

# **FINAL QUALITY REPORT**

## **EU-SILC 2006 OPERATION POLAND**

**Warsaw, November 2008**

<b>CONTENTS</b>	<b>Page</b>
<b>PREFACE .....</b>	<b>3</b>
<b>1. COMMON LONGITUDINAL EUROPEAN UNION INDICATORS .....</b>	<b>4</b>
<b>2. ACCURACY .....</b>	<b>4</b>
2.1. Sample design .....	4
2.2. Sampling errors .....	8
2.3. Non-sampling errors.....	13
2.4. Mode of data collection.....	22
2.5. Imputation procedure .....	24
2.6. Imputed rent .....	27
2.7. Company cars .....	27
<b>3. COMPARABILITY .....</b>	<b>27</b>
3.1. Basic concepts and definitions .....	27
3.2. Components of income.....	29
3.3. Tracing rules.....	30
<b>4. COHERENCE .....</b>	<b>31</b>
4.1. Comparison of EU-SILC and HBS results.....	31
4.2. Comparison of income data from SNA for the household sector and EU-SILC .....	34

## **PREFACE**

The present quality report is the final quality report of EU-SILC 2006 in Poland according to grant agreement No. 36400.2005.001-2005.435, as provided for in Council Regulation No 1177/2003 and it following the structure outlined in Commission Regulation No. 28/2004. This report provides information on accuracy, comparability and coherence of data with external sources.

In Poland EU-SILC operations started in 2005, so by the year 2006 the panel data has not fully matured yet. Currently, 3 of the 4 sub-samples which form the total sample of the EU-SILC operation 2006 are a panel of 2 consecutive years duration: 2005 and 2006.

## **1. COMMON LONGITUDINAL EUROPEAN UNION INDICATORS**

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

## **2. ACCURACY**

### **2.1. Sample design**

#### *Type of sampling design*

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

#### *Sampling units*

The first-stage primary sampling units (PSU) were census areas, while at the second stage dwellings were selected.

#### *Stratification and substratification criteria*

The strata were the voivodships (NUTS2), while within voivodships primary sampling units were classified by class of locality. In urban areas census enumeration 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 poviats (NUTS4). Altogether 211 strata were distinguished.

#### *Sample size and allocation criteria*

It was decided that the sample should include some 24 000 dwellings. Proportional allocation of dwellings to particular strata was applied. The number of dwellings selected from a particular stratum was in proportion to the population number of dwellings in the stratum. Furthermore, the number of the 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 from each PSU were selected. Altogether 5912 census areas and 24044 dwellings were selected for the sample. The subsample 5 was selected for the survey in 2006 in order to replace the sub-sample 1. It consisted of 1476 census areas and 6002 dwellings. Sub-samples 2, 3, 4 formed the longitudinal (panel) component.

#### *Sample selection schemes*

Census areas were selected according to the Hartley-Rao scheme. Prior to selection census areas were put in random order separately for each stratum and then the determined number of PSU 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.

### ***Sample distribution over time***

The sample is not distributed over time.

### ***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 another one, selected independently as described above. For the 2006 survey the subsample 5 was selected as a replacement of the subsample 1. Rotation comprised first-stage units.

### ***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 PSU selected from the h-th stratum,

$m'_h$  - number of dwellings selected from 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 completeness indicator, estimated for each class of locality separately:

<b>Code of class of locality (p)</b>	<b>Class of locality</b>	<b>Completeness rate (<math>R_{ap} * R_{hp}</math>)</b>
	Poland	0.699
1	Warsaw	0.399
2	Towns 500 000 – 1 000 000 inhabitants	0.567
3	Towns 100 000 – 500 000 inhabitants	0.636
4	Towns 20 000 – 100 000 inhabitants	0.695
5	Towns less than 20 000 inhabitants	0.740
6	Rural areas	0.823

The adjusted weights were calculated according to the formula:

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

The weights DB080 and  $DB080^{corrected}$  were calculated for the subsample 5. The next step consisted in calculating the weights DB090 and RB050 for the households of the subsample 5 with the use of the integrated calibration method as described below in section *Adjustments to external data*.

### 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 (14 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 were derived from the 2002 Census and current demographic estimates.

### Final longitudinal weight

Panel weights are described in section below.

### Non-response adjustments

For the subsamples 2, 3 and 4, surveyed for the second time, the base weights were determined by the correction of the base weights from the previous year. The base weight of 2005 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 performed on the subsets of the so called *sample persons* i.e. those who were in the surveyed sample at the age of 14 and over in 2005 and who should be surveyed in 2006. The modifying factor was determined for each subsample (2, 3 and 4) separately 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 (t = 1, 2),

$M$  – 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 2005 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’ weights were averaged and all the members of a given household were ascribed such a mean weight. Then for the base weights thus obtained the trimming of extreme weights was applied.

#### Adjustments to external data

Adjustment to external data was not applied.

#### Final longitudinal weight

The panel weight RB062 was calculated by dividing the base weights by 3.

#### Final household cross-sectional weight

The last stage of calculations consisted in combining the four independent subsamples, applying the integrated calibration as described below (for sample 5 repeatedly) and trimming of extreme weights. As a result the following cross-sectional weights were calculated for households and individuals from samples 2, 3, 4 and 5 in EU-SILC 2006:

DB090 – weight for households,

RB050 – weight for all household members but

$$RB050_{ij} = DB090_i$$

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 is obtained by the adjustment of RB050 separately in the groups according to gender and age in each voivodship according to urban and rural area,

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.

#### ***Substitutions***

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

## 2.2. Sampling errors

### *Standard error and effective sample size*

Estimation of standard errors was based on resampling approach. We used a bootstrap method which resamples 200 times from each stratum  $n_h - 1$  PSU's (primary sampling units) with replacement (McCarthy and Snowden method (1985)), where  $n_h$  denotes the sample size of PSU's 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 also implemented linearization method of variance estimation for main poverty indicators, and the results of comparisons with bootstrap method were very similar.

### *Cross-sectional component*

The mean, the total number of observations (before and after imputation) and the standard errors for the following income components (mean and standard errors based on weighted data while the number of observations based on unweighted results).

Income components	Mean <sup>1</sup>	Standard error	Mean <sup>2</sup>	Standard error	Number of observations	
					Before imputation	After imputation
Total household gross income (HY010)	36728	244	36689	243	6267	14900
Total disposable household income (HY020)	27725	170	27709	170	10892	14906
Total disposable household income before social transfers other than old-age and survivors benefits (HY022)	25964	179	25511	175	10830	14684
Total disposable household income including old-age and survivors benefits (HY023)	20861	195	18501	171	9838	13351
<b>Net income components at household level</b>						
HY040N	6048	511	90	11	156	204
HY050N	2090	49	422	11	3454	3529
HY060N	1440	67	73	4	790	806
HY070N	1262	36	68	3	800	831
HY080N	4889	193	354	19	925	1002
HY090N	4085	448	60	8	113	191
HY110N	2181	117	67	5	551	560
HY120N	215	5	103	3	6835	7467
HY130N	3600	159	199	12	768	809
HY140N	9147	80	8888	78	6184	14538
HY145N	-450	36	-210	17	6557	6953

<sup>1</sup> Taking into account only households/persons receiving such income.

<sup>2</sup> Taking into account whole population (households/persons) surveyed.

Income components	Mean <sup>1</sup>	Standard error	Mean <sup>2</sup>	Standard error	Number of observations	
					Before imputation	After imputation
<b>Gross income components at household level</b>						
HY040G	7017	574	105	12	174	204
HY050G	2159	54	436	12	3320	3529
HY060G	1440	67	73	4	790	806
HY070G	1262	36	68	3	800	831
HY080G	4889	193	354	19	925	1002
HY090G	4990	547	74	10	62	191
HY110G	2342	136	72	5	499	560
HY120G	215	5	103	3	6835	7467
HY130G	3600	159	199	12	768	809
HY140G	8917	85	8678	83	6166	14563
<b>Net income components at personal level</b>						
PY010N	16475	158	6849	89	11243	13840
PY020N	4389	700	20	4	35	114
PY035N	1733	49	71	4	963	1220
PY050N	12391	348	1098	38	2081	3242
PY080N	10566	7329	2	1	4	6
PY090N	5449	148	208	9	1243	1347
PY100N	12313	74	3001	42	8174	8886
PY110N	7478	252	111	7	528	570
PY120N	1876	187	9	1	148	164
PY130N	7208	74	485	12	2285	2469
PY140N	2107	188	24	3	461	484
<b>Gross income components at personal level</b>						
PY010G	23153	228	9626	127	5792	13840
PY020G	4389	700	20	4	35	114
PY035G	1733	49	71	4	963	1220
PY050G	16780	464	1664	56	1852	3653
PY080G	12574	8721	2	2	1	6
PY090G	6239	178	238	10	765	1347
PY100G	14252	88	3473	49	5798	8886
PY110G	8570	302	127	8	328	570
PY120G	2165	230	10	1	92	164
PY130G	8174	87	550	14	1612	2469
PY140G	2107	188	24	3	461	484
PY200G	1899	18	722	10	30757	32388

The mean, the number of observations (before and after imputation) and the standard errors for the equivalised disposable income breakdown by sex, age groups and household size (mean and standard errors based on weighted data while the number of observations based on unweighted results).

Equivalised disposable income	Mean <sup>3</sup>	Standard error	Mean <sup>4</sup>	Standard error	Number of observations	
					Before imputation	After imputation
Subclasses by household size						
1 household member	14395	242	14372	242	2374	2726
2 household members	17622	234	17611	234	5938	7518
3 household members	17384	230	17378	230	6396	9297
4 and more	13197	128	13197	128	16225	25567
Population by age group						
<25	13336	114	13333	114	10163	15263
25 to 34	16410	207	16401	207	3891	5979
35 to 44	15480	231	15472	231	3752	5615
45 to 54	15435	193	15432	193	4810	7238
55 to 64	16021	197	16013	197	3501	4882
65+	14908	140	14908	140	4816	6131
Population by sex						
Male	14953	105	14946	105	14715	21754
Female	14863	97	14861	97	16218	23354

### *Longitudinal component*

The mean, the total number of observations (before and after imputation) and the standard errors for the following income components (mean and standard errors based on weighted data while the number of observations based on unweighted results).

Income components	Mean <sup>5</sup>	Standard error	Mean <sup>6</sup>	Standard error	Number of observations	
					Before imputation	After imputation
Total household gross income (HY010)	36078	332	36047	331	4610	10805
Total disposable household income (HY020)	27367	230	27358	230	7948	10810
Total disposable household income before social transfers other than old-age and survivors benefits (HY022)	25570	231	25134	230	7896	10646
Total disposable household income including old-age and survivors benefits (HY023)	20110	263	17713	248	7145	9649

<sup>3</sup> Taking into account only households/persons receiving such income.

<sup>4</sup> Taking into account whole population (households/persons) surveyed.

<sup>5</sup> Taking into account only households/persons receiving such income.

<sup>6</sup> Taking into account whole population (households/persons) surveyed.

Income components	Mean <sup>5</sup>	Standard error	Mean <sup>6</sup>	Standard error	Number of observations	
					Before imputation	After imputation
<b>Net income components at household level</b>						
HY040N	5528	510	78	10	116	148
HY050N	2087	53	423	14	2511	2561
HY060N	1429	73	72	5	556	570
HY070N	1315	44	72	4	578	606
HY080N	4587	200	313	20	673	724
HY090N	4450	562	62	10	73	131
HY110N	2308	147	69	6	386	394
HY120N	209	6	100	3	4934	5353
HY130N	3483	157	196	12	567	594
HY140N	8842	112	8617	109	4567	10550
HY145N	-499	36	-225	16	4612	4907
<b>Gross income components at household level</b>						
HY040G	6471	588	92	12	127	148
HY050G	2149	58	436	15	2421	2561
HY060G	1429	73	72	5	556	570
HY070G	1315	44	72	4	578	606
HY080G	4587	200	313	20	673	724
HY090G	5436	687	76	12	40	131
HY110G	2488	171	75	6	346	394
HY120G	209	6	100	3	4934	5353
HY130G	3483	157	196	12	567	594
HY140G	8599	113	8393	111	4550	10565
<b>Net income components at personal level</b>						
PY010N	16225	171	6337	93	8141	9966
PY020N	5175	913	22	5	24	81
PY035N	1712	55	66	4	686	877
PY050N	12457	366	1038	37	1497	2344
PY080N	2450	1002	0	0	3	4
PY090N	5475	166	210	10	924	999
PY100N	12328	77	3076	49	6026	6525
PY110N	7127	277	104	7	386	414
PY120N	1701	167	8	1	108	121
PY130N	7179	86	472	13	1676	1790
PY140N	1854	206	20	3	319	334
<b>Gross income components at personal level</b>						
PY010G	22749	247	8885	132	4204	9966
PY020G	5175	913	22	5	24	81
PY035G	1712	55	66	4	686	877
PY050G	16852	493	1570	55	1348	2636
PY080G	2915	1193	0	0		
PY090G	6266	199	241	11	563	999
PY100G	14267	92	3560	57	4298	6525
PY110G	8174	329	119	8	240	414
PY120G	1950	206	9	1	69	121
PY130G	8137	101	535	15	1189	1790
PY140G	1854	206	20	3	319	334

The mean, the number of observations (before and after imputation) and the standard errors for the equivalised disposable income breakdown by sex, age groups and household size (mean and standard errors based on weighted data while the number of observations based on unweighted results).

Equivalised disposable income	Mean <sup>7</sup>	Standard error	Mean <sup>8</sup>	Standard error	Number of observations	
					Before imputation	After imputation
Subclasses by household size						
1 household member	13937	258	13930	258	1730	1981
2 household members	17279	265	17271	265	4316	5418
3 household members	17174	278	17168	277	4608	6654
4 and more	13140	162	13140	162	11979	18742
Population by age group						
<25	13207	144	13206	144	7425	11085
25 to 34	15961	260	15950	260	2794	4290
35 to 44	15117	257	15117	257	2727	4051
45 to 54	15187	214	15187	214	3548	5312
55 to 64	16034	230	16024	229	2603	3588
65+	14921	158	14921	158	3536	4469
Population by sex						
Male	14756	131	14752	131	10732	15797
Female	14690	118	14688	118	11901	16998

<sup>7</sup> Taking into account only households/persons receiving such income.

<sup>8</sup> Taking into account whole population (households/persons) surveyed.

## 2.3. Non-sampling errors

### *Sampling frame and coverage errors*

The samples for EU-SILC 2005 and EU-SILC 2006 were selected from the sampling frame based on the TERYT system, i.e. the *Domestic Territorial Division Register*. Two kinds of primary sampling units (PSU) were distinguished in the sampling frame:

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

In the longitudinal (panel) component consisting of the subsamples 2, 3 and 4 some 7.6% of dwellings were found to be non-existing (cancelled, changed for non-residential units) as well as uninhabited or temporarily inhabited.

### *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.

According to the interviewers, who after the household and individual interview completion were obliged to answer a few questions concerning interview performance for EU-SILC 2005, over 70% of the respondents showed a favourable attitude towards the survey, while about 5% were unwilling towards it. In the interviewers' opinion, in about 86% of questionnaires (both household and individual ones) the quality of non-income data collected could be recognised as good or very good and in 1% - as doubtful.

For EU-SILC 2006 on the basis of this material it is possible to state that about three quarters of respondents (78% of those filling in the household questionnaire and 75% of those filling in the individual questionnaire) showed a favourable attitude towards the survey, while about 3% (both in the case of the household and individual interview) were unwilling towards it. In the interviewers' opinion, in about 88% of questionnaires (both household and individual ones) the quality of non-income data collected could be recognised as good or very good and in 1% - as doubtful.

The quality of income data in 2005 and 2006 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 EU-SILC was carried out in May/June 2005 and 2006.

During the years 2005 and 2006 the data collection was 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 were applicable.

The organisation and performance of the survey in the field was within the responsibility of regional statistical offices. Many interviewers were regular employees of the statistical offices having experience in other social surveys. Survey performance in the field was preceded by a series of trainings in 2005 and in 2006. Regional survey coordinators were instructed by CSO Social Statistics Division staff members and then the regional survey coordinators trained interviewers at the regional statistical offices.

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

The interviewers received written instructions concerning the survey performance.

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

Data recording and check-up took place in regional statistical offices and was done with the use of Microsoft Visual FoxPro. After all the questionnaires for a given household had been recorded (the identifiers being voivodship number, dwelling number and household number), it was possible to make the household screening which consisted of logical and calculation check-up at the section, inter-section and inter-questionnaire levels. The regional files were then transferred to the CSO Computing Centre and combined together to make up the general files at the national level. The national file completeness was also checked with the use of Microsoft Visual FoxPro. Additional check-up was made with SAS checking programmes. On the basis of overall data files it was possible to create files for Eurostat. Some of the primary target variables could be found directly in the questionnaires, others had to be calculated with the algorithms especially prepared for this purpose.

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

### *Non-response errors*

#### Achieved sample size

Sample size	Wave	
	1	2
A	12147	10813
B	29542	25441
C	36525	32801

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

B - number of person of 16 years or older who are members of the households for which the interview is accepted for database, and who completed a personal interview

C - number of selected person who are members of the households for which the interview is accepted for the database, and who completed a personal interview

Unit non-response

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

$Ra = 0.992$

$Rh = 0.702$

$NRh = 30.36$

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

$Rp = 0.950$

$NRp = 5.00$

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

$*NRp = 33.84$ .

Response rate for household:

- Wave response rate = 0.890

(percentage of households successfully interviewed (DB135 = 1) which were passed on to wave (from wave  $t-1$ ) or newly created or added during wave  $t$ , excluding those out of scope (under the tracing rules) or non-existent)

- Achieved samples size ratio = 0.890

(ratio of the number of households accepted for the database (DB135 = 1) in wave  $t$  to the number of households accepted for the database (DB135 = 1) in wave  $t-1$ )

Response rate for persons:

- Wave response rate of sample persons = 0.965

(percentage of sample person successfully interviewed (RB250 = 11, 12, 13) among those passed on to wave  $t$  (from wave  $t-1$ ) or newly created or added during wave  $t$ , excluding those out scope (under the tracing rules).

- Wave response rate of co-residents = 0.000

(percentage of co-residents selected in wave 1 successfully interviewed (RB250 = 11, 12, 13) among those passed on to wave  $t$  (from wave  $t-1$ ))

- Longitudinal follow-up rate = 0.956

(percentage of sample person successfully interviewed (RB250 = 11, 12, 13) in wave  $t$  out of all sample person selected, excluding those who have died or been found ineligible (out of scope), breakdown by causes of non-response).

- Achieved samples size ratio = 0.907

(ratio of the number of completed personal interviews (RB250 = 11, 12, 13) in wave  $t$  to the number of completed personal interviews in wave  $t-1$ . This ratio will be defined for sample persons and for all persons including non-sample persons aged 16+ and for co-residents aged 16+ selected in first wave).

- Response rate for non-sample persons = 0.496

(ratio of the number of completed personal interviews (RB250 = 11, 12, 13) of non-sample persons aged 16+ in wave  $t$  to all non-sample persons aged 16+ listed in the households accepted for the database (DB135 = 1) in wave  $t$  or listed in the most recently conducted household interviews for households, which were forwarded from wave  $t-1$  to wave  $t$  for follow-up, but could not be successfully interviewed in wave  $t$ ).

*Distribution of households by household status (DB110), by record of contact at address (DB120), by household questionnaire result (DB130) and by household interview acceptance (DB135)*

## Wave 1

### Household questionnaire result

<b>DB120=11</b>	<b>Total</b>	<b>%</b>
Total	17297	100.0
11 – household questionnaire completed,	12248	70.8
21 – refusal to co-operate,	3642	20.0
22 – entire household temporarily away for duration of fieldwork,	941	5.4
23 – household unable to respond (illness, incapacity,...)	485	2.8
24 – other reasons.	161	0.9

### Household interview acceptance

<b>DB135</b>	<b>Total</b>	<b>%</b>
Total	12248	100.0
1 – interview accepted for database	12147	99.2
2 – interview rejected	101	0.8

## Wave 2

### Household status

<b>DB110</b>	<b>Total</b>	<b>%</b>
Total	12262	100.0
1 – at the same address as last interview	11601	94.6
2 – entire household moved to a private household within the country	214	1.7
3 – entire household moved to a collective household or institution within the country	4	0.0
4 – household moved outside the country	42	0.3
5 – entire household died	70	0.6
6 – household does not contain sample person	13	0.1
7 – address non-contacted (unable to access, loss of information on record on what happened to the household-)	199	1.6
8 – split –off household	115	0.9
10 – fusion	4	0.0

### Record of contact at address

<b>DB120</b>	<b>Total</b>	<b>%</b>
Total	333	100.0
11 – address contacted	214	64.3
21 – address cannot be located	2	0.6
22 – address unable to access	0	0.0
23 – address does not exist or is non-residential address or is unoccupied or not principal residence	113	33.9
Missing	4	1.2

### Household questionnaire result

<b>DB130</b>	<b>Total</b>	<b>%</b>
Total	12121	100.0
11 – household questionnaire completed	10814	69.2
21 – refusal to co-operate	744	6.1
22 – entire household temporarily away for duration of fieldwork	159	1.3
23 – household unable to respond (illness, incapacity,...)	72	0.6
24 – other reasons	26	0.2
Missing	306	2.5

### Household interview acceptance

DB135	Total	%
Total	10814	100.0
1 – interview accepted for database ,	10813	100.0
2 – interview rejected	1	0.0

### Distribution of persons for membership status (RB110)

#### Distribution of persons for membership status (RB110)

	Total	Current household members				No current household members		
		RB110=1	RB110=2	RB110=3	RB110=4	RB120 = 2 to 4	RB120=6	RB120=7
Total	33701	31910	178	425	288	358	0	0
%	100.0	94.7	0.5	1.3	0.9	1.1	0.0	0.0

#### Distribution of persons moving out by variable RB120.

	Total	RB110 = 5				
		RB120 = 1		RB120 = 2	RB120 = 3	RB120 = 4
		A	B			
Total	691	166	167	108	209	41
%	100	24.0	24.2	15.6	30.2	6.0

A – this person is a current household number of a household this wave

B - this person is not a current household member

*Item non-response (income variables)*

2005

Item non-response	(A)	(B)	(C)
	% of households having received an amount	% of households with missing values	% of households with partial information
Total household gross income	39.66	6.50	53.34
Total disposable household income	67.65	5.71	26.34
Total disposable household income before social transfers other than old-age and survivor's benefits	67.32	8.50	22.18
Total disposable household income before social transfers, including old-age and survivor's benefits	60.62	11.30	16.50
<b>Net income components at household level</b>			
HY040N	1.20	0.10	0.00
HY050N	23.33	0.30	0.60
HY060N	4.22	0.13	0.00
HY070N	5.96	0.19	0.00
HY080N	5.20	0.62	0.00
HY090N	1.01	0.68	0.00
HY110N	2.29	0.12	0.00
HY120N	42.54	4.72	0.00
HY130N	5.02	0.29	0.00
HY140N	39.93	33.80	22.98
HY145N	44.87	3.53	0.03
<b>Gross income components at household level</b>			
HY040G	1.20	0.10	0.00
HY050G	22.85	0.30	1.08
HY060G	4.22	0.13	0.00
HY070G	5.96	0.19	0.00
HY080G	5.20	0.62	0.00
HY090G	0.54	0.68	0.48
HY110G	1.95	0.12	0.34
HY120G	42.54	4.72	0.00
HY130G	5.02	0.29	0.00
HY140G	40.01	33.75	23.50
HY040G	1.20	0.10	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	29.21	6.89	0.07
PY020N	0.10	0.39	0.00
PY035N	3.01	0.65	0.00
PY050N	4.62	3.13	0.20
PY080N	0.01	0.02	0.00
PY090N	3.73	0.24	0.02
PY100N	22.45	1.38	0.34
PY110N	1.36	0.09	0.00
PY120N	0.40	0.05	0.00
PY130N	6.40	0.27	0.02
PY140N	0.84	0.02	0.00
PY010N	29.21	6.89	0.07
<b>Gross income components at personal level</b>			
PY010G	15.86	6.85	13.46
PY020G	0.10	0.39	0.00
PY035G	3.01	0.65	0.00
PY050G	4.70	1.51	3.31
PY080G	0.00	0.02	0.01
PY090G	2.19	0.24	1.56
PY100G	15.28	1.36	7.52
PY110G	0.82	0.09	0.54
PY120G	0.07	0.05	0.34
PY130G	4.32	0.27	2.11
PY140G	0.84	0.02	0.00

2006

Item non-response	(A)	(B)	(C)
	% of households having received an amount	% of households with missing values	% of households with partial information
Total household gross income	42.63	5.03	52.26
Total disposable household income	73.50	4.49	21.98
Total disposable household income before social transfers other than old-age and survivor's benefits	73.02	6.01	19.42
Total disposable household income before social transfers, including old-age and survivor's benefits	66.07	8.22	14.95
<b>Net income components at household level</b>			
HY040N	1.07	0.19	0.10
HY050N	23.22	0.22	0.24
HY060N	5.14	0.08	0.05
HY070N	5.35	0.23	0.03
HY080N	6.22	0.47	0.00
HY090N	0.68	0.54	0.00
HY110N	3.57	0.07	0.00
HY120N	45.63	3.88	0.00
HY130N	5.24	0.25	0.00
HY140N	42.24	30.35	24.98
HY145N	42.65	2.73	0.00
<b>Gross income components at household level</b>			
HY040G	1.17	0.19	0.00
HY050G	22.39	0.22	1.07
HY060G	5.14	0.08	0.05
HY070G	5.35	0.23	0.03
HY080G	6.22	0.47	0.00
HY090G	0.37	0.54	0.31
HY110G	3.20	0.07	0.37
HY120G	45.63	3.88	0.00
HY130G	5.24	0.25	0.00
HY140G	42.08	30.20	25.42
HY040G	1.17	0.19	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	30.56	6.77	0.08
PY020N	0.09	0.21	0.00
PY035N	2.58	0.72	0.00
PY050N	5.62	2.85	0.33
PY080N	0.01	0.00	0.00
PY090N	3.47	0.28	0.00
PY100N	22.62	1.65	0.23
PY110N	1.45	0.11	0.00
PY120N	0.41	0.05	0.00
PY130N	6.29	0.41	0.02
PY140N	1.20	0.06	0.00
PY010N	30.56	6.77	0.08
<b>Gross income components at personal level</b>			
PY010G	15.78	6.77	14.86
PY020G	0.09	0.21	0.00
PY035G	2.58	0.72	0.00
PY050G	5.06	1.72	3.12
PY080G	0.00	0.00	0.01
PY090G	2.11	0.28	1.36
PY100G	16.13	1.65	6.71
PY110G	0.90	0.11	0.55
PY120G	0.26	0.05	0.15
PY130G	4.46	0.41	1.85
PY140G	1.20	0.06	0.00

#### **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 were applicable.

## Wave 1

### Distribution of household members by RB250

#### Household members 16+ (RB245 = 1 to 3)

	Total	RB250=11	RB250=21	RB250=23	RB250=31	RB250=32	RB250=33
Total	29542	28064	119	687	581	91	0
%	100.0	95.0	0.4	2.3	2.0	0.3	0.0

### Distribution of household members by RB260

#### Household members 16+ (RB245 = 1 to 3 and RB250 = 11 or 13)

	Total	RB260 = 1	RB260 = 2	RB260 = 3	RB260 = 4	RB260 = 5
Total	28064	22641	0	0	0	5423
%	100.0	80.7	0.0	0.0	0.0	19.3

## Wave 2

### Distribution of household members by RB250

#### Household members 16+ (RB245 = 1 to 3)

	Total	RB250=11	RB250=21	RB250=23	RB250=31	RB250=32	RB250=33
Total	26639	25441	81	602	453	60	2
%	100.0	95.5	0.3	2.3	1.7	0.2	0.0

#### Sample persons 16+ (RB2345 = 1 and RB100 = 1)

	Total	RB250=11	RB250=21	RB250=23	RB250=31	RB250=32	RB250=33
Total	26290	25128	80	583	442	55	2
%	100.0	95.6	0.3	2.2	1.7	0.2	0.0

#### Co-resident 16+ (RB245 = 1 to 3 and RB = 2)

	Total	RB250=11	RB250=21	RB250=23	RB250=31	RB250=32	RB250=33
Total	349	313	1	19	11	5	0
%	100.0	89.7	0.3	5.4	3.2	1.4	0.0

### Distribution of household members by RB260

#### **Household members 16+ (RB245 = 1 to 3) and RB245 = 11 or 13**

	<b>Total</b>	<b>RB260 = 1</b>	<b>RB260 = 2</b>	<b>RB260 = 3</b>	<b>RB260 = 4</b>	<b>RB260 = 5</b>
Total	25441	20632	0	0	0	4809
%	100.0	81.1	0	0	0	18.9

#### **Sample persons 16+ (RB245 = 1 to 3 and RB100 = 1) and RB250 = 11 or 13**

	<b>Total</b>	<b>RB260 = 1</b>	<b>RB260 = 2</b>	<b>RB260 = 3</b>	<b>RB260 = 4</b>	<b>RB260 = 5</b>
Total	25128	20408	0	0	0	4720
%	100.0	81.2	0.0	0.0	0.0	18.8

#### **Co-residents 16+ (RB245 = 1 to 3 and RB100 = 2) and RB250 = 11 or 13**

	<b>Total</b>	<b>RB260 = 1</b>	<b>RB260 = 2</b>	<b>RB260 = 3</b>	<b>RB260 = 4</b>	<b>RB260 = 5</b>
Total	313	224	0	0	0	89
%	100.0	71.6	0.0	0.0	0.0	28.4

As for individual interviews, in 2005 and in 2006 a relatively high share (19% and 18.6%) 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. Imputation procedures**

Imputation is aimed at obtaining complete records at the level of target variables. Thus the target variable level is the highest level of aggregation at which imputation can be made. This approach is applied wherever it does not cause loss of significant information from the file.

In the situation where:

- a target variable includes components of different character (e.g. taking different but highly predictable values, like benefits, or dependent on explanatory variables and thus easier to be modelled separately),

- there are many components of a target variable and it is often the case that in some of them there are missing data, while in others there are correct ones which could be lost during the imputation of the aggregated variable,

imputation is carried out at the level of particular components of target variables, frequently at the level of questionnaire variables. In some cases the target variables are identical with the questionnaire variables.

There are several methods of component imputation. They can be classified as deterministic and stochastic methods. In case of deterministic methods the method and the set of explanatory variables (algorithm) determines the imputation value for each record. In stochastic methods the imputation value is determined randomly so that with the same algorithm and the same data file each algorithm realisation may 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, allowing for the correct estimation of random error. Deterministic imputation causes variable variance reduction in the file and underestimation of random error; it also distorts the correlation structure (increasing correlations with explanatory variables). According to item 2.7 of Decision 1981/2003 it is recommended that for EU-SILC imputation the methods retaining distribution characteristics should be used, 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 categorising variables occur in the hot-deck method, a random representative is selected out of the records showing adequate values of auxiliary variables. If it proved impossible to find a donor of the equivalent values for all the auxiliary variables, the so called sequence approach was applied. The categorising variables were ranked from the most to the least significant ones. If there are no donors in the group, grouping is carried out with the subsequent explanatory variables left out, starting from the least significant ones so as to obtain a subset containing donors.

- Regression imputation with simulated residuals

Auxiliary variables are the explanatory variables of the regression model. The model takes a logarithmic form because of the income variable distribution. It is fitted on the basis of the correct records. The imputed value (its logarithm) is a sum of the theoretical value derived from the model and the pseudo-random number of the normal distribution with variance corresponding to the estimated variance of an error term in the model.

Out of the deterministic methods the following are applied:

- Regression deterministic imputation

A theoretical value from the model is taken 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 their most significant components, stochastic imputation is applied in order to retain the variable properties distribution as required by Decision 1981/2003.

The employment of regression imputation with simulated residuals 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 data set. That is why in the majority of cases, where in accordance with the assumption referred to above stochastic imputation is required, the hot-deck method is applied. 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. Regression imputation with randomly generated residuals is applied to incomes from hired employment, as:

- it is an important category of income, declared by a significant percentage of respondents and, if present, having 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 makes it possible to design a well-fitted model.

Deterministic imputation is applied where missing data concern less significant components of target variables (taxes, burdens to the main component, additions, etc.) and 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 availability 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, when expressing an economically important relationship or providing a grouping condition (interpretation criterion) in the calculation algorithm, have been retained, even if their effect on the imputed variable has not been statistically significant.

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

The imputation of the missing individual questionnaires is made with the use of hot-deck method.

The data on the donor's total income: gross, net and taxes as well as those on the sums of individual income components used for the calculation of the obligatory target variables at the household level are transferred to the taker's record.

## **2.6. Imputed rent**

For 2005 and 2006 imputed rents haven't been calculated.

## **2.7. Company cars**

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 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.

## **3. COMPARABILITY**

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

#### *The reference population*

There were no essential differences between the national concepts and standard EU-SILC concepts.

The survey unit was a household and all the household members who had completed 16 years of age by:

- December 31, 2004 for EU-SILC 2005;
- December 31, 2005 for EU-SILC 2006.

The survey did not cover collective accommodation households (such as boarding house, workers' hostel, pensioners' house or monastery), 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.

#### *The private household definition*

No difference to the common definition in either wave (EU-SILC 2005 and EU-SILC 2006).

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.

### *The household membership*

No difference to the common definition in either wave (EU-SILC 2005 and EU-SILC 2006).

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.

### *The income reference period(s) used*

No difference to the common definition in either wave. The income reference year for EU-SILC 2005 was 2004 and for EU-SILC 2006 the year 2005.

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

No difference to the common definition in either wave (EU-SILC 2005 and EU-SILC 2006).

The reference period for income tax prepayment and compulsory social insurance contributions were again years 2004, 2005. The account clearance with the Treasury Office (including payments and returns) effected in 2005 refers to the income for 2004, and in 2006 for 2005.

### *The reference period for taxes on wealth*

No difference to the common definition in either wave (EU-SILC 2005 and EU-SILC 2006). Taxes on wealth paid during the income reference period were recorded, properly 2004 year or 2005.

### *The lag between the income reference period and current variables*

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

### *The total duration of the data collection of the sample*

EU-SILC was performed on the territory of the whole country in 2005 year between May 2 and June 17, in 2006 between May 2 and June 19.

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

#### Differences concerning EU-SILC 2005:

Variables were not recorded:

- Change of job since last year (PL160);
- Reason for change (PL170);
- Most recent change in the individual's activity status (PL180).

Starting from EU-SILC 2006, they have been taken into account.

#### Differences concerning EU-SILC 2006:

In EU-SILC 2006 the definition of retired person (now this definition is in accordance with international recommendations) was changed. In EU-SILC 2005 people obtaining disability pensions were included in PL085 and in variable PL210 in category 6 while in EU-SILC 2006 they are included in variable PL090 and in variable PL210 in category 8.

## **3.2. Components of income**

### *Differences between the national definitions and standards EU-SILC definitions, and an assessment:*

Income components where no difference between national and standard definitions can be found are not mentioned. The differences between the national and the EUROSTAT definitions refer to two waves likewise unless it is indicated.

#### *PY010 - Cash or near-cash employee income*

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),
- benefit granted to the families when the only person providing income for the family is called up to the active military service; since this benefit is only granted when the only family supporter has been called to the military service, it has been included in the category of „Family related allowances' (HY050).

*PY020 - Non-cash employee income*

The information collected only refers to the income gained from the use of the company car for private purposes.

*PY030 – Employer’s social insurance contributions*

Variables was collected since EU-SILC 2006.

*PY050 - Cash profits or losses from self-employment (including royalties)*

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.

For EU-SILC 2006 (the income reference year 2005):

All family benefits was collected for the whole year, with except advance payment of alimony (was collected for the maximum of 4 months).

*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.

*The form in which income variables at 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.

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

The gross income was obtained by summing up net value, 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.

### **3.3. Tracing rules**

Standard EU-SILC tracing rules are applied.

## 4. COHERENCE

The calculation in point 4.1 and 4.2 was made taking into account the cross-sectional data of EU-SILC 2006.

### 4.1. 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 the HBS the main source of data is the so called diary. Two additional questionnaires are also 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 households which refused to participate in the survey are replaced with those from the so called reserve list. No replacement is applied in EU-SILC;
- Slightly different weighting of results.

In some tables below socio-economic groups' breakdown is used. The household survey results are traditionally prepared by CSO according to the so called socio-economic groups of households. 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 2006	HBS 2006
	in %	
Total	100.0	100.0
0-14	16.5	18.1
15-24	15.7	16.6
25-54	44.1	41.6
55-64	10.5	11.3
65+	13.3	12.4

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

Specification	EU-SILC 2006	HBS 2006
	in %	
Total	100.0	100.0
No school education	2.5	0.9
Completed primary	19.2	20.0
Lower secondary	4.9	6.5
Elementary vocational	26.6	26.8
Secondary	33.5	33.5
Higher	13.2	12.2

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

Households	Households		Persons in households	
	EU-SILC 2006	HBS 2005	EU-SILC 2006	HBS 2005
Total	13318760	13332605	37794851	37744302
<b>Total = 100</b>				
Employees	47.9	43.1	57.4	51.3
Farmers	2.5	4.7	3.7	7.0
Self-employed	5.0	5.9	5.6	7.0
Retirees	27.8	27.0	19.7	18.8
Pensioners	9.8	11.8	7.1	8.8
Maintained from non-earned sources	6.9	7.5	6.5	7.0

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

Households	Disposable income		Income from hired work	
	EU-SILC 2006	HBS 2005	EU-SILC 2006	HBS 2005
Total	14902	13444	8701	6636
Employees	16872	14325	14117	11695
Farmers	9224	12229	689	1216
Self-employed	18271	18271	2970	2565
Retirees	13700	13309	1270	1266
Pensioners	9649	9622	1004	902
Maintained from non-earned sources	7213	8552	1387	727

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

Households	Disposable income		Income from hired work	
	EU-SILC 2006	HBS 2005	EU-SILC 2006	HBS 2005
Total	14902	13444	8701	6636
1-person	14372	13704	4439	3673
2-persons	17611	15961	7247	5589
3-persons	17378	14954	11691	8839
4-persons	14928	13582	10713	8382
5-persons	12349	11318	7569	5888
6-persons and more	10959	9964	6113	3970

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

Specification	EU-SILC 2006	HBS 2006
	in %	
Fixed telephone	74.0	71.9
Mobile telephone	70.7	73.1
Television set	96.9	98.5
Computer	44.4	43.7
Printer	31.8	29.5
Internet connection	28.5	28.4
Microwave oven	34.1	38.0
Dishwasher	8.0	6.2
Refrigerator	96.9	98.6
Washing machine	96.2	96.8
Passenger car	50.7	49.5

#### **4.2. Comparison of 2005 results of SNA and EU-SILC 2006 for Poland**

Comparison of SNA data for the household sector and EU-SILC 2006 (data for 2005) in the area of incomes

The comparison covered the disposable income and its main components: income from hired employment, self-employment (in and outside farming) and social benefits.

It was confirmed that in EU-SILC 2006 the disposable income was 57% of the respective category in SNA. This has been brought about by the following reasons:

1. The household sector in SNA includes collective households which are not covered by EU-SILC.
2. Each of the systems applies a different method of measuring income from self-employment.
3. The primary and secondary income distribution in SNA used as a basis for the calculation of disposable income includes some items which were not present in EU-SILC 2006, the most important one being imputed rents.

In SNA income from self-employment is calculated as the so called operation surplus which is a balance between global production and current production inputs, i.e. intermediate consumption and hired employees' remunerations. This amount is reduced by taxes and increased by subsidies. The operation surplus calculated in this way is allocated to households' consumption needs as well as dwelling- and business-related investment. In the Polish EU-SILC the question about income from self-employment refers only to the amount spent on household's consumption and its dwelling-related investment. Besides, SNA takes into account consumption from own production, which was not covered by EU-SILC 2006 for farming. These differences are responsible for the fact that income from self-employment in EU-SILC 2006 amounted only to 27% of the operation surplus in SNA (after section K deduction).

The income from self-employment in EU-SILC 2006 is equal to 96% of the respective category in SNA, while social benefits – 92%, respectively, which seems to be a satisfactory outcome.

As compared with EU-SILC 2005, the coherence between EU-SILC 2006 and SNA data improved: in case of disposable income by 3 percentage points, for income from hired employment – by 6 p.p. and for income from self-employment and social benefits - by 3 p.p. This improvement was possible due to better quality of data, particularly the interviewers' work.

<i>Category in SNA</i>	<i>Variables in EU-SILC 2006</i>	<i>Category description in EU-SILC 2006</i>	<i>SNA in mln PLN</i>	<i>EU-SILC in mln PLN</i>	<i>SNA = 100%</i>	<i>SNA =100% EU-SILC 2005</i>
Gross disposable income (net)	HY020	Total disposable household income (net)	651 512	369 046	57	54
Wages, salaries and other income connected with hired work (gross)	PY010G	Employee cash or near cash income (gross)	303 358	290 140	96	90
Gross operating surplus (gross) with the exception of section K	PY050G	Self-employment income (gross) - value allocated to household's consumption and dwelling-related investment	189 378	50 167	27	24
Social security benefits and social assistance benefits (gross)	PY90G + PY100G + PY110G + PY120G + PY130G + PY140G + HY050G + HY060G + HY070G	Social benefits (gross)	153 946	141 334	92	89

Remarks:

1. Remarks in brackets: "net" or "gross" refer to including or not including income tax and social security contributions while the word "gross" in SNA names of categories refer to including of depreciation of fixed assets.
2. Data for gross operating surplus in SNA has been taken into consideration with the exception of section K what allows for better comparability with EU-SILC data on self-employment income (PY050G). The data for section K includes mainly imputed rents, not calculated in EU-SILC 2006 (data for 2005), and market income from renting of real estate included in EU-SILC as the variable HY040G.