INTERMEDIATE QUALITY REPORT

EU-SILC 2007 OPERATION POLAND

Warsaw, November 2008

CONTENTS Page

PREFACE

The present quality report is the intermediate quality report of EU-SILC 2007 in Poland according to grant agreement No. 36401.2006.001-2006.199 and follows the structure outlined in the Commission Regulation No. 1177/2003.

The mentioned regulation consists of four chapters.

The first chapter describes the common cross-sectional indicators and other indicators of interest computed on the basis of EU-SILC 2007.

The second chapter deals with accuracy i.e. here should be described all the factors that affect the precision of estimations and results.

The third chapter reports on comparability and describes all differences between the standard EU definitions and the definitions applied in the survey in Poland.

The fourth and last chapter, reporting on coherence, presents the comparisons of the EU-SILC 2007 data with external sources.

As this is the third intermediate quality report for EU-SILC in Poland some chapters and sections resemble the corresponding chapters and sections of the preceding reports.

1. COMMON CROSS-SECTIONAL EUROPEAN UNION INDICATORS

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

	Indicator	Value
1	At-risk-of-poverty rate after social transfers - total	**17**
2	At-risk-of-poverty rate after social transfers - men total	**18**
3	At-risk-of-poverty rate after social transfers - women total	**17**
4	At-risk-of-poverty rate after social transfers - 0-17 years	**24**
5	At-risk-of-poverty rate after social transfers - 18-64 years	**17**
6	At-risk-of-poverty rate after social transfers - 65+ years	**8**
7	At-risk-of-poverty rate after social transfers - men 18-64 years	**18**
8	At-risk-of-poverty rate after social transfers - men 65+ years	**6**
9	At-risk-of-poverty rate after social transfers - women 18-64 years	**17**
10	At-risk-of-poverty rate after social transfers - women 65+ years	**9**
11	At-risk-of-poverty threshold - single	**8187 PLN**
12	At-risk-of-poverty threshold - 2 adults, 2 children	**17192 PLN**
13	Relative median at-risk-of-poverty gap - total	**24**
14	Relative median at-risk-of-poverty gap - men total	**25**
15	Relative median at-risk-of-poverty gap - women total	**23**
16	Relative median at-risk-of-poverty gap - 0-17 years	**26**
17	Relative median at-risk-of-poverty gap - 18-64 years	**25**
18	Relative median at-risk-of-poverty gap - 65+ years	**14**
19	Relative median at-risk-of-poverty gap - men, 18-64 years	**25**
20	Relative median at-risk-of-poverty gap - men, 65+ years	**15**
21	Relative median at-risk-of-poverty gap - women, 18-64 years	**24**
22	Relative median at-risk-of-poverty gap - women, 65+ years	**14**
23	Inequality of income distribution S80/S20 income quintile share ratio	**5.26**
24	In work at-risk-of-poverty rate - total	**12**
25	In work at-risk-of-poverty rate - men total	**13**
26	In work at-risk-of-poverty rate - women total	**10**
27	Relative median income ratio people aged 65+/0-64	**1.04**
28	Relative median income ratio people aged 65+/0-64 - men	**1.12**
29	Relative median income ratio people aged 65+/0-64 - women	**0.99**
30	Aggregate replacement ratio pensions 65-74/earnings 50-59	**0.57**
31	Aggregate replacement ratio pensions 65-74/earnings 50-59 - men	**0.63**
32	Aggregate replacement ratio pensions 65-74/earnings 50-59 - women	**0.56**
	Before social transfers except old-age and survivors' benefits	
33	At-risk-of-poverty rate before social transfers - total	**27**
34	At-risk-of-poverty rate before social transfers - men total	**27**
35	At-risk-of-poverty rate before social transfers - women total	**26**
36	At-risk-of-poverty rate before social transfers - 0-17 years	**35**
37	At-risk-of-poverty rate before social transfers - 18-64 years	**27**
38	At-risk-of-poverty rate before social transfers - 65+ years	**12**
39	At-risk-of-poverty rate before social transfers - men, 18-64 years	**28**
40	At-risk-of-poverty rate before social transfers - men, 65+ years	**9**
41	At-risk-of-poverty rate before social transfers - women, 18-64 years	**26**
42	At-risk-of-poverty rate before social transfers - women, 65+ years	**13**

	Indicator	Value
	Before social transfers including old-age and survivors' benefits	
43	At-risk-of-poverty rate before social transfers - total	**47**
44	At-risk-of-poverty rate before social transfers - men total	**45**
45	At-risk-of-poverty rate before social transfers - women total	**49**
46	At-risk-of-poverty rate before social transfers - 0-17 years	**41**
47	At-risk-of-poverty rate before social transfers - 18-64 years	**41**
48	At-risk-of-poverty rate before social transfers - 65+ years	**87**
49	At-risk-of-poverty rate before social transfers - men, 18-64 years	**40**
50	At-risk-of-poverty rate before social transfers - men, 65+ years	**87**
51	At-risk-of-poverty rate before social transfers - women, 18-64 years	**42**
52	At-risk-of-poverty rate before social transfers - women, 65+ years	**87**
53	Mean equivalised disposable income	**16166 PLN**

2. ACCURACY

2.1. Sample design

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.

Sampling units

The first-stage sampling units (primary sampling units - PSU) 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.

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 poviats (NUTS4). Altogether 211 strata were distinguished.

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 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 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.

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 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.

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.

Weightings

<u>*Design factor*</u>

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 a PSU in the h-th stratum,
M_h – number of dwellings in the h-th stratum.

<u>*Non-response adjustments*</u>

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

Code of class of locality (p)	Class of locality	Completeness rate ($cr_p = R_{ap} * Rh_p$)
	Poland	0.660
1	Warsaw	0.439
2	Towns 500 000 – 1 000 000 inhabitants	0.506
3	Towns 100 000 – 500 000 inhabitants	0.594
4	Towns 20 000 – 100 000 inhabitants	0.665
5	Towns less than 20 000 inhabitants	0.703
6	Rural areas	0.775

The adjusted weights were calculated according to the formula:

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

Weights DB080 and DB080^{corrected} were calculated for the subsample 6. The next step consisted in calculating the weights DB090 and RB050 for the households of the subsample 6 with the use of the integrated calibration method. For the subsamples 3 and 4 surveyed for the third time and the subsample 5 surveyed for the second time the base weights were determined by the correction of the base weights from the previous year.

For the subsample 5 the following method was used:
The base weight of 2006 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 2006 and who should be surveyed in 2007. 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 2006 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 3 and 4 (surveyed for the third time) algorithm based on method described for the subsample 5 was used. Additionally, re-entries occurrence was taken into account i.e. persons who were surveyed in 2005, not surveyed in 2006, and again surveyed in 2007 year. The base weights for such persons were computed by correction of base weights from year

2005 on data for years 2005 and 2007 (without information from 2006 year). Inclusion of re-entries to the subsamples surveyed in 2007 year 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 2005. Computed coefficients are shown in the following table:

class of locality	Correction for subsample 3	Correction for subsample 4
1	0.927	0.918
2	0.976	0.970
3	0.976	0.977
4	0.983	0.976
5	0.981	0.989
6	0.990	0.986

The last stage of the base weight calculation for the third year of the survey consisted in receiving weights of mothers by children of "sample persons" and zero weights by "coresidents' 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 in combining the four independent subsamples, applying the integrated calibration as described below (for the sample 6 repeatedly) and trimming. As a result, DB090 and RB050 weights are obtained for households and individuals from the samples 3, 4, 5 and 6.

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.

8

Final cross-sectional weight

In EU-SILC 2007 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 is obtained by the adjustment of RB050 separately in the groups according to gender and age in each voivodship broken up 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 a resampling approach. We used a bootstrap method which resamples 200 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.

Indicator	Value	Standard error	Achieved sample size	Design effect	Effective sample size
At-risk-of-poverty rate after social transfers - total	17.34	0.42	42852	4.45	9636
At-risk-of-poverty rate after social transfers - men total	17.58	0.47	20553	2.62	7848
At-risk-of-poverty rate after social transfers - women total	17.11	0.44	22299	2.45	9104

At-risk-of-poverty rate after social transfers - 0-17 years	24.18	0.76	9331	2.36	3951
At-risk-of-poverty rate after social transfers - 18-64 years	17.18	0.44	27560	2.89	9523
At-risk-of-poverty rate after social transfers - men 18-64 years	17.81	0.48	13437	1.89	7116
At-risk-of-poverty rate after social transfers - women 18-64 years	16.57	0.45	14123	1.60	8814
At-risk-of-poverty rate after social transfers - 65+ years	7.83	0.43	5961	1.31	4558
At-risk-of-poverty rate after social transfers - men 65+ years	5.56	0.58	2376	1.10	2152
At-risk-of-poverty rate after social transfers - women 65+ years	9.19	0.56	3585	1.19	3024
At-risk-of-poverty threshold - single	8187	66	42852	3.78	11351
At-risk-of-poverty threshold - 2 adults, 2 children	17192	138	42852	3.78	11351
Relative median at-risk-of-poverty gap – total	23.97	0.90	42852	3.40	12586
Relative median at-risk-of-poverty gap – men total	25.43	0.99	20553	2.94	6989
Relative median at-risk-of-poverty gap - women total	22.81	0.70	22299	2.72	8212
Relative median at-risk-of-poverty gap – 0-17 years	26.18	1.00	9331	2.96	3151
Relative median at-risk-of-poverty gap – 18-64 years	24.64	1.04	27560	2.70	10225
Relative median at-risk-of-poverty gap - men, 18-64 years	25.29	1.11	13437	1.74	7701
Relative median at-risk-of-poverty gap - women, 18-64 years	23.78	0.99	14123	1.70	8309
Relative median at-risk-of-poverty gap – 65+ years	14.45	0.69	5961	1.16	5118
Relative median at-risk-of-poverty gap - men, 65+ years	15.31	1.61	2376	0.98	2416
Relative median at-risk-of-poverty gap - women, 65+ years	14.38	0.71	3585	1.10	3253
Inequality of income distribution S80/S20 income quintile share ratio	5.26	0.09	42852	2.96	14477
In work at-risk-of-poverty rate - total	11.68	0.38	14552	1.56	9337
In work at-risk-of-poverty rate - men total	12.64	0.45	7850	1.09	7220
In work at-risk-of-poverty rate - women total	10.48	0.43	6702	1.12	6000
Relative median income ratio people aged 65+/0-64	1.04	0.01	42852	1.76	24291
Relative median income ratio people aged 65+/0-64 - men	1.12	0.02	20553	1.29	15930
Relative median income ratio people aged 65+/0-64 - women	0.99	0.01	22299	1.32	16944
Aggregate replacement ratio pensions 65-74/earnings 50-59	0.57	0.01	5523	1.24	4449
Aggregate replacement ratio pensions 65-74/earnings 50-59 - men	0.63	0.02	2750	1.00	2739
Aggregate replacement ratio pensions 65-74/earnings 50-59 - women	0.56	0.02	2773	1.10	2524

Before social transfers except old-age and survivors' benefits

At-risk-of-poverty rate before social transfers - total	26.52	0.50	42852	4.93	8692
At-risk-of-poverty rate before social transfers - men total	26.97	0.56	20553	2.60	7917
At-risk-of-poverty rate before social transfers - women total	26.10	0.51	22299	2.34	9513
At-risk-of-poverty rate before social transfers - 0-17 years	34.51	0.83	9331	2.19	4267
At-risk-of-poverty rate before social transfers - 18-64 years	27.11	0.54	27560	3.56	7742
At-risk-of-poverty rate before social transfers - men, 18-64 years	27.79	0.59	13437	1.91	7030
At-risk-of-poverty rate before social transfers - women, 18-64 years	26.45	0.56	14123	1.73	8145
At-risk-of-poverty rate before social transfers - 65+ years	11.57	0.51	5961	1.33	4487
At-risk-of-poverty rate before social transfers - men, 65+ years	9.16	0.73	2376	1.10	2162
At-risk-of-poverty rate before social transfers - women, 65+ years	13.02	0.62	3585	1.10	3255

Before social transfers including old-age and survivors' benefits

At-risk-of-poverty rate before social transfers - total	47.11	0.56	42852	4.22	10145
At-risk-of-poverty rate before social transfers - men total	44.90	0.60	20553	2.39	8599
At-risk-of-poverty rate before social transfers - women total	49.17	0.59	22299	2.38	9375
At-risk-of-poverty rate before social transfers – 0-17 years	41.30	0.86	9331	2.00	4666
At-risk-of-poverty rate before social transfers - 18-64 years	40.79	0.60	27560	3.26	8456
At-risk-of-poverty rate before social transfers - men, 18-64 years	39.86	0.64	13437	1.92	7014
At-risk-of-poverty rate before social transfers - women, 18-64 years	41.70	0.64	14123	1.86	7574
At-risk-of-poverty rate before social transfers - 65+ years	87.03	0.57	5961	1.40	4252
At-risk-of-poverty rate before social transfers - men, 65+ years	87.32	0.85	2376	1.17	2033
At-risk-of-poverty rate before social transfers - women, 65+ years	86.86	0.65	3585	1.10	3260
Mean equivalised disposable income	16166	129.13	42852	3.55	12071

2.3. Non-sampling errors

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 (PSU) 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) as well as 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.

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 (80% of those filling in the household questionnaire and 78% 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 89% 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 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 2007 was carried out in May/June.

EU-SILC, as it was in 2005 and 2006, 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.

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. Survey performance in the field was preceded by a series of trainings. 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. 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-up was made with SAS checking programmes.

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

Non-response errors

Achieved sample size

Sample size	Rotational group				
	3	4	5	6	**Total**
A	3403	3421	3632	3830	14286
B	7860	7939	8396	8606	32801
C	10224	10249	10978	11401	42852

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 $NRh = [1 - (Ra*Rh)]*100$,

$Ra = 0.995$
$Rh = 0.837$

Ra – the address contact rate
Rh – the proportion of complete household interviews accepted for the database

$NRh = 16.72$

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

$Rp = 0.94$
$NRp = 5.98$

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

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

$*NRp = 21.7$

Information on non-response	Rotational group				
	3	4	5	6	**Total**
Ra	1.000	0.999	0.999	0.985	0.995
Rh	0.878	0.886	0.907	0.720	0.837
NRh	12.20	11.49	9.39	29.08	16.72
Rp	0.943	0.944	0.942	0.931	0.940
NRp	5.65	5.57	5.76	6.87	5.98
*NRp	17.12	16.39	14.62	33.93	21.70

Distribution of households

- DB120 - Contact at address

DB120	Rotational group				
	3	4	5	6	**Total**
Address contacted (11)	3874	3863	4006	5318	17061
Address cannot be located (21)	0	1	3	77	81
Address impossible to access (22)	0	0	0	1	1
Address does not exist or is non-residential or is unoccupied or not the principal residence (23)	24	26	26	800	876
Total	3898	3890	4035	6196	18019

- DB130 - Household questionnaire result

DB130	Rotational group				
	3	4	5	6	**Total**
Household questionnaire completed (11)	3403	3421	3632	3834	14290
Refusal to co-operate (21)	349	310	234	1171	2064
Entire household temporarily away for duration of fieldwork (22)	82	82	86	166	416
Household unable to respond (illness, incapacity,…) (23)	28	38	33	114	213
Other reasons (24)	12	12	21	33	78
Total	3874	3863	4006	5318	17061

- DB135 - Household interview acceptance

DB135	Rotational group				
	3	4	5	6	**Total**
Interview accepted for database (1)	3403	3421	3632	3830	14286
Interview rejected (2)	0	0	0	4	4
Total	3403	3421	3632	3834	14290

Item non-response (income variables)

Item non-response	(A) % of households having received an amount	(B) % of households with missing values	(C) % of households with partial information
Total household gross income	36.21	5.37	58.31
Total disposable household income	73.24	5.26	21.42
Total disposable household income before social transfers other than old-age and survivor's benefits	73.06	7.36	18.21
Total disposable household income before social transfers, including old-age and survivor's benefits	65.15	9.99	13.87
Net income components at household level			
HY040N	0.87	0.21	0.17
HY050N	21.87	0.31	0.39
HY060N	5.02	0.11	0.06
HY070N	4.84	0.12	0.00
HY080N	5.75	0.60	0.00
HY090N	1.21	0.84	0.00
HY100N	0.92	1.78	0.00
HY110N	3.91	0.15	0.01
HY120N	49.09	4.90	0.00
HY130N	4.34	0.29	0.00
HY140N	35.11	37.72	25.09
HY145N	35.72	2.53	0.03
Gross income components at household level			
HY040G	1.04	0.21	0.00
HY050G	20.84	0.31	1.43
HY060G	5.02	0.11	0.06
HY070G	4.84	0.12	0.00
HY080G	5.75	0.60	0.00
HY090G	0.54	0.83	0.68
HY100G	0.92	1.78	0.00
HY110G	3.52	0.15	0.40
HY120G	49.09	4.90	0.00
HY130G	4.34	0.29	0.00
HY140G	34.97	37.22	25.82

	% of persons 16+ having received an amount	% of persons 16+ with missing values	% of persons 16+ with partial information
Net income components at personal level			
PY010N	31.23	7.99	0.07
PY020N	7.37	3.02	0.93
PY021N	0.19	0.21	0.00
PY035N	2.39	0.67	0.00
PY050N	5.79	2.82	0.30
PY070N	5.95	1.47	0.00
PY080N	0.02	0.00	0.00
PY090N	2.58	0.41	0.03
PY100N	23.28	1.85	0.22
PY110N	1.23	0.17	0.00
PY120N	0.37	0.05	0.00
PY130N	6.06	0.55	0.02
PY140N	1.33	0.13	0.00
	% of persons 16+ having received an amount	% of persons 16+ with missing values	% of persons 16+ with partial information
Gross income components at personal level			
PY010G	15.42	7.99	15.89
PY020G	7.37	3.02	0.93
PY021G	0.19	0.21	0.00
PY030G	2.61	19.60	0.00
PY035G	2.30	0.67	0.00
PY050G	5.68	2.07	2.83
PY070G	5.95	1.47	0.00
PY080G	0.01	0.00	0.01
PY090G	1.24	0.41	1.36
PY100G	14.01	1.85	9.49
PY110G	0.52	0.17	0.71
PY120G	0.19	0.05	0.17
PY130G	3.10	0.55	2.97
PY140G	1.33	0.13	0.00
PY200G	28.50	7.69	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 2006 (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 2006, explanatory variables derived from the current data file are applied. Wherever the same type of income is found in the data for 2006, 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 the common cross-sectional European Union 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	42852	14236
At-risk-of-poverty rate after social transfers - men total	20553	7032
At-risk-of-poverty rate after social transfers - women total	22299	7204
At-risk-of-poverty rate after social transfers - 0-17 years	9331	3191
At-risk-of-poverty rate after social transfers - 18-64 years	27560	9832
At-risk-of-poverty rate after social transfers - men 18-64 years	13437	4905
At-risk-of-poverty rate after social transfers - women 18-64 years	14123	4927
At-risk-of-poverty rate after social transfers - 65+ years	5961	1213
At-risk-of-poverty rate after social transfers - men 65+ years	2376	488
At-risk-of-poverty rate after social transfers - women 65+ years	3585	725
At-risk-of-poverty threshold - single	42852	14236
At-risk-of-poverty threshold - 2 adults, 2 children	42852	14236
Relative median at-risk-of-poverty gap - total	42852	14236
Relative median at-risk-of-poverty gap - men total	20553	7032
Relative median at-risk-of-poverty gap - women total	22299	7204
Relative median at-risk-of-poverty gap - 0-17 years	9331	3191
Relative median at-risk-of-poverty gap - 18-64 years	27560	9832
Relative median at-risk-of-poverty gap - men, 18-64 years	13437	4905
Relative median at-risk-of-poverty gap - women, 18-64 years	14123	4927
Relative median at-risk-of-poverty gap - 65+ years	5961	1213
Relative median at-risk-of-poverty gap - men, 65+ years	2376	488
Relative median at-risk-of-poverty gap - women, 65+ years	3585	725
Inequality of income distribution S80/S20 income quintile share ratio	42852	14236

In work at-risk-of-poverty rate - total	14552	4904
In work at-risk-of-poverty rate - men total	7850	2569
In work at-risk-of-poverty rate - women total	6702	2335
Relative median income ratio people aged 65+/0-64	42852	14236
Relative median income ratio people aged 65+/0-64 - men	20553	7032
Relative median income ratio people aged 65+/0-64 - women	22299	7204
Aggregate replacement ratio pensions 65-74/earnings 50-59	5523	2450
Aggregate replacement ratio pensions 65-74/earnings 50-59 - men	2750	1255
Aggregate replacement ratio pensions 65-74/earnings 50-59 - women	2773	1195
Before social transfers except old-age and survivors' benefits		
At-risk-of-poverty rate before social transfers - total	42852	13881
At-risk-of-poverty rate before social transfers - men total	20553	6852
At-risk-of-poverty rate before social transfers - women total	22299	7029
At-risk-of-poverty rate before social transfers - 0-17 years	9331	3083
At-risk-of-poverty rate before social transfers - 18-64 years	27560	9598
At-risk-of-poverty rate before social transfers - men, 18-64 years	13437	4785
At-risk-of-poverty rate before social transfers - women, 18-64 years	14123	4813
At-risk-of-poverty rate before social transfers - 65+ years	5961	1200
At-risk-of-poverty rate before social transfers - men, 65+ years	2376	486
At-risk-of-poverty rate before social transfers - women, 65+ years	3585	714
Before social transfers including old-age and survivors' benefits		
At-risk-of-poverty rate before social transfers - total	42852	13430
At-risk-of-poverty rate before social transfers - men total	20553	6661
At-risk-of-poverty rate before social transfers - women total	22299	6769
At-risk-of-poverty rate before social transfers - 0-17 years	9331	3051
At-risk-of-poverty rate before social transfers - 18-64 years	27560	9413
At-risk-of-poverty rate before social transfers - men, 18-64 years	13437	4695
At-risk-of-poverty rate before social transfers - women, 18-64 years	14123	4718
At-risk-of-poverty rate before social transfers - 65+ years	5961	966
At-risk-of-poverty rate before social transfers - men, 65+ years	2376	400
At-risk-of-poverty rate before social transfers - women, 65+ years	3585	566
Mean equivalised disposable income	42852	14236

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				
	3	4	5	6	**Total**
Information completed only from interview (11)	7860	7939	8396	8606	32801
Individual unable to respond (illness, incapacity, etc) (21)	30	31	23	35	119
Refusal to co-operate (23)	210	233	261	275	979
Person temporarily away and no proxy possible (31)	201	162	183	234	780
No contact for another reasons (32)	30	42	44	91	207
Information not completed: reason unknown (33)	0	0	2	0	2
Total	8331	8407	8909	9241	34888

- RB260 – Type of interview

RB260	Rotational group				
	3	4	5	6	**Total**
Face to face (1)	6557	6553	6922	7134	27166
Proxy interview (2)	1303	1386	1474	1472	5635
Total	7860	7939	8396	8606	32801

As for individual interviews, in 2007 a relatively high share (17,2%) 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 34 minutes, while the average individual interview duration was about 24 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 88 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 and 2006.

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, 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

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

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

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

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

There were no differences between the national concepts and standard EU-SILC concepts.
The income reference period was last calendar year (2006).

The period for taxes on income and social insurance contributions

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

The reference period for taxes on wealth

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

Taxes on wealth paid during the income reference period (2006) were recorded.

The lag between the income reference period and current variables

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

The total duration of the data collection of the sample

EU-SILC was performed on the territory of the whole country between May 2 and June 19 2007.

Basic information on activity status during the income reference period

Since 2006 there were no differences between the national concepts and standard EU-SILC concepts.

3.2. Components of income

Differences between the national definitions and standard EU-SILC definitions, and an assessment:

Variables collected since 2007 but not included in the household income (according to recommendations Eurostat):

− imputed rent (variable has been calculated in foothold about econometric model);
− interest paid on mortgages;
− non-cash employee income;
− company car – 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 estimated the amount gained) imputation is applied with the use of hot-deck and regression imputation with simulated residuals methods;
− value of goods produced for own consumption;
− employers' social insurance contributions.

Pension from individual private plans
This variable has been included in the income of household. We have received information of exclusion after transmission.

Cash or near-cash 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).

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.
It is performed detailed analysis of this data in each wave of research and it decides about choice of fairest method. It employ second method in 2007 year also.

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

Sickness benefits
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.

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 065/04:

Capacity to afford paying for one week annual holiday away from home – the question included the expression "if the household wants"

Leaking roof, damp walls/ floors/foundation, or rot in window frames or floor – the question was formulated in a different way, namely: "Do you think your dwelling requires renovation because of…?"

Indoor flushing toilet for sole use of the household – the toilet could have been shared with other households.

There were no other major divergences from common definitions.

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.

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 2007	HBS 2007
	in %	
Total	100.0	100.0
0-14	16.0	18.1
15-24	15.5	16.1
25-54	44.0	41.9
55-64	11.1	11.6
65+	13.4	12.3

Tab. 2. Structure of population by level of education

Specification	EU-SILC 2007	HBS 2007
	in %	
Total	100.0	1000
No school education	2.1	0.9
Completed primary	18.5	19.1
Lower secondary	5.1	6.6
Elementary vocational	26.8	26.6
Secondary	33.6	33.9
Higher	13.8	12.9

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

Households	Households		Persons in households	
	EU-SILC 2007	HBS 2006	EU-SILC 2007	HBS 2006
Total	13281985	13332332	37719639	37703168
Total = 100				
Employees	49.5	45.1	59.3	53.6
Farmers	2.6	4.6	3.6	7.0
Self-employed	4.8	6.1	5.6	7.1
Retirees	27.6	27.9	19.4	19.5
Pensioners	9.3	10.3	6.4	7.4
Maintained from non-earned sources	6.1	6.0	5.9	5.4

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

Households	Disposable income		Income from hired work	
	EU-SILC 2007	HBS 2006	EU-SILC 2007	HBS 2006
Total	16166	14767	9649	7482
Employees	18140	15455	15195	12677
Farmers	10550	13891	1042	1367
Self-employed	17828	20416	2700	2805
Retirees	14670	14278	1466	1423
Pensioners	10888	10478	1175	1022
Maintained from non-earned sources	8723	9272	1665	736

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

Households	Disposable income		Income from hired work	
	EU-SILC 2007	HBS 2006	EU-SILC 2007	HBS 2006
Total	16166	14769	9649	7484
1-person	15412	14686	4713	4103
2-persons	18725	17225	7980	6145
3-persons	18561	16511	12689	9955
4-persons	16235	14862	11714	9393
5-persons	13733	12716	8889	6735
6-persons and more	12456	11187	7166	4684

Tab. 6. Households provided with selected durables

Specification	EU-SILC 2007	HBS 2007
	in %	
Fixed telephone	71.6	67.9
Mobile telephone	75.5	79.3
Television set	97.1	98.5
Computer	48.7	50.1
Printer	35.7	33.6
Internet connection	34.8	36.6
Microwave oven	37.9	42.4
Dishwasher	9.3	7.4
Refrigerator	97.6	98.9
Washing machine	96.6	97.1
Passenger car	53.6	52.5

4.2. Comparison of Laeken Indicators based on EU-SILC 2006 and EU-SILC 2007

The preliminary analysis of the results of EU-SILC 2007 indicates that between 2005 and 2006 further improvement of the households' income condition occurred in Poland. Moreover, income inequality among the Polish population showed a decreasing trend. In 2006 the at-risk-of-relative-poverty rate estimated at 60% of the disposable income median was 17% as compared with 19% noted in 2005, while Gini coefficient amounted to 0.32 (in 2005 – 0.33 respectively). The income quintile share ratio (S80/S20) also went down slightly – being 5.3 as compared with 5.6 in 2006. The trends observed on the basis of EU-SILC 2007 as for the income condition and poverty reduction were also confirmed by the household budget survey results. Both surveys reflect positive macroeconomic trends (including gradual improvement of the labour market situation showed by lower unemployment figures, increase of real wages and retirement pays, higher incomes of the farmers' households brought about among others by the EU financial support).

4.3. Comparison of 2006 results of SNA and EU-SILC 2007 (data for 2006) for Poland

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 58% 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 estimates of primary and secondary distribution of income in SNA, used as a basis for the calculation of disposable income refer to some items not covered by EU-SILC 2007 or not taken into account when calculating EU-SILC results. The most important of these items is 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 2007 for farming. These differences are responsible for the fact that income from self-employment in EU-SILC 2007 amounted only to 26% of the operation surplus in SNA (after section K deduction).
The income from self-employment in EU-SILC 2006 is equal to 99% of the respective category in SNA, while social benefits – 93%, respectively, which seems to be a satisfactory outcome.
As compared with EU-SILC 2006, there was a higher convergence between EU-SILC 2007 data and SNA: for disposable income by 1 percentage point, for incomes from hired employment – by 3 p.p. and for social benefits – by 1 p.p. This marks further improvement of the quality of data. Some improvement of the data convergence with SNA was already noticed for 2005. The only decrease in convergence of data between SNA and EU-SILC was noted for the income from self-employment which dropped in EU-SILC 2007 by 1 p.p. as compared with EU-SILC 2006.

Comparison of 2006 results of SNA and EU-SILC 2007 for Poland

Category in SNA	Variables in EU-SILC 2005	Category description in EU-SILC 2005	SNA in mln PLN	EU-SILC in mln PLN	SNA = 100%	SNA = 100% EU-SILC 2006
Gross disposable income (net)	HY020	Total disposable household income (net)	683 483	398 939	58	57
Wages, salaries and other income connected with hired work (gross)	PY010G	Employee cash or near cash income (gross)	325 030	323 956	99	96
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	201 601	53 219	26	27
Social security benefits and social assistance benefits (gross)	PY90G + PY100G + PY110G + PY120G + PY130G + PY140G + HY050G + HY060G + HY070G	Social benefits (gross)	161 336	149 258	93	92

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 included in the results of EU-SILC 2007 (data for 2006), and market income from renting of real estate included in EU-SILC as the variable HY040G.