU average (24.1%). However, in Poland a gradual decline in the value of this indicator can be observed. Since 2005, this indicator dropped by 18.1 percentage points.

In 2011 in Poland, 17.7% of the population was at risk of population. In 2011 in Poland the poverty threshold per person amounted to 3015 EUR and poverty threshold for a family with two children amounted to 6332 EUR.

Since 2005 the value of material deprivation indicator for Poland is steadily decreasing. While in 2005, the situation affected every third person in 2011 it affected only every seventh person. Despite this, the value of material deprivation rate for Poland is still much higher than the average value for EU27 ( 13% for Poland compared to 8.8% for EU).

The value of low-intensity work rate for Poland in 2011 was 6.9% and for several years has been below the EU average (which for 2011 was 10%).

## **4.3. Comparison between SNA results for the household sector and EU-SILC 2011 (data for 2010) in the scope of incomes**

The comparison covered disposable income and its main components: income from hired employment, self-employment (in and outside farming), property income, social benefits as well as current taxes on income.

It was found out that the disposable income in EU-SILC 2011 made up 63.7% of the corresponding category in SNA. This was due to the following factors:

1. The household sector in SNA includes collective households which do not enter EU-SILC.
2. Both systems employ different methods of measuring income from self-employment.
3. Accounts of primary and secondary income distribution in SNA used for the determination of disposable income include some items not covered by EU-SILC or not taken into account in the calculation of its results. The most important of them are imputed rents.

In order to minimize the impact of the last item, the analysis was made not only on aggregate variables (EU-SILC dataset), but also based on variables from the national dataset.

The table presents the results of this analysis.

<b>Macro Code</b>	<b>Macro Name</b>	<b>Macro amount (Ma)</b>	<b>Micro Code</b>	<b>Micro name</b>	<b>Micro amount (Mi)</b>	<b>Coverage rate (Mi/Ma) in %</b>
D1R	Compensation of employees, received: - wages and salaries;	467 337	PY010G PY020G	Employee cash or near cash income; Non-cash employee income;	451 182	96,5
	<b>Changes on macro data</b> adding: - employers` social contributions	534 275		<b>Changes on micro data</b> deducting: - employee cash or near cash income from abroad; - self-employment treated as employment (hired work) gross adding: - one-time retirement pay gross; - one-time redundancy pay gross; - employer`s social insurance contributions	482 110	90,2
B2G	Gross operation surplus with the exception of section L (NACE rev.2)	284 801	PY050G	Cash benefits or losses from self-employment	82 265	28,9
B2G, B3G	<b>Changes on macro data</b> Gross operation surplus and mixed income, received	336 825		<b>Changes on micro data</b> adding: - self-employment treated as employment (hired work) gross; - imputed rent; - value of goods produced for own consumption	183 651	54,5
D4R	Property income, received <b>Changes on macro data</b> deducting: - property income attributed to insurance policy holders	44 555	HY090G	Interests, dividends, profit from capital investments in unincorporated business	5797	13,0

<b>Macro Code</b>	<b>Macro Name</b>	<b>Macro amount (Ma)</b>	<b>Micro Code</b>	<b>Micro name</b>	<b>Micro amount (Mi)</b>	<b>Coverage rate (Mi/Ma) in %</b>
D62R	Social benefits other than social transfers in kind, received	210 072	HY050G, HY060G, HY070G, PY090G, PY100G, PY110G, PY120G, PY130G, PY140G	Social benefits (the sum of all the nine categories)	183 177	87,2
D5P, D6P	Current taxes on income, wealth, etc. and social contributions paid	246 751	HY120G, HY140G, PY030G	Regular taxes on wealth, tax on income and social contributions, employer's social insurance contributions	202 582	82,1
B6g	Gross disposable income	889 437	HY020  <b>Changes on micro data</b> deducting: - company car; - interest repayment on mortgage adding: - non-cash employee income; - value of goods produced for own consumption; - imputed rent	Total disposable household income	566 236	63,7
					659 197	74,1

Modification of individual categories consisted in retaining the common components wherever possible.

Gross disposable income in the EU-SILC makes up about 63.7% of the corresponding category in the SNA.

After the modifications on microdata the coverage rate increased by 10.4 percentage points and reached 74.1%, which seems to be a good result.

The main components of disposable income, i.e. income from hired employment and social benefits, at the microdata level are similar to the results obtained based on the SNA.

After the EU-SILC income modification the income from paid employment reached 90.2% and income from social benefits 87.2%.

Income from self-employment is the only income category that does not give so good results (54.5%). In the Polish EU-SILC the question about income from self-employment concerns just the amount allocated to the household's consumer needs and its housing-related investment.

Property income is the least comparable component (13.0%). The EU-SILC survey is carried out by the direct (face-to-face) interview method. It becomes rather difficult to get from the respondents information about the occurrence and amount of this income. We are investigating a possibility to obtain some of the income variables from administrative sources.

Tax and social insurance make up a position which has also gained a relatively high coverage rate (82.1%).