



EU-SILC 2007 Operation

Intermediate quality report

Czech Republic



December 2008

TABLE OF CONTENTS

1. Common cross-sectional European Union indicators	4
1.1 Common cross-sectional European Union indicators based on the cross-sectional component of EU-SILC	4
1.2 Other Indicators	8
2. Accuracy	8
2.1 Sampling design	8
2.1.1 Type of sampling	8
2.1.2 Sampling units	9
2.1.3 Stratification criteria	9
2.1.4 Sample size and allocation criteria	10
2.1.5 Sample selection schemes	10
2.1.6 Sample distribution over time	10
2.1.7 Renewal of the sample: Rotational groups	10
2.1.8 Weightings	10
2.1.8.1 Design factor	10
2.1.8.2 Non-response adjustments	10
2.1.8.3 Adjustments to external data	10
2.1.8.4 Final cross-sectional weights	11
2.1.9 Substitutions	11
2.2 Sampling errors	11
2.2.1 Standard errors and effective sample size	11
2.3 Non-sampling errors	12
2.3.1 Sampling frame and coverage errors	12
2.3.2 Measurement and processing errors	12
2.3.2.1 Measurement errors	12
2.3.2.2 Processing errors	13
2.3.3 Non-response errors	14
2.3.3.1 Achieved sample size	14
2.3.3.2 Unit non-response	16
2.3.3.3 Distribution of households by 'record of contact at address' (DB120), by 'household questionnaire result' (DB130) and by 'household interview acceptance' (DB135)	17
2.3.3.4 Distribution of substituted units	20
2.3.3.5 Item non-response	20
2.4 Mode of data collection	22
2.5 Interview duration	23
3. Comparability	23
3.1 Basic concepts and definitions	23
3.2 Components of income	23
3.2.1 Differences between the national definitions and standard EU-SILC definitions	23
3.2.2 The source or procedure used for collection of income variables	23
3.2.3 The form in which income variables at component level have been obtained	24
3.2.4 The method used for obtaining the income target variables in required form	24
4. Coherence	25
4.1 Comparison of income target variables and number of persons with external sources	25

TABLE OF CONTENTS

Table 1 At-Risk-of-Poverty Rate by Age and Gender	4
Table 2 At-Risk-of-Poverty Rate by Most Frequent Activity and Gender	4
Table 3 At-Risk-of-Poverty Rate by Household Type	5
Table 4 At-Risk-of-Poverty Rate by Accommodation Tenure Status, Gender and Selected Age groups	5
Table 5 At-Risk-of-Poverty Threshold (illustrative values)	6
Table 6 Relative Median Poverty Risk Gap by Age and Gender	6
Table 7 Dispersion around the At-Risk-of-Poverty Threshold by Gender and Selected Age Group	7
Table 8 At-Risk-of-Poverty Rate before Social Transfers by Gender and Selected Age Groups (Except Pensions).....	8
Table 9 At-Risk-of-Poverty Rate before Social Transfers by Age and Gender.....	8
Table 10 Mean, number of observations and standard errors for income components	11
Table 11 Response by interviewers' characteristics (%).....	13
Table 12 Sample size	14
Table 13 Regional disparities in response	15
Table 14 First wave: Distribution of households by 'record of contact at address' (DB120).....	17
Table 15 First wave: Distribution of address contacted by 'household questionnaire result' (DB130, DB135).....	17
Table 16 Second wave: Distribution of households by 'record of contact at address' (DB120)	18
Table 17 Second wave: Distribution of address contacted by 'household questionnaire result' (DB130, DB135).....	18
Table 18 Third wave: Distribution of households by 'record of contact at address' (DB120)	18
Table 19 Third wave: Distribution of address contacted by 'household questionnaire result' (DB130, DB135).....	19
Table 20 Total sample: Distribution of households by 'record of contact at address' (DB120)	19
Table 21 Total sample: Distribution of address contacted by 'household questionnaire result' (DB130, DB135).....	19
Table 23 Distribution of household members by type of interview (RB260)	23
Table 24 Overview of the collection of income data (net/gross values)	24
Table 25 Social income – comparison with administrative sources (Ministry of Labour and Social Affairs) – in million CZK	26
Table 26 Income – comparison with national accounts – in million CZK.....	26

1. Common cross-sectional European Union indicators

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

Primary Laeken indicators of social cohesion

At-risk-of-poverty rate after social transfers

The percentage of persons (over the total population) with an income below 60% of the median national income.

Table 1 At-Risk-of-Poverty Rate by Age and Gender

age	sex	rounded value
total	total	10
	men	9
	women	10
0_17 years	total	16
18_64 years	total	8
	men	8
	women	9
65+ years	total	5
	men	2
	women	8

Table 2 At-Risk-of-Poverty Rate by Most Frequent Activity and Gender

activity	sex	rounded value
employment	total	3
	men	3
	women	3
unemployment	total	48
	men	53
	women	44
non employment	total	13
	men	13
	women	14
retired	total	6
	men	4
	women	8
other inactive	total	13
	men	12
	women	14

Table 3 At-Risk-of-Poverty Rate by Household Type

household type	rounded value
total	10
households with no dependent children	6
one adult younger than 64 years	18
one adult older than 65 years	13
single female	17
single male	14
two adults, at least one aged 65 years and over	2
two adults younger than 65 years	5
three or more adults	2
households with dependent children	13
single parent with dependent children	37
two adults with one dependent child	7
two adults with two dependent children	8
two adults with three or more dependent children	29
three or more adults with dependent children	12

Table 4 At-Risk-of-Poverty Rate by Accommodation Tenure Status, Gender and Selected Age groups

age	tenure status	sex	rounded value
total	owner	total	7
		men	6
		women	7
	rent	total	19
		men	18
		women	20
0_17 years	owner	total	10
	rent	total	33
18_64 years	owner	total	6
		men	6
		women	7
	rent	total	17
		men	15
		women	18
65+ years	owner	total	5
		men	2
		women	7
	rent	total	8
		men	3
		women	11

Table 5 At-Risk-of-Poverty Threshold (illustrative values)

household type	currency	rounded value
single person	EUR	3251
	NAC	92146
	PPS	5348
two adults with two children younger than 14 years	EUR	6828
	NAC	193506
	PPS	11231

Inequality of income distribution S80/S20 income quintile share ratio

S80/S20 income quintile share ratio: Ratio of total income received by the 20% of the country's population with the highest income (top quintile) to that received by the 20% of the country's population with the lowest income (lowest quintile).

rounded value	3.5
---------------	-----

Relative median at-risk-of- poverty gap

Difference between the median income of persons below the at-risk-of poverty threshold, and the at-risk-of- poverty threshold; expressed as a percentage of the at-risk-of-poverty threshold.

Table 6 Relative Median Poverty Risk Gap by Age and Gender

age	sex	rounded value
total	total	18
	men	19
	women	17
0_17 years	total	19
18_64 years	total	19
	men	21
	women	19
65+ years	total	7
	men	14
	women	7

Secondary Laeken indicators of social cohesion**Dispersion around the risk-of-poverty threshold**

The percentage of persons (over the total population) with an income below 40 %, 50 % and 70 % of the national median income

Table 7 Dispersion around the At-Risk-of-Poverty Threshold by Gender and Selected Age Group

% of the national median income	age	sex	rounded value
40%	total	total	2
		men	2
		women	3
	0_17 years	total	4
	18_64 year	total	2
		men	2
		women	2
	65+ year	total	0
		men	0
women		0	
50%	total	total	5
		men	5
		women	6
	0_17 year	total	10
	18_64 year	total	5
		men	4
		women	5
	65+ year	total	2
		men	1
women		2	
70%	total	total	17
		men	15
		women	19
	0_17 year	total	25
	18_64 year	total	15
		men	13
		women	16
	65+ year	total	19
		men	10
women		25	

At-risk-of-poverty rate before transfers

The 'at-risk-of-poverty rate before social transfers except old-age and survivors' benefits' shows the percentage (over the total population) of the population having an equivalised disposable income before social transfers except old-age and survivors' benefits below the national 'at-risk-of-poverty threshold'.

The 'at-risk-of-poverty rate before social transfers including old-age and survivors' benefits' shows the percentage (over the total population) of the population having an equivalised disposable income before social transfers including old-age and survivors' benefits below the national 'at-risk-of-poverty threshold'.

Table 8 At-Risk-of-Poverty Rate before Social Transfers by Gender and Selected Age Groups (Except Pensions)

age	sex	rounded value
total	total	20
	men	19
	women	21
0_17 years	total	31
18_64 years	total	19
	men	18
	women	20
65+ years	total	12
	men	9
	women	14

Table 9 At-Risk-of-Poverty Rate before Social Transfers by Age and Gender

age	sex	rounded value
total	total	38
	men	35
	women	41
0_17 years	total	33
18_64 years	total	29
	men	26
	women	32
65+ years	total	89
	men	91
	women	88

Inequality of income distribution: Gini coefficient

The relationship of cumulative shares of the population arranged according to the level of income, to the cumulative share of the total income received by them.

rounded value	25
---------------	----

1.2 Other Indicators

Equivalised disposable income

The average of the equivalised disposable income of each person.

Mean equalised disposable income (HY020 equalised)	171 021
--	---------

The gender pay gap

The gender pay gap is not calculated from EU-SILC.

2. Accuracy

2.1 Sampling design

2.1.1 Type of sampling

The survey was carried out on the whole territory of the Czech Republic. The sample size of newly selected dwelling (first wave in 2007) was 4250 dwellings. Dwellings were selected using stratified two-stage sampling design. Small geographical areas (CEUs - census enumeration units) were first sampled as primary sampling units with probability proportional to their size. In the second stage, 10 dwellings were sampled in each sampled CEU.

2.1.2 Sampling units

Census Enumeration Districts (CEUs) constitute the first-stage sampling units. CEUs are small geographical areas covering the whole territory of the country. They are used as enumeration districts during the census, but their use is more general. Continuously updated geographical register is maintained by the CZSO, where these units form the basic geographical layer, on which subsequent aggregations are based. This register is the base for an integrated hierarchical geographical information system and is the base for databases of regional indicators and statistical data.

For each CEU, a list of all buildings is maintained in the register. This list is updated from administrative data of the construction authorities (new buildings', flats' or commercial premises' acceptance protocols, demolitions' protocols). For each building, the number of dwelling units is recorded.

CEUs vary considerably in size measured in number of dwelling units in them. Before drawing of the first stage sample, the sampling frame of CEUs had to be adjusted in two ways:

- As noted above, CEUs have wider use than sampling of dwellings and there are CEUs not containing any buildings dwellings (like industrial areas, railway stations and the like). These CEUs, where the number of dwellings is zero, are dropped from the sampling frame.
- In order to enable incorporation of small census enumeration units into the sampling process (to reach the required full geographical coverage of the national territory), small CEUs (with less than 20 inhabited dwellings) were merged with adjacent CEUs and this larger merged CEU entered the first stage of sampling. Therefore, in some cases, the 10 dwellings sampled in the second stage belong to two, in exceptional cases even more, real administrative CEUs. The survey design variable DB060 (PSU) is later coded according to this adjusted structure of the sampling frame, to keep the dwellings together as they were actually sampled.

In the second stage, 10 dwellings were sampled in each sampled CEU. CZSO's regional fieldwork units (each covering one of the 14 NUTS3 administrative regions) received the list of selected dwellings (address + identification number of the flat in buildings with more than one flat). Before the actual fieldwork, the regional fieldwork units' staff carried out identification of the selected dwellings and filled in the contact names on the list of selected dwellings for interviewers.

The ultimate sampling unit was the dwelling, i.e. all persons with usual residence in that dwelling (their only place of residence or their main place of residence, according to the EU-SILC definition) were included in the survey. This includes also foreign nationals and sub-tenants living in the selected dwelling.

The household definition is based on the sharing of expenditures concept, in line with the definition of Paragraph 115 of the national Civil Code – based on the declaration of the persons in sampled dwelling unit that they permanently live together and finance together expenditures to cover their needs.

2.1.3 Stratification criteria

The sampling of CEUs is stratified by region (NUTS4) and municipality size with following four categories:

- below 2 000 inhabitants
- 2000 – 9999 inhabitants
- 10 000 – 49 999 inhabitants
- 50 000 and more inhabitants

2.1.4 Sample size and allocation criteria

The total sample size was 11 926 dwellings (12063 households) from which 4250 addresses were newly selected and 7458 dwellings (7568 households) were revisited (second and third wave). The new sample was allocated to the strata using proportional algorithm (proportionally to the number of dwellings in the sampling frame).

2.1.5 Sample selection schemes

In the first stage, CEUs were sampled with probability proportional to size (number of dwellings). Simple random sampling without replacement is used for sampling of constant number of 10 dwellings in each sampled CEU.

2.1.6 Sample distribution over time

Due to the limited duration of the fieldwork period, the survey was organized as a one-shot survey. Sample was not distributed into separate waves over the duration of the fieldwork.

2.1.7 Renewal of the sample: Rotational groups

The survey will in the long term use the integrated four-year rotational panel design. Since the 2005 operation was the first year of the survey, there was only one sample replication and no rotation was applied. Due to the relatively small sample size in 2005, all responding households were carried over to the 2006 operation. One new sample replication was added in 2006 and 2007. The rotational scheme with four replications will be functional starting in 2009, when the households from the 2005 operation will be dropped from the sample.

The sample rotation will be at the level of CEUs as primary sampling units (whole CEUs will be added to/dropped from the sample).

2.1.8 Weightings

2.1.8.1 Design factor

The sample was designed as a self-weighting sample. Design factor for all sampled dwellings is equal to 1.

2.1.8.2 Non-response adjustments

The original sample was designed as a self-weighting probability sample. However, non-ignorable level of non-response biased the structure of the sample of achieved interviews. For example, compared to the available demographic statistics and external data, the achieved average household size was significantly smaller. There was under-representation of the self-employed, of the unemployed as well as of persons living in larger cities. On the other hand, there was overrepresentation of persons in the retirement age and of persons living in family houses.

Due to the limited information on non-respondents of the first wave restricted only to the geographical information obtainable from the sampling frame, the possibilities for modelling using propensity to response models were quite limited. There was an option by second wave households to utilize information, which was obtained from previous SILC wave, and to adjust their previous year weights for attrition. In that case it would be difference between first and next wave weighting procedures. Experimental computations show that this method would entail excessive weights variability increase. Therefore, united calibration for all the waves was used as the method for correcting non-response.

The achieved sample was re-weighted using the integrated calibration technique (producing the same weights on household and personal level). This technique ensures that the weighted sample structure corresponds to a set of known external population characteristics. The calculations were implemented using the CALMAR software in SAS.

2.1.8.3 Adjustments to external data

The following calibration variables were used:

- number of inhabited dwellings in each NUTS3 region, subdivided into family houses (detached and semi-detached houses) and flats, based on the 2001 Census continuously updated from administrative sources of construction authorities
- population characteristics in each NUTS 3 region:
 - o population totals from demographic statistics
 - o economic activity characteristics in each NUTS3 region:
 - number of pensioners (excl. pensions for orphans), based on the administrative data from social security administration
 - number of unemployed (registered unemployed from administrative source of the Ministry of Labour and Social Affairs, corrected for unregistered unemployment using the Labour Force Survey data)
 - number of self-employed (estimate based on the Labour Force Survey)
 - number of children aged 0-15 (from demographic statistics)
- population characteristics at the national level:
 - o age groups 0-15, 16-24, 25-34, 35-44, 45-54, 55-64, 65+ - based on the demographic statistics)
 - o gender at the national level (based on the demographic statistics)
 - o municipality size at the national level (below 2 000 inhabitants, 2 000 - 9 999, 10 000 - 49 999, 50 000+ inhabitants)

Since the target population of the survey were persons living in private households, the demographic statistics aggregate data were adjusted by subtracting institutionalized population (from social security administrative data) and persons in prisons.

2.1.8.4 Final cross-sectional weights

Final household cross-sectional weight was result of Calmar calibration.

	N	Minimum	Maximum	Mean	Std. Dev.
Weights DB0909675	100	3475	417.92	205.52	

The number of cross-sectional weights (number of DB090 > 0 is 9691) differs from the number of successfully interviewed households by 16. There can be more than one household in the dwelling and in these cases occurred that one of the households in the dwelling refused the interview, was unable to respond, moved or the households have merged while at least one of the households in the dwelling was successfully interviewed. Since the calibration is performed at the dwelling level, these households get also non-zero weight. Nevertheless the number of successfully interviewed households is 9675.

2.1.9 Substitutions

Substitutions were not used.

2.2 Sampling errors

2.2.1 Standard errors and effective sample size

The estimated standard errors, confidence intervals and design effects for the main indicators are provided below:

Table 10 Mean, number of observations and standard errors for income components

Indicator	Value	Std.error	95% C.I.		Deff
Calculated at household level:					
Mean disposable income (HY020)	294 252	3032	288 308	300 195	1.38
Mean equalised disposable income (HY020 equalised)	171 021	1534	168 014	174 027	1.36
Calculated at individual level:					
At-risk-of poverty rate (with fixed poverty line)	9.8%	0.4%	9.0%	10.7%	1.29

The estimated standard errors take into account the complex sampling scheme used in the survey (stratification, two-stage design). Results were obtained using the linearisation method. The computations were done in R 2.4.0 software, survey package 3.6-5.

2.3 Non-sampling errors

2.3.1 Sampling frame and coverage errors

Sampling frame covers existing buildings with the information on number of dwelling units in each building (see part on sampling units for description of the register of CEUs).

Out of the 4250 newly sampled dwelling unit records (in the first wave), 217 were found to be ineligible for the survey (4.4 %). Fieldwork staff undertaking pre-fieldwork identification of sampled dwelling units and interviewers must declare clear confirmation of the fact, that the dwelling unit was not located.

2.3.2 Measurement and processing errors

2.3.2.1 Measurement errors

Development of the questionnaires

Data collection had the form of an interview and interviewers filled in the answers into paper questionnaires (PAPI data collection).

The survey was conducted using paper questionnaires designed for OCR technology data capture (scanning). The first SILC questionnaires were developed in 2004. The inputs for designing the questionnaires were the questionnaires from Microcensus surveys (national income survey), the harmonised description of EU-SILC target variables (technical document SILC 065) and the blueprint questionnaire in English used for previous SILC pilots in old Member States. Basic questionnaire structure follows the practice already well established in the Microcensus, with three main forms: dwelling unit questionnaire with household membership roster, household questionnaire and personal questionnaire. The questionnaires were first tested in pilot survey of 600 randomly sampled households (Spring 2004). The pilot project involved 14 future regional coordinators of the survey and small group of experienced interviewers (2-3 per region). After this fieldwork test, questionnaire was updated and partly re-designed, with active involvement of the regional staff and the participating interviewers. Together with the questionnaires, detailed interviewers guidelines were developed with binding instructions to all questions.

The content of the survey was divided into three questionnaires with different units of reference:

Questionnaire A (dwelling unit questionnaire): contained the roster with the list of all persons with usual residence in the selected dwelling, their basic demographic and social characteristics, information on sharing of expenses to determine household units¹ and relationship of each person to the main user of the dwelling and to the head of household.

Questionnaire B (household questionnaire): filled in for each household, contained information on housing, childcare, financial situation of the household, consumer durables, inter-household transfers paid and received, consumption from household own production (i.e. small scale farming and similar activities), family social benefits, rental income and paid regular taxes on wealth (buildings and land).

Questionnaire C (personal questionnaire): filled in by each household member aged 16+ as of 31.12.2006 (i.e. persons born in 1990 and earlier). This questionnaire contained information on labour status and employment, personal income, participation in private pension plans, health,

¹ Since the household definition is based on sharing of expenditures (housekeeping concept), there are dwelling units with more than one household. If this was the case, all households in selected dwellings were included as eligible for the survey.

education and selected biographical information. The questionnaire C was supplemented with the EU-SILC Module 2007 (housing conditions).

Reference periods

- Age: 31.12.2006
- Other demographic variables: marital status, education: at the date of the interview
- Current employment variables (current employment status, occupation, ...): at the date of the interview
- Income data: calendar year 2006
- Housing, consumer durables, financial and social situation of household: at the date of the interview, unless the question specifically refers to some other reference period

Interviewers

The survey was performed by 777 interviewers (approximately almost 15 households per interviewer). The following table shows the successfulness of the interviewers by their basic characteristics (if there are more than one household in the dwelling, at least one interviewed household is considered as successfully surveyed).

Table 11 Response by interviewers' characteristics (%)

Interviewers' characteristics	Total	Wave 1	Wave 2	Wave 3
Age:				
Age ≤ 40	80.96	61.95	91.52	94.22
Age 41-60	82.95	64.15	92.40	93.45
Age > 60	86.16	68.74	93.73	95.96
Sex:				
Male	81.74	61.55	90.26	93.93
Female	83.86	65.87	93.42	94.34
Education:				
Primary	85.63	70.05	93.51	95.81
Lower secondary	87.48	71.56	92.65	95.23
Upper secondary	83.48	65.98	92.65	93.92
Tertiary education	78.32	54.22	91.77	94.17
Economic activity:				
Employed	81.06	61.18	92.32	93.38
Student	82.89	64.75	91.97	92.69
Retired	86.42	69.80	93.16	95.58
Unemployed	80.23	69.57	93.10	90.91
Other	81.26	63.10	89.62	95.42
Experience with surveys:				
SILC 2005 - yes	87.31	63.31	92.67	95.30
- no	80.46	65.20	92.51	91.76
SILC 2006 - yes	86.27	64.30	93.12	94.44
- no	73.83	65.16	89.35	92.40
Other	84.32	65.93	89.35	94.62
Different interviewer in 2006			88.13	90.92
Same interviewer as in 2006			94.20	95.06
Total	83.27	64.69	92.55	94.22

2.3.2.2. Processing errors

Data processing

Data were captured using OCR technology (scanning). After the data collection in the field, the regional fieldwork staff gathers the questionnaire material. While accepting the material from each interviewer, the initial check is performed – the way, how the questionnaires are filled, completeness of the questionnaires, basic consistence checks. Then, control sum of numerical values on each page is calculated and filled by the regional coding staff. Larger tables, with more numerical data, have their own control sums. At the same time, the coding staff coded some variables – occupation (ISCO), sector of employment (NACE) and country codes for country of birth and citizenship variables.

After this preparatory phase, questionnaires are scanned into raw data files. CZSO has three specialised scanning units with technical equipment and expertises in this data capture technology. This technology is also used extensively in business and agricultural surveys. Control sums are automatically checked during scanning. Whenever the sum of captured values does not match the control sum or when some number is not properly recognised, that position of the questionnaire appears as image on the screen of the operator for verification. Images of the scanned questionnaires are also stored with the captured data with unique filenames allowing linking of each data record with the image of the questionnaire, from which the data were captured.

The raw data files are then subject to initial centrally performed checks – checking the integrity of identification numbers, consistency with the sample, completeness of the questionnaire sets for all dwellings. Regional staff is responsible for further checking of the data for their respective region, using a special software application containing a set of logical controls, captured data and linked images of the questionnaires. Three kinds of errors are distinguished: critical errors (must be corrected, limited to a small set of key consistency issues), errors to verify (must be commented, involving contacting the interviewer in charge of that household, if additional information is necessary) and informative flags (extraordinary or unusual situations, which should be looked at).

2.3.3 Non-response errors

2.3.3.1 Achieved sample size

4 250 new dwellings entered the survey (1st wave) and 7 676 dwellings were revisited - 7458 at the last year's address and 218 were tracked to their new home. The fieldwork revealed that among the total of 11 926 dwellings in the sample there were 428 dwellings (4 %) unoccupied, unlocated or ineligible because the households had moved. Since there was no substitution for these ineligible units, the survey was conducted in 11 498 dwellings and 11 611 households. There were 113 additional interviewed households in these dwellings, since in 107 dwellings there are more households in one dwelling unit (household definition is based on sharing of expenses).

The overview of the survey response can be summarised by Table 12:

Table 12 Sample size

	Households				%			
	total	1st wave	2nd wave	3rd wave	total	1st wave	2nd wave	3rd wave
Response, total	9675	2654	3344	3677	80.2	62.1	89.5	90.8
Non-response, total	1936	1437	272	227	19.8	37.9	10.5	9.2
					100	100	100	100
Refusals (unwillingness to give information)	1515	1138	211	166	78.3	79.2	77.6	73.1
Household not contacted. temporarily absent	363	257	54	52	18.8	17.9	19.9	22.9
Household unable to respond (health limitation)	47	31	7	9	2.4	2.2	2.6	4.0
Other reasons (linguistic etc.)	11	11	0	0	0.6	0.8	0.0	0.0
	Persons 16+							

	total	1st wave	2nd wave	3rd wave
Response, total	19384	5330	6646	7408

Refusals also include situations when the household did not refuse the survey as such, but did not accept to provide the information on income to the extent, which would qualify the household as successfully interviewed. The definition of successfully interviewed household allowed missing income data for only one person and the person must not be the head of the household. Non-contacts, temporarily absent category cover situations, when the interviewer did not establish contact with the selected household, despite the prescribed minimum number of three attempts of personal contact.

Table 13 Regional disparities in response

Region (NUTS3)	total			1 st wave			2 nd and 3 rd wave		
	HHs in survey	response		HHs in survey	response		HHs in survey	response	
		count	%		count	%		count	%
Praha	1237	864	69.8	538	254	47.2	699	610	87.3
Stredocesky	1207	1006	83.3	462	326	70.6	745	680	91.3
Jihočeský	704	612	86.9	250	188	75.2	454	424	93.4
Plzeňský	669	562	84.0	210	129	61.4	459	433	94.3
Karlovarský	364	328	90.1	124	101	81.5	240	227	94.6
Ústecký	951	787	82.8	334	216	64.7	617	571	92.5
Liberecký	476	391	82.1	172	111	64.5	304	280	92.1
Královéhradecký	611	513	84.0	209	134	64.1	402	379	94.3
Pardubický	596	513	86.1	198	143	72.2	398	370	93.0
Vysočina	586	510	87.0	180	120	66.7	406	390	96.1
Jihomoravský	1176	948	80.6	433	252	58.2	743	696	93.7
Olomoucký	778	666	85.6	260	171	65.8	518	495	95.6
Zlínský	657	576	87.7	218	163	74.8	439	413	94.1
Moravskoslezský	1599	1399	87.5	503	346	68.8	1096	1053	96.1
CZ total	11611	9675	83.3	4091	2654	64.9	7520	7021	93.4

The lowest achieved response rate was in the City of Prague region, almost 70 percent. This result has its objective reasons, as in any other large city, the social environment and dwelling structure in this metropolitan region is the least favourable for conducting household surveys. On the other hand, Karlovarsky region (West Bohemia) is the region with exceptionally high response rate, above 90 percent. For the remaining regions, the differences between response rates are not large. As in other surveys, the highest response rates (above 87 percent) were achieved in the Eastern part of the country (Moravskoslezsky, Vysocina, Zlinsky regions). The other regions have response rates also above 80 percent.

Participation in the national EU-SILC survey is voluntary, there is no duty imposed on households to provide the required information, like it is for example in the population census. The household must be informed about the content of the survey and that its participation is voluntary and left to its decision. The main reasons for refusal reported from the field are privacy reasons (objections against giving personal information and fear of misuse of the personal data), unwillingness to report income, fear of contact with interviewers as strangers. There is a considerable group of persons, who, as a matter of principle, strictly refuse to give any information about them and their households.

SILC data files non-response characteristics, with the SILC harmonised response rates²:

Achieved sample size is 9 675.

Number of households for which an interview is accepted for the database: 9 675

Number of persons of 16 years or older, who are members of the households and for whom the interview is accepted for the database: 19 384

2.3.3.2 Unit non-response

New replication

• Household non-response rates (NRh)

$$NRh = (1 - (Ra * Rh)) * 100$$

Where

$$Ra = \frac{\text{Number of addresses successfully contacted}}{\text{Number of valid addresses selected}} \\ = \frac{\sum [DB120 = 11]}{\sum [DB120 = all] - \sum [DB120 = 23]} = \frac{4091}{4277 - 0} = 0.956512$$

$$Rh = \frac{\text{Number of household interviews completed and accepted for the database}}{\text{Number of eligible households at contacted addresses}}$$

$$= \frac{\sum [DB135 = 1]}{\sum [DB130 = all]} = \frac{2654}{4277} = 0.620528$$

$$NRh = (1 - 0.956512 * 0.620528) * 100 = 40,64754$$

The household non-response rate is about 40,65 %.

• Individual non-response rates (NRp)

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

Where

$$Rp = \frac{\text{Number of personal interview completed}}{\text{Number of eligible individuals}} = \frac{5330}{5330} = 1$$

$$NRp = (1 - 1) * 100 = 0 \%$$

So, the individual non-response rate is 0 %.

• Overall individual non-response rates (*NRp)

$$*NRp = (1 - (Ra * Rh * Rp)) * 100$$

$$*NRp = (1 - (0.95605 * 0.65459 * 1)) * 100 = 37.41813$$

So, the overall individual non-response rate is about 37,42 %.

Total sample

• Household non-response rates (NRh)

² For the more detailed definitions of the SILC database variables, please refer to the SILC UDB Documentation.

$$Ra = 11611/(12063 - 265) = 0.98415$$

$$Rh = 9675/11611 = 0.833262^3$$

$$NRh = (1-0.98415*0.833262)*100 = 17.99458$$

• **Individual non-response rates (NRp)**

$$Rp = 19384/19384 = 1$$

$$NRp = (1-1)*100 = 0 \%$$

• **Overall individual non-response rates (*NRp)**

$$*NRp=(1-(Ra*Rh*Rp))*100$$

$$*NRp = (1-(0.98415*0.833262*1))*100 = 17.99458$$

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

First wave

Table 14 First wave: Distribution of households by 'record of contact at address' (DB120)

	Number	Percentage
Total (DB120 = 11 to 23)	4277	100.00
Address contacted (DB120 = 11)	4091	95.65
Address non-contacted (DB120 = 21 to 23)	186	4.35
Total address non-contacted (DB120 = 21 to 23)	186	100.00
Address cannot be located (DB120 = 21)	186	100.00
Address unable to access (DB120 = 22)	0	0.00
Address does not exist or is non-residential address or is unoccupied or not principal residence (DB120 = 23)	0	0.00

Table 15 First wave: Distribution of address contacted by 'household questionnaire result' (DB130, DB135)

	Number	Percentage
Total	4091	100.00
Household questionnaire completed (DB130 = 11)	2654	64.87
Interview not completed (DB130 = 21 to 24)	1437	35.13
Total interview not completed (DB130 = 21 to 24)	1437	100.00
Refusal to co-operate (DB130 = 21)	1138	79.19
Entire household temporarily away for duration of fieldwork – i.e. non-contacts (DB130 = 22)	257	17.88
Household unable to respond (illness, incapacity, etc.) (DB130 = 23)	31	2.16
Other reasons (DB130 = 24)	11	0.77

³ There were more than one household units in some interviewed dwellings (107 cases, with 113 additional households, out of which 110 were successfully interviewed). These 106 households are included in the database. Their inclusion in the non-response calculation slightly bias upwards the non-response calculated at the household level – assuming that at least in some of the non-responding dwellings can also include more than one household unit, the denominator should be higher than 11 611. This difference is unknown, but is likely to be quite small.

Household questionnaire completed (DB135 = 1+ 2)	2654	100.00
Interview accepted for data base (DB135 = 1)	2654	100.00
Interview rejected (DB135 = 2)	0	0.00

Second wave

Table 16 Second wave: Distribution of households by 'record of contact at address' (DB120)

	Number	Percentage
Total (DB120 = 11 to 23)	3737	100.00
Address contacted (DB120 = 11)	3616	96.76
Address non-contacted (DB120 = 21 to 23)	121	3.24
Total address non-contacted (DB120 = 21 to 23)	121	100.00
Address cannot be located (DB120 = 21)	0	0.00
Address unable to access (DB120 = 22)	0	0.00
Address does not exists or is non-residential address or is unoccupied or not principal residence (DB120 = 23)	121	100.00

Table 17 Second wave: Distribution of address contacted by 'household questionnaire result' (DB130, DB135)

	Number	Percentage
Total	3616	100.00
Household questionnaire completed (DB130 = 11)	3344	92.48
Interview not completed (DB130 = 21 to 24)	272	7.52
Total interview not completed (DB130 = 21 to 24)	272	100.00
Refusal to co-operate (DB130 = 21)	211	77.57
Entire household temporarily away for duration of fieldwork – i.e. non-contacts (DB130 = 22)	54	19.85
Household unable to respond (illness, incapacity, etc.) (DB130 = 23)	7	2.57
Other reasons (DB130 = 24)	0	0.00
Household questionnaire completed (DB135 = 1+ 2)	3344	100.00
Interview accepted for database (DB135 = 1)	3344	100.00
Interview rejected (DB135 = 2)	0	0.00

Third wave

Table 18 Third wave: Distribution of households by 'record of contact at address' (DB120)

	Number	Percentage
Total (DB120 = 11 to 23)	4049	100.00
Address contacted (DB120 = 11)	3904	96.42
Address non-contacted (DB120 = 21 to 23)	145	3.58
Total address non-contacted (DB120 = 21 to 23)	145	100.00
Address cannot be located (DB120 = 21)	1	0.69
Address unable to access (DB120 = 22)	0	0.00
Address does not exists or is non-residential address or is unoccupied or not principal residence (DB120 = 23)	144	99.31

Table 19 Third wave: Distribution of address contacted by 'household questionnaire result' (DB130, DB135)

	Number	Percentage
Total	3904	100.00
Household questionnaire completed (DB130 = 11)	3677	94.19
Interview not completed (DB130 = 21 to 24)	227	5.81
Total interview not completed (DB130 = 21 to 24)	227	100.00
Refusal to co-operate (DB130 = 21)	166	73.13
Entire household temporarily away for duration of fieldwork – i.e. non-contacts (DB130 = 22)	52	22.91
Household unable to respond (illness, incapacity, etc.) (DB130 = 23)	9	3.96
Other reasons (DB130 = 24)	0	0.00
Household questionnaire completed (DB135 = 1+ 2)	3677	100.00
Interview accepted for data base (DB135 = 1)	3677	100.00
Interview rejected (DB135 = 2)	0	0.00

Total sample

Table 20 Total sample: Distribution of households by 'record of contact at address' (DB120)

	Number	Percentage
Total (DB120 = 11 to 23)	12063	100.00
Address contacted (DB120 = 11)	11611	96.25
Address non-contacted (DB120 = 21 to 23)	452	3.75
Total address non-contacted (DB120 = 21 to 23)	452	100.00
Address cannot be located (DB120 = 21)	187	41.37
Address unable to access (DB120 = 22)	0	0.00
Address does not exist or is non-residential address or is unoccupied or not principal residence (DB120 = 23)	265	58.63

Table 21 Total sample: Distribution of address contacted by 'household questionnaire result' (DB130, DB135)

	Number	Percentage
Total	11611	100.00
Household questionnaire completed (DB130 = 11)	9675	83.33
Interview not completed (DB130 = 21 to 24)	1936	16.67
Total interview not completed (DB130 = 21 to 24)	1936	100.00
Refusal to co-operate (DB130 = 21)	1515	78.25
Entire household temporarily away for duration of fieldwork – i.e. non-contacts (DB130 = 22)	363	18.75
Household unable to respond (illness, incapacity, etc.) (DB130 = 23)	47	2.43
Other reasons (DB130 = 24)	11	0.57
Household questionnaire completed (DB135 = 1+ 2)	9675	100.00
Interview accepted for data base (DB135 = 1)	9675	100.00
Interview rejected (DB135 = 2)	0	0.00

2.3.3.4 Distribution of substituted units

Substitutions were not used.

2.3.3.5 Item non-response

In Table 20 an overview of the item non-response for all income variables is presented. The percentage households having received an amount, the percentage of households with missing values and the percentage of households with partial information is calculated.

These percentages are calculated as follows:

% of households having received an amount: number of households (or persons) who have received something (yes to a filter) / total

% of households with missing values: number of households (or persons) who said that they have received something but did not give any amount (no partial information) / number of households (or persons) who have received something (yes to a filter)

% of households with partial information: number of households (or persons) who said that they have received something but gave partial information (amounts were not given for all components) / number of households (or persons) who have received something (yes to a filter)

Table 22 Overview of the non-response for the income variables - % households having received an amount, % of households with missing values and % of households with partial information

Item non-response (overview for different income components) ⁴	% of households having received an amount	% of households with missing values (before imputation)	% of households with partial information (before imputation)
Total gross household income (HY010)	99.99	0.00	0.07
Total disposable household income (HY020)	99.99	0.00	0.07
Total disposable household income before social transfers except old-age and survivor's benefits (HY022)	98.69	0.00	0.07
Total disposable household income including social transfers except old-age and survivor's benefits (HY023)	88.11	0.00	0.08
Net income components at household level			
Income from rental of a property or land (HY040N)	4.03	0.00	0.00
Family related allowances (HY050N)	26.83	0.00	0.00
Social exclusion not elsewhere classified (HY060N)	3.47	0.00	0.00
Housing allowance (HY070N)	3.99	0.00	0.00
Regular inter-household cash transfer received (HY080N)	7.50	0.00	0.00
Income received by people aged < 16 (HY110N)	0.00	0.00	0.00
Regular taxes on wealth (HY120N)	61.09	0.00	0.00
Regular inter-household cash transfer paid (HY130N)	7.02	0.00	0.00
Tax on income and social contributions (HY140N)	66.75	0.00	0.00
Gross income components at household level			
Income from rental of a property or land (HY040G)	4.03	0.00	0.00
Family related allowances (HY050G)	26.83	0.00	0.00
Social exclusion not elsewhere classified (HY060G)	3.47	0.00	0.00
Housing allowance (HY070G)	3.99	0.00	0.00
Regular inter-household cash transfer received (HY080G)	7.50	0.00	0.00
Interests, dividends, etc. (HY090G)	13.96	0.00	0.00
Interest repayments on mortgage (HY100G)	8.44	0.00	0.00
Regular taxes on wealth (HY120G)	61.09	0.00	0.00
Regular inter-household cash transfer paid (HY130G)	7.02	0.00	0.00
Tax on income and social contributions (HY140G)	66.75	0.00	0.00

⁴ For the more detailed definitions of the SILC income variables, please refer to the SILC UDB Documentation

	% of persons 16+ having received an amount	% of persons with missing values (before imputation)	% of persons with partial information (before imputation)
Net income components at personal level			
Employee cash or near cash income (PY010N)	47.89	0.06	0.00
Contributions to individual private pension plans (PY035N)	36.71	0.04	0.00
Value of goods produced by own-consumption (PY070N)	19.89	3.22	0.00
Pension from individual private plans (PY080N)	0.51	0.00	0.00
Unemployment benefits (PY090N)	2.89	0.18	0.00
Old age benefits (PY100N)	29.88	0.00	0.00
Survivor' benefits (PY110N)	8.25	0.00	0.00
Sickness benefits (PY120N)	8.32	0.06	0.00
Disability benefits (PY130N)	8.02	0.00	0.00
Education-related allowances (PY140N)	0.76	0.00	0.00
Gross income components at personal level			
Employee cash or near cash income (PY010G)	47.89	0.06	0.00
Non cash employee income (PY020G)	27.26	0.02	0.00
Contributions to individual private pension plans (PY035G)	36.71	0.04	0.00
Cash benefits or losses from self-employment (PY050G)	7.58	2.52	0.00
Value of goods produced by own-consumption (PY070G)	19.89	3.22	0.00
Pension from individual private plans (PY080G)	0.51	0.00	0.00
Unemployment benefits (PY090G)	2.89	0.18	0.00
Old age benefits (PY100G)	29.99	0.00	0.00
Survivor' benefits (PY110G)	8.25	0.00	0.00
Sickness benefits (PY120G)	8.32	0.06	0.00
Disability benefits (PY130G)	8.02	0.00	0.00
Education-related allowances (PY140G)	0.76	0.00	0.00

2.4 Mode of data collection

Distribution of household members by data status (RB250)

Registers are not used at all. Due to strict definition of response, there are any “not completed interviews” at individual level or “not contacted individuals” (all such cases were filled as proxy or were self-administered by respondents).

Distribution of household members by type of interview (RB260)

The data collection method was PAPI (paper-and-pencil interview). Most of the questionnaires were filled during fact-to-face interview with the interviewer. Some personal questionnaires were filled as proxy interviews – information for household member not present at the time of the interview was provided by another household member. In some case, where this was agreed with the household, interviewer left the personal questionnaire for some household member and collected it later (self-administered questionnaire).

Table 23 Distribution of household members by type of interview (RB260)

Method	Total		First wave		Second wave		Third wave	
	Count	%	Count	%	Count	%	Count	%
Face-to-face with paper questionnaire	17 610	90.8	4 839	90.8	6 222	91.3	6 549	90.5
Face-to-face with computer (CAPI)	not used	-	not used	-	not used	-	not used	-
Telephone interviews (CATI)	not used	-	not used	-	not used	-	not used	-
Self administered questionnaire	56	0.3	14	0.3	15	0.2	27	0.4
Proxy face-to-face interview (information from another household member)	1 718	8.9	477	8.9	581	8.5	660	9.1
Total	19 384	100	5 330	100	6 818	100	7 236	100

2.5 Interview duration

The average interview duration in successfully interviewed households (the whole interview time: household + all personal questionnaires combined) was 84.0 minutes.

3. Comparability

3.1 Basic concepts and definitions

- The reference period: no differences between the national and standard EU-SILC concept
- The private household definition: no differences (there can be more households in one dwelling eligible for the survey)
- The household membership: no differences
- The income reference period used: last calendar year
- The period for taxes and social contributions: taxes and social insurance contribution refer to the income received during the income reference period
- The reference period for taxes on wealth: income reference period
- The lag between the income reference period and current variables: three to four months (the survey took place from the end of February to the end of April 2006)
- The total duration of the data collection of the sample: 8 weeks
- Basic information on activity status during the income reference period: no differences

3.2 Components of income

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

The concepts and definitions used in the survey are those set in the EU-SILC documentation (definitions of target variables, as they are set in the EU-SILC regulations and technical document "Description of Target Variables – Doc. SILC 065). There is only one deliberate deviation from the used concepts:

Variable PY070 Value of goods produced by own-consumption, which is defined at the level of individual household members, is collected at the household level and later assigned to the head of household. This is due to the difficult attribution of this income in kind to individual household members (includes mainly small scale farming activities for own-consumption or own-consumption from family businesses).

3.2.2 The source or procedure used for collection of income variables

All the income variables are obtained by interview. The EU-SILC income target variables were divided to more subcomponents. The subcomponents were defined according to the Czech benefit system. These subcomponents were surveyed.

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

Table 24 Overview of the collection of income data (net/gross values)⁵

income component	% collected net of taxes and social contributions	% collected gross ⁶
PY010G	26.0	74.0
PY010N	26.0	74.0
PY020G	0.0	100.0
PY020N	-	-
PY035G	100.0	0.0
PY035N	100.0	0.0
PY050G	18.9	81.1
PY050N	-	-
PY070G	0.0	100.0
PY070N	100.0	0.0
PY080G	100.0	0.0
PY080N	100.0	0.0
PY090G	0.0	100.0
PY090N	100.0	0.0
PY100G	0.0	100.0
PY100N	100.0	0.0 ⁷
PY110G	0.0	100.0
PY110N	100.0	0.0
PY120G	0.0	100.0
PY120N	100.0	0.0
PY130G	0.0	100.0
PY130N	100.0	0.0
PY140G	0.0	100.0
PY140N	100.0	0.0

Both alternatives (gross amounts, net amount – net of taxes and social insurance contributions) were available to respondents for income from employment and self-employment income. In addition, information on claimed tax deductions was collected from respondents. Algorithms based on detailed application of the national tax rules were then used to calculate the complementary net/gross amount. Social benefits are generally tax-exempt – therefore there is no difference between gross and net values – they can be collected as one value and assigned to both gross and net.

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

Situation of **missing income data** for one of the household members was rare (7 cases). For these persons, the income was **imputed by the simple hot-deck method** (using randomly chosen person with similar characteristics from another household).

Another source of bias, which needs to be taken into account, stems from the interviewing. Data on income obtained during interviews with household members have the tendency to underestimate certain sources of income or data on some components is missing (**item non-response**).

⁵ For the definitions of the SILC database income variables, please refer to the SILC UDB Documentation.

⁶ Gross amount does not include social insurance contributions for the self-employed – where these are treated in our national system as part of the tax-deductible costs and not as part of the gross self-employment income.

⁷ In 22 cases variable of net series is not filled because variable of gross series is filled (Flag –5)

Underestimation of income is a natural consequence of the fact, that respondents either tends to give lower than actual values or simply did not recall certain irregular or small incomes. It is, more or less, a non-sampling error, affected substantially by the incomes themselves and by their source. The possibilities to eliminate this underestimation of the survey data are limited. In the presented survey, only such adjustments were done, where there was sufficiently reliable external statistical source or which can be based on the legislation.

Data on gross income from employment were compared with corresponding data from wage statistics broken into sectors of activity (NACE). Different from the last year's survey and in accordance with experience from other income surveys, income from work was underestimated (roughly by 5.4 %). Primarily, this underestimation concerned those incomes that were recorded as yearly lump sums. Such incomes were moderately boosted so that the average monthly gross pay by sectors approached the data from wage statistics. There was no need for corrections with income from private enterprise.

In case of social benefits for which there is a legal entitlement (parental leave benefit, child birth benefit, death grant provided to families of the deceased, to some extent also maternity leave benefit), a check on their receiving by the eligible households was applied and amounts provided were corrected according to the amounts fixed by the legislation. Old age benefits (pension from the social security system) were not corrected, since their underestimation is quite low.

Amounts declared by the unemployed as unemployment benefits were overestimated. Unemployed respondents tend to report their income from social benefits as unemployment benefits and do not distinguish them from the minimum income support benefits (claimed on the basis of the legal minimum subsistence amounts). In cases where the duration of unemployment and the reported amounts did not match the rules of the unemployment benefits provision, the reported amounts were re-classified as minimum income support benefits.

It was not possible to correct the underestimation of the sickness benefits (where respondents tend to forget spells of short-term illness over the 12 months income reference period), means-tested social benefits whose claims depend on the previous income (prior to the income reference periods), capital income and non-monetary income generated by own-consumption.

The value of goods produced by own-consumption was an estimate of the household based on the amount of consumed food and other goods, own production and goods from own business during the year 2006 (for example food and animals from own small-scale non-commercial farming activity, value of meals from own restaurant, bread from own bakery and the like).

4. Coherence

4.1 Comparison of income target variables and number of persons with external sources

The numbers of recipients of most of the incomes were used as calibration variables. The total gross income can be divided into four components: income of employees, income of self-employed, social income and other income. Any other sufficiently reliable source of household income is not available. The only part of income that can be reliably compared with the external source (administrative source) is the social income.

Table 25 Social income – comparison with administrative sources (Ministry of Labour and Social Affairs) – in million CZK

	EU-SILC 2006	Administrative source	Ratio*
Total social income	311 514	317 556	98.1
Sickness benefits	13 478	32 774	41.1
Pensions (all)	257 300	261 464	98.4
Unemployment benefits	5 719	7 308	78.3
Child benefits	11 123	11 033	100.8
Parental allowances	13 345	13 526	98.7
Housing allowances	2 182	2 287	95.4

* (EU-SILC/Administrative source)*100

The other income components except to social income can be only compared to national accounts for household sector. Comparison of the aggregated income from this survey with the household sector aggregates of the national accounts (even after their modification taking into account the items, which are not covered by household income surveys) is relatively difficult. Concerning its aggregated value the income obtained by direct questioning in the households will always be lower. The more important fact for evaluation of their credibility is that the trend in development of household income is in line with the trends in the national accounts. From this viewpoint, the presented results of SILC 2007 are in full agreement with data from the previous year and with related statistics from developed nations of the European Union.

Table 26 Income – comparison with national accounts – in million CZK

	EU-SILC 2006	National Accounts*	Ratio**
Income of employees	858 998	1 013 832	84.7
Income of self-employed	214 352	252 972	84.7
Total gross income	1 435 704	1 516 1147***	93.5
Total net income	1 203 297	1 441 284***	83.5

* Preliminary results

** (EU-SILC/National Accounts)*100

***Excluding imputed rent