



# EU-SILC 2005 Operation

*Final quality report*

*Czech Republic*



20.12.2007

## **TABLE OF CONTENTS**

<b>1. Common longitudinal indicators .....</b>	<b>4</b>
<b>2. Accuracy.....</b>	<b>4</b>
2.1 Sample design .....	4
2.1.1 Type of sampling.....	4
2.1.2 Sampling units.....	4
2.1.3 Stratification criteria.....	5
2.1.4 Sample size and allocation criteria.....	5
2.1.5 Sample selection schemes .....	5
2.1.6 Sample distribution over time .....	5
2.1.7 Renewal of the sample: Rotational groups.....	5
2.1.8 Weightings .....	6
2.1.8.1 Design factor .....	6
2.1.8.2. Non-response adjustments.....	6
2.1.8.3. Adjustments to external data (level, variables used and sources).....	6
2.1.8.4 Final household cross-sectional weight.....	7
2.1.9 Substitutions .....	7
2.2 Sampling errors .....	7
2.3 Non-sampling errors.....	9
2.3.1 Sampling frame and coverage errors.....	9
2.3.2 Measurement and processing errors.....	9
2.3.3 Non-response errors .....	11
2.3.3.1 Achieved sample size .....	11
2.3.3.2 Unit non-response .....	12
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).....	13
2.3.3.4. Distribution of persons for membership status (RB110) .....	14
2.3.3.5 Item non-response – overview for income variables.....	14
2.4 Mode of data collection.....	16
2.5 Imputation procedure .....	17
2.6 Imputed rent .....	17
2.7 Company cars.....	17
<b>3. Comparability.....</b>	<b>17</b>
3.1 Basic concepts and definitions .....	17
3.2 Components of income.....	17
3.2.1 Differences between the national definitions and standard EU-SILC definitions ..	17
3.2.2 The source or procedure used for collection of income variables.....	18
3.2.3 The form in which income variables at component level have been obtained.....	18
3.2.4 The method used for obtaining the income target variables in required form .....	19
3.3 Tracing rules.....	20
<b>4. Coherence.....</b>	<b>20</b>
4.1 Comparison of income target variables and number of persons with external sources .	20

## ***TABLE OF CONTENTS***

Table 1 Mean, number of observations and standard errors for income components.....	7
Table 2 Sample size.....	11
Table 3 Regional disparities in non-response .....	12
Table 4 Distribution of households by ‘record of contact at address’ (DB120) .....	14
Table 5 Distribution of address contacted by ‘household questionnaire result’ (DB130, DB135).....	14
Table 6 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.....	15
Table 7 Distribution of household members by type of interview (RB260).....	17
Table 8 Overview of the collection of income data (net/gross values).....	18
Table 9 Social income – comparison with administrative sources (Ministry of Labour and Social Affairs) – in million CZK.....	20
Table 10 Income – comparison with national accounts – in million CZK .....	20

## **1. Common longitudinal indicators**

2005 was the initial year of EU-SILC, therefore we do not present any longitudinal indicators.

## **2. Accuracy**

### ***2.1 Sample design***

#### **2.1.1 Type of sampling**

The survey was carried out on the whole territory of the Czech Republic. The sample size was 7 000 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 CSU, 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 with 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**

Sample size for the 2005 survey was mainly dictated by the available fieldwork capacity in terms of human resources and financing. The sample size was 7 000 dwellings. The 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 restricted only to the geographical information obtainable from the sampling frame, the possibilities for modelling using propensity to response models was quite limited. Therefore, calibration 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 (level, variables used and sources)*

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:

- age groups 0-15, 16-24, 25-34, 35-44, 45-54, 55-64, 65+ - based on the demographic statistics)
- gender at the national level (based on the demographic statistics)
- 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 institutionalised population (from social security administrative data) and persons in prisons.

#### 2.1.8.4 Final household cross-sectional weight

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

	N	Minimum	Maximum	Mean	Std. Dev.
Weights DB090	4351	100	3118	922	425

The number of cross-sectional weights (number of DB090 > 0 is 4354) differs from the number of successfully interviewed households by 3. There can be more than one household in the dwelling and in three cases occurred that the second household in the dwelling refused the interview while the first was successfully interviewed. Since the calibration is performed at the dwelling level, the second household get non-zero weight although the interview was refused. Nevertheless the number of successfully interviewed households is 4351.

#### 2.1.9 Substitutions

Substitutions were not used.

#### 2.2 Sampling errors

**Table 1** Mean, number of observations and standard errors for income components\*

	Mean	Number of observations		Standard error
		Before imputation	After imputation	
<b>Income components</b>				
Total disposable household income (HY020)	260336	4349	4351	4000.8
Total disposable household income before social transfers other than old-age and survivor's benefits (HY022)	236289	4268	4270	4113.7
Total disposable household income before social transfers including old-age and survivor's benefits (HY023)	187984	3638	3640	4265.3
<b>Net income components at household level</b>				
Income from rental of a property or land (HY040N)	1467	179	179	396.9
Family/Children related allowances (HY050N)	7383	1173	1173	300.8
Social exclusion not elsewhere classified (HY060N)	1904	169	169	262
Housing allowances (HY070N)	622	253	253	69.7
Regular inter-household cash transfer received (HY080N)	2083	319	319	204

Income received by people aged under 16 (HY110N)	1	2	2	0.9
Regular inter-household cash transfer paid (HY130N)	1016	220	220	105.9
<b>Gross income components at household level</b>				
Income from rental of a property or land (HY040G)	1725	179	179	466.9
Interest, dividends ... (HY090G)	2165	732	732	405.4
Income received by people aged under 16 (HY110G)				
<b>Net income components at personal level</b>				
Employee cash or near cash income (PY010N)	67894	4123	4133	1296.1
Contributions to individual private pension plans (PY035N)	1731	2779	2782	54.5
Value of goods produced by own-consumption (PY070N)	882	1582	1582	60.5
Pension from individual private plans (PY080N)	114	49	49	27.5
Unemployment benefits (PY090N)	897	345	348	70.6
Old-age benefits (PY100N)	21501	2518	2518	603.8
Survivor' benefits (PY110N)	1518	733	733	86.8
Sickness benefits (PY120N)	1502	547	547	92.1
Disability benefits (PY130N)	4019	607	607	20.4
Education-related allowances (PY140N)	91	120	120	20.4
<b>Net income components at personal level</b>				
Employee cash or near cash income (PY010G)	89128	4123	4133	1815.8
Cash benefits or losses from self-employment (PY050G)	22893	637	641	1753.1

	Mean	Number of observations		Standard error
		Before imputation	After imputation	
<b>Equivalised disposable income</b>				
Subclasses by household size:				
1 household member	128268	1228	1228	4223
2 household members	157580	1469	1469	2912
3 household members	169056	730	731	4872
4 and more	150277	922	923	3186
Population by age group:				
<25	147047	2881	2881	3165
25-34	166976	1501	1501	3512
35-44	157381	1257	1257	5224
45-54	172160	1493	1493	4240
55-64	160052	1487	1487	2704
65+	126099	1714	1714	1570
Population by sex:				
Male	157846	4916	4916	2300
Female	150657	5417	5417	2178

\*Imputation on household level means imputed income just for some household members.



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 7 000 sampled dwelling unit records, 353 were found to be ineligible for the survey (5 %). 5 addresses were not located in the field and in 348 cases address did not exist, was non-residential or not occupied. Fieldwork staff undertaking **pre-fieldwork** identification of sampled dwelling units and interviewers must declare clear confirmation of the fact, that the dwelling unit is in fact non-residential or unoccupied. In case of doubts or no information on the status of the dwelling, the case was assumed to be eligible for the survey and coded as non-contact.

### **2.3.2 Measurement and processing 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 co-ordinators 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 four 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 units<sup>1</sup> 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 31.12.2004 (i.e. persons born in 1988 and earlier). This questionnaire contained information on labour status and employment, personal income, participation in private pension plans, health, education and selected biographical information.

Questionnaire CM (SILC Module 2005): questions of the module on the intergenerational transmission of poverty.

### **Reference periods**

- Age: 31.12.2004
- 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 2004
- 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

### **Data processing**

Data were captured using OCR technology (scanning). After the data collection in the field, the questionnaire material is gathered by the regional fieldwork staff. While accepting the material from each interviewers, 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. CSU has three specialised scanning units with technical equipment and expertise 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.

---

<sup>1</sup> 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.

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

The initial gross sample contained addresses of 7 000 dwellings. 353 (5%) addresses were unoccupied or not located. Since there was no substitution of these ineligible units, the survey was conducted in 6 647 dwellings. One another household could be considered as ineligible since it was not included in the sample and this household is involved in other reasons of non-response. There were 68 additional interviewed households in these dwellings, since in some cases 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 2:

**Table 2** Sample size

	Number	Share
<b>Gross sample size:</b>	<b>7 000</b>	<b>100,0 %</b>
Ineligible addresses	353	5,0 %
<b>Dwellings included in the survey:</b>	<b>6 647</b>	<b>100,0 %</b>
Dwellings successfully interviewed:	4 283	64,4 %
+ 68 additional households (2 <sup>nd</sup> ,3 <sup>rd</sup> ,4 <sup>th</sup> household in the dwelling)		
<b>Households successfully interviewed:</b>	<b>4 351</b>	
<b>Non-response:</b>	<b>2 364</b>	<b>100,0 %</b>
- Refusals	1 784	75,5 %
- Non-contacts, temporary absent	464	19,6 %
- Incapacity to participate	96	4,1 %
- Other reasons	20	0,8 %

Response rates on regional (NUTS3) level differ from the national average by approximately  $\pm 10$  percentage points:

**Table 3** Regional disparities in non-response

Region (NUTS3)	HHs (total)	Successfully interviewed		Region (NUTS3)	HHs (total)	Successfully interviewed	
		count	%			count	%
City of Prague	917	469	51,1	Královéhradecký	364	229	62,9
Středočeský	721	459	63,7	Pardubický	304	207	68,1
Jihočeský	396	249	62,9	Vysočina	317	233	73,5
Plzeňský	375	275	73,3	Jihomoravský	708	425	60,0
Karlovarský	193	118	61,1	Olomoucký	414	308	74,4
Ústecký	560	362	64,6	Zlínský	358	241	67,3
Liberecký	272	174	64,0	Moravskoslezský	815	602	73,9

The lowest achieved response rate was in the City of Prague region, slightly above the 50 percent mark. 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. For the remaining regions, the differences between response rates are not large. As in other surveys, the highest response rates were achieved in the Eastern part of the country (Olomoucky, Moravskoslezsky, Vysocina regions). Plzensky region (West Bohemia) is the remaining region with response rate above 70 percent. The other regions have response rates between 60 and 70 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 rates<sup>2</sup>:**

Achieved sample size is 4351.

Number of households for which an interview is accepted for the database: 4351

Number of persons of 16 years or older, who are members of the households and for whom the interview is accepted for the database: 8628

#### *2.3.3.2 Unit non-response*

- **Household non-response rates (NRh)**

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

Where

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

$$Ra = \frac{\text{Number of addresses successfully contacted}}{\text{Number of valid addresses selected}}$$

$$= \frac{\sum[DB120=11]}{\sum[DB120=all] - \sum[DB120=23]} = \frac{6715}{7068 - 348} = 0.99926$$

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

$$= \frac{\sum[DB135=1]}{\sum[DB130=all]} = \frac{4351}{6715} = 0.64795^3$$

$$NRh = (1 - 0.99926 * 0.64795) * 100 = 35.25298$$

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

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

Where

$$Rp = \frac{\text{Number of personal interview completed}}{\text{Number of eligible individuals}} = \frac{8628}{8628} = 1$$

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

So, the individual non-response rate is 0 %

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

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

$$(1 - (0.99926 * 0.64795 * 1)) * 100 = 35.25298$$

So, the overall individual non-response rate is 35.3%

**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)**

DB110 will be provided from 2006 onwards.

---

<sup>3</sup> There were more than one household units in some interviewed dwellings (62 cases, with 68 additional households, out of which 65 were successfully interviewed). These 62 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 some of the non-responding dwellings can also include more than one household unit, the denominator should be higher than 7 068. This difference is unknown, but is likely to be quite small.

**Table 4** Distribution of households by ‘record of contact at address’ (DB120)

	Number	Percentage
<b>Total</b> (DB120 = 11 to 23)	7068	100.00%
Address contacted (DB120 = 11)	6715	95.01%
Address non-contacted (DB120 = 21 to 23)	353	4.99%
<b>Total address non-contacted</b> (DB120 = 21 to 23)	353	100.00%
Address cannot be located (DB120 = 21)	5	1.42%
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)	348	98.58%

**Table 5** Distribution of address contacted by ‘household questionnaire result’ (DB130, DB135)

	Number	Percentage
<b>Total</b>	<b>6715</b>	<b>100.00%</b>
Household questionnaire completed (DB130 = 11)	4351	64.80%
Interview not completed (DB130 = 21 to 24)	2364	35.20%
<b>Total interview not completed (DB130 = 21 to 24)</b>	<b>2364</b>	<b>100.00%</b>
Refusal to co-operate (DB130 = 21)	1784	75.47%
Entire household temporarily away for duration of fieldwork – i.e. non-contacts (DB130 = 22)	464	19.63%
Household unable to respond (illness, incapacity, etc.) (DB130 = 23)	96	4.06%
Other reasons	20	0.85%
<b>Household questionnaire completed (DB135 = 1+ 2)</b>	<b>4351</b>	<b>100.00%</b>
Interview accepted for data base (DB135 = 1)	4351	100.00%
Interview rejected (DB135 = 2)	0	0.00%

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

RB110 will be provided from 2006 onwards.

#### 2.3.3.5 Item non-response – overview for income variables

In table 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 6** 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

<b>Item non-response (overview for different income components)<sup>4</sup></b>	<b>% of households having received an amount</b>	<b>% of households with missing values (before imputation)</b>	<b>% of households with partial information (before imputation)</b>
Total gross household income (HY010)	100.00%	0.00%	0.39%
Total disposable household income (HY020)	100.00%	0.00%	0.39%
Total disposable household income before social transfers except old-age and survivor's benefits (HY022)	98.14%	0.00%	0.39%
Total disposable household income including social transfers except old-age and survivor's benefits (HY023)	83.66%	0.00%	0.39%
<b><i>Net income components at household level</i></b>			
Income from rental of a property or land (HY040N)	4.11%	0.00%	0.00%
Family related allowances (HY050N)	26.96%	0.00%	0.00%
Social exclusion not elsewhere classified (HY060N)	3.88%	0.00%	0.00%
Housing allowance (HY070N)	5.81%	0.00%	0.00%
Regular inter-household cash transfer received (HY080N)	7.33%	0.00%	0.00%
Income received by people aged < 16 (HY110N)	0.05%	0.00%	0.00%
Regular taxes on wealth (HY120N)	44.84%	0.00%	0.00%
Regular inter-household cash transfer paid (HY130N)	5.06%	0.00%	0.00%
Tax on income and social contributions (HY140N)	68.17%	0.00%	0.00%
<b><i>Gross income components at household level</i></b>			
Income from rental of a property or land (HY040G)	4.11%	0.00%	0.00%
Family related allowances (HY050G)	26.96%	0.00%	0.00%
Social exclusion not elsewhere classified (HY060G)	3.88%	0.00%	0.00%
Housing allowance (HY070G)	5.81%	0.00%	0.00%
Regular inter-household cash transfer received (HY080G)	7.33%	0.00%	0.00%
Interests, dividends, etc. (HY090G)	16.82%	0.00%	0.00%
Interest repayments on mortgage (HY100G)	7.70%	0.00%	0.00%
Regular taxes on wealth (HY120G)	44.84%	0.00%	0.00%
Regular inter-household cash transfer paid (HY130G)	5.06%	0.00%	0.00%
Tax on income and social contributions (HY140G)	68.17%	0.00%	0.39%

<sup>4</sup> 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)
<b><i>Net income components at personal level</i></b>			
Employee cash or near cash income (PY010N)	47.90%	0.12%	0.00%
Contributions to individual private pension plans (PY035N)	32.24%	0.03%	0.00%
Value of goods produced by own-consumption (PY070N)	18.34%	0.00%	0.00%
Pension from individual private plans (PY080N)	0.57%	0.00%	0.00%
Unemployment benefits (PY090N)	4.03%	0.03%	0.00%
Old age benefits (PY100N)	29.18%	0.00%	0.00%
Survivor' benefits (PY110N)	8.50%	0.00%	0.00%
Sickness benefits (PY120N)	6.34%	0.00%	0.00%
Disability benefits (PY130N)	7.04%	0.00%	0.00%
Education-related allowances (PY140N)	1.39%	0.00%	0.00%
<b>Gross income components at personal level</b>			
Employee cash or near cash income (PY010G)	47.90%	0.12%	0.00%
Non cash employee income (PY020G)	1.74%	0.01%	0.00%
Contributions to individual private pension plans (PY035G)	32.24%	0.03%	0.00%
Cash benefits or losses from self-employment (PY050G)	7.43%	0.05%	0.00%
Value of goods produced by own-consumption (PY070G)	18.34%	0.00%	0.00%
Pension from individual private plans (PY080G)	0.57%	0.00%	0.00%
Unemployment benefits (PY090G)	4.03%	0.03%	0.00%
Old age benefits (PY100G)	29.23%	0.00%	0.00%
Survivor' benefits (PY110G)	8.50%	0.00%	0.00%
Sickness benefits (PY120G)	6.34%	0.00%	0.00%
Disability benefits (PY130G)	7.04%	0.00%	0.00%
Education-related allowances (PY140G)	1.39%	0.00%	0.00%

## ***2.4 Mode of data collection***

### **Distribution of household members by data status (RB250)**

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).

Registers are not used at all. Due to strict definition of response, there are any “not completed interviews” at individual level (all cases were filled as proxy).



**Table 7** Distribution of household members by type of interview (RB260)

Method	Count	%
Face-to-face with paper questionnaire	7759	89,9%
Face-to-face with computer (CAPI)	not used	-
Telephone interviews (CATI)	not used	-
Self administered questionnaire	65	0,8%
Proxy face-to-face interview (information from another household member)	804	9,3%
Total	<b>8628</b>	<b>100,0%</b>

## *2.5 Imputation procedure*

Situation of **missing income data** for one of the household members was relatively rare (18 cases). For these persons, the income was **imputed by the simple hot-deck method** (using randomly chosen person with similar characteristics from another household).

## *2.6 Imputed rent*

We do not calculate imputed rent in 2005.

## *2.7 Company cars*

The lowest possible amount applicable for taxation in the tax law is added to the non-monetary income of the employee (CZK 1000/month).

## **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: four to five months (the survey took place in April and May 2005)
- The total duration of the data collection of the sample: 6 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 8** Overview of the collection of income data (net/gross values)<sup>5</sup>

income component	% collected net of taxes and social contributions	% collected gross <sup>6</sup>
PY010G	47,5%	52,5%
PY010N	47,5%	52,5%
PY020G	0,0%	100,0%
PY020N	-	-
PY035G	100,0%	0,0%
PY035N	100,0%	0,0%
PY050G	16,2%	83,8%
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%
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%

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

<sup>6</sup> 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.

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**

One source of potential 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**).

Underestimation of income is a natural consequence of the fact, that respondents either tend to give lower than actual values or simply did not recall certain irregular or small incomes. Previous experience from Microcensus income surveys had shown the underestimation of about 10 %, but with varying degree dependent on the level and source of income. 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). Compared to the previous findings, the average underestimation from this comparison was negligible (2.8 %). Having in mind the limited available number of cases, if broken into subgroups by sectors, the decision has been made not to apply any corrections on the wage data. In case of self-employment income, detailed analysis have shown that in some cases the reported gross income from self-employment were in fact most likely revenues (before deducting the costs). In this case, the disproportionally high gross income values were substituted by gross amount from modelled relation between gross profits, net profits and collected amount of paid social insurance.

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 2004 (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).

### 3.3 Tracing rules

Tracing rules were not applied since the EU-SILC 2005 was the first wave.

## 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 which can be reliably compared with the external source (administrative source) is the social income.

**Table 9** Social income – comparison with administrative sources (Ministry of Labour and Social Affairs) – in million CZK

	EU-SILC 2005	Administrative source	Ratio*
<b>Total social income</b>	<b>277 621</b>	<b>287 150</b>	<b>96,7</b>
Sickness benefits	12 737	29 563	43,1
Pensions	225 493	229 538	98,2
Unemployment benefits	7 611	6 994	108,8
Child benefits	14 816	15 664	94,6
Parental allowances	9 753	10 425	93,6
Housing allowances	2 497	2 548	98,0

\* (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 income aggregates of the national accounts (even after their modification taking into account the items, which are not covered by household income surveys) is problematic. Concerning its aggregated value the income based on household survey data 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.

**Table 10** Income – comparison with national accounts – in million CZK

	EU-SILC 2005	National Accounts	Ratio*
Income of employees	755 931	895 108	84,5
Income of self-employed	193 698	243 407	79,6
<b>Total gross income</b>	<b>1 269 973</b>	<b>1 370 8300**</b>	<b>92,6</b>
Total net income	1 053 273	1 386 870	76,8

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

\*\*Excluding imputed rent