

- Full view -

**SILC\_ESQRS\_A\_CZ\_2017\_0000**

National Reference Metadata in ESS Standard for Quality Reports Structure (ESQRSSI)

Compiling agency: Czech Statistical Office

Time Dimension: 2017-A0

Data Provider: CZ1

Data Flow: SILC\_ESQRS\_A

**Eurostat metadata**

## Reference metadata

[1. Contact](#)  
[2. Statistical presentation](#)  
[3. Statistical processing](#)  
[4. Quality management](#)  
[5. Relevance](#)  
[6. Accuracy and reliability](#)  
[7. Timeliness and punctuality](#)  
[8. Coherence and comparability](#)  
[9. Accessibility and clarity](#)  
[10. Cost and Burden](#)  
[11. Confidentiality](#)  
[12. Comment](#)  
[Related Metadata](#)  
[Annexes](#) (including footnotes)

For any question on data and metadata, please contact: [EUROPEAN STATISTICAL DATA SUPPORT](#)

1. Contact <a href="#">Top</a>	
1.1. Contact organisation	Czech Statistical Office
1.2. Contact organisation unit	Social Surveys Unit Household Surveys Department
1.3. Contact name	Tana Dvornakova
1.4. Contact person function	Methodology
1.5. Contact mail address	Na padesatem 81, 100 82 Praha 10, Czech Republic
1.6. Contact email address	tana.dvornakova@czso.cz
1.7. Contact phone number	+420 274 054 360
1.8. Contact fax number	

2. Statistical presentation <a href="#">Top</a>	
2.1. Data description	
Not available. New concept added with the migration to SIMS 2.0. Information (content) will be available after the next collection.	
2.2. Classification system	
Not available. New concept added with the migration to SIMS 2.0. Information (content) will be available after the next collection.	
2.3. Coverage - sector	
Not available. New concept added with the migration to SIMS 2.0. Information (content) will be available after the next collection.	
2.4. Statistical concepts and definitions	
Total hh gross income (HY010) - F Total disposable hh income (HY020) - F Total disposable hh income before social transfers other than old-age and survivors' benefits (HY022) - F Total disposable hh income before all social transfers (HY023) - F  Imputed rent (HY030) - P Income from rental of property or land (HY040) - F Family/ Children related allowances (HY050) - F Social exclusion payments not elsewhere classified (HY060) - F Housing allowances (HY070) - F Regular inter-hh cash transfers received (HY080) - F Interest, dividends, profit from capital investments in incorporated businesses (HY090) - F Interest paid on mortgage (HY100) - F Income received by people aged under 16 (HY110) - NC Regular taxes on wealth (HY120) - F Regular inter-hh transfers paid (HY130) - F Value of goods produced for own consumption (HY170) - F  Cash or near-cash employee income (PY010) - F	

Other non-cash employee income (PY020) - F  
Income from private use of company car (PY021) - F  
Employers social insurance contributions (PY030) - F  
Cash profits or losses from self-employment (PY050) - F  
Unemployment benefits (PY090) - F  
Old-age benefits (PY100) - F  
Survivors benefits (PY110) - F  
Sickness benefits (PY120) - F  
Disability benefits (PY130) - F  
Education-related allowances (PY140) - F  
Gross monthly earnings for employees (PY200) - NC

#### The source or procedure used for the 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.

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

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.

See Annex 8 Overview of the collection of income data (net/gross values), cross sectional sample 2017.

#### The method used for obtaining target variables in the required form

Situation of missing income data for one of the household members was 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).

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).

Underestimation of income is a natural consequence of the fact, those 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. 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 2016 (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).

#### Annexes:

[Annex 8 Collection of income data C17](#)

#### 2.5. Statistical unit

Not available.

New concept added with the migration to SIMS 2.0.

Information (content) will be available after the next collection.

#### 2.6. Statistical population

Not available.

New concept added with the migration to SIMS 2.0.

Information (content) will be available after the next collection.

#### 2.7. Reference area

Not available.

New concept added with the migration to SIMS 2.0.

Information (content) will be available after the next collection.

#### 2.8. Coverage - Time

Not available.

New concept added with the migration to SIMS 2.0.

Information (content) will be available after the next collection.

#### 2.9. Base period

Not available.

New concept added with the migration to SIMS 2.0.

Information (content) will be available after the next collection.

### 3. Statistical processing

[Top](#)

Detailed information concerning sampling frame, sampling design, sampling units, sampling size, weightings and mode of data collection can be found in this section. Such information is mainly used for the computation of the accuracy measures.

#### 3.1. Source data

**Reference periods**

- Age: December 31, 2016.
- Other demographic variables - marital status, education, housing, financial situation: the date of the interview.
- Work activity was collected for each month of 2016 as well as currently. Work activity figures are gathered by self-definition of the respondent (respondents themselves choose among different types of activity the one that fits them the most). Its value primarily depends on the respondent's main occupation and on the time spent in it. Subsequently, other data related to the respondent's work activity (status at work, profession) were collected as of the date of the interview. Parallel activities were surveyed (second job, study), together with data on receipt of pensions and parental benefits.
- Economic activity was not collected but derived from the monthly/yearly data (if monthly data was the basis, the activity with the highest incidence was coded as the yearly value). For those who completed their education in 2016 the latter half of the year was considered.
- Income data (both monetary and in kind): calendar year 2016.
- Subjective questions focused on housing and financial problems: the date of the interview.
- Health problems: last six or twelve months.
- Housing, consumer durables, financial and social situation of household: the date of the interview, unless the question specifically refers to some other period.

**3.1.1. Sampling design and procedure****Type of sampling design**

The survey was carried out on the whole territory of the Czech Republic. The sample size of newly selected dwelling (first wave in 2017) was 4 750 dwellings. Dwellings were selected using stratified two-stage sampling design. At the first sampling stage small geographical areas (CEUs – Census Enumeration Units) were selected by probability sampling. In the second stage selection a sample of 10 dwellings was drawn from each CEU.

**Sample size and allocation criteria**

In 2017, the total sample size was 10 866 dwellings (10 904 households) from which 4 750 addresses were newly selected, 6 710 dwellings were revisited from previous waves and 150 dwellings into which from the previous year moved some sample persons. The new sample was allocated to the strata using proportional algorithm (proportionally to the number of dwellings in the sampling frame).

**Stratification and sub stratification criteria**

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

- below 2 000 inhabitants
- 2 000 – 9 999 inhabitants
- 10 000 – 49 999 inhabitants
- 50 000 and more inhabitants

**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.

**Sample distribution over time**

Due to the limited duration of the fieldwork period, the survey was organized as a one-short survey. The fieldwork started on the 4th February and ended on the 14th May (PAPI) or on the 28th May (CAPI). Sample was not distributed into separate waves over the duration of the fieldwork.

**3.1.2. Sampling unit****Sampling unit**

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 subtenants 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.

**3.1.3. Sampling rate and sampling size****Sampling rate and sampling size**

Concerning the SILC instrument, three different sample size definitions can be applied:

- the actual sample size which is the number of sampling units selected in the sample: 11 610 (without ineligible units: 10 866)
- the achieved sample size which is the number of observed sampling units (household or individual) with an accepted interview
  - accepted household interviews: 8 701
  - accepted individual interviews: 16 286
- the effective sample size which is defined as the achieved sample size divided by the design effect with regards to the at-risk-of poverty rate indicator: 4 750

Accepted interviews by waves

	Total	1 <sup>st</sup> wave	2 <sup>nd</sup> wave	3 <sup>rd</sup> wave	4 <sup>th</sup> wave
Accepted household interviews	8 701	2 478	2 357	1 990	1 876
Accepted personal interviews	16 286	4 618	4 462	3 702	3 504
Number of persons aged 16 years and older	16 286	4 618	4 462	3 702	3 504
Sample persons	16 131		4 438	3 660	3 415
Co-resident	155		24	42	89

Cross-sectional sample 2017, Longitudinal 2014-2017

Achieved household sample size, sub-sample

2016-17	2015-16-17	2014-15-16-17
6 182	3 818	1 847

Longitudinal 2014-2017

Achieved individual sample size

2016-17				2015-16-17				2014-15-16-17			
All present	16+ present	Sample person present	Co-resident present	All present	16+ present	Sample person present	Co-resident present	All present	16+ present	Sample person present	Co-resident present
14 014	11 474	12 025	1 989	8 488	6 977	6 977	1 135	4 103	3 352	3 566	537

Longitudinal 2014-2017

#### Achieved sample size

4 750 new dwellings entered the survey (1st wave) and 6 860 dwellings were revisited – 6 710 at the last year's address and 150 were tracked to their new home. The fieldwork revealed that among the total of 11 610 dwellings in the sample there were 744 dwellings (6.4 %) unoccupied, unlocated or ineligible because the households had moved. Since there was no substitution for these ineligible units, the survey was conducted in 10 866 dwellings and 10 904 households. There were 38 additional interviewed households in these dwellings, since in 38 dwellings there are more households in one dwelling unit (household definition is based on sharing of expenses).

See annex 9 Sample size - which include overview of sample size of household by waves

Participation in the sample survey is voluntary; unlike the population census, households were not obliged to provide any information. A selected household has to be informed about the content of the survey and about the fact that their participation in the survey is voluntary. Whether to respond or not is left to the household's own deliberation. The main reasons for refusal are privacy reasons (objections against giving personal information and fear of abuse 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.

Refusals instead unwillingness to give information 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.

The lowest achieved response rate was in the Capital city Praha region, about 74 percent. On the other hand, the highest response rate, around 85 percent was in Jihocesky region. For the remaining regions, the differences between response rates are not large.

See annex 10 Regional disparities in response

#### Annexes:

[Annex 9 Sample size C17](#)

[Annex 10 Regional disparities in response C17](#)

#### 3.2. Frequency of data collection

##### Fieldwork

Data collection lasted from February 4th to May 14th (PAPI) or to May 28th (CAPI) 2017.

##### Renewal of sample: rotational groups

The survey uses 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. The rotational scheme with four replications was begun in 2008. In 2009 first rotational panel was ended and the household from the 2005 operation was dropped from the sample. In 2017, households from the 2013 operation were dropped from the sample. Each next year, one sub-sample rotates out and a new one is drawn and substituted for.

The longitudinal dataset contains households sampled from 2014 (first interviews), 2015 (second interviews), 2016 (third interviews) and 2017 (fourth interviews).

new in 2014				
2014	wave 1	new in 2015		
2015	wave 2	wave 1	new in 2016	
2016	wave 3	wave 2	wave 1	new in 2017
2017	wave 4	wave 3	wave 2	wave 1

Longitudinal sample: 2014 - 2017      2015 - 2017      2016 - 2017

#### 3.3. Data collection

##### Mode of data collection

The data collection methods were PAPI (Paper Assistance Personal Interview) around 37 percent, and CAPI (Computer Assistance Personal Interview) around 34 percent. Each questionnaire was filled during fact-to-face interview with the interviewer. PAPI contain interview that was carried out by PAPI and then feed into electronic questionnaire. Some personal questionnaires were filled as proxy interviews (29 percent) – information about household member was not present at the time of the interview was provided by another household member.

1-PAPI	2-CAPI	6-proxy PAPI	7-proxy CAPI
(% of total)	(% of total)	(% of total)	(% of total)
37.0	33.8	12.3	17.0

Cross-sectional sampling 2017

See Annex 11 Mode of data collection by type of interview and by waves, cross-sectional sample 2017, longitudinal sample 2014-2017

#### Annexes:

[Annex 11 Mode of data collection](#)

<b>3.4. Data validation</b>												
Not requested by Reg. 28/2004												
<b>3.5. Data compilation</b>												
Not requested by Reg. 28/2004												
<b>3.5.1. Weighting procedure</b>												
<b>Design factor</b> The sample was designed as a self-weighting sample within the region. Due to the exact numbers of CEUs, the design factor for all sampled dwellings is in the range of 0.94 to 1.09. The CEU are determined to keep the self-weighting principle as much as possible. This change took place in 2016.												
<b>Non-response adjustments</b> 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 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 over-representation 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.												
<b>Adjustment to external data</b> 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 and 2011 Census continuously updated from administrative sources of construction authorities Population characteristics: <ul style="list-style-type: none"><li>Population totals in each NUTS3 region (from demographic statistics)</li><li>Economic activity characteristics for the Czech Republic<ul style="list-style-type: none"><li>Number of employees - derived from the number of employees in the economy based on the Labour Force Survey (LFS) results and company reporting</li></ul></li><li>Demographic characteristics at the national level (based on the demographic statistics):<ul style="list-style-type: none"><li>Age groups (0-15, 16-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75+)</li><li>Gender at the national level</li><li>Municipality size at the national level (below 2 000 inhabitants, 2 000 - 9 999, 10 000 - 49 999, 50 000+ inhabitants)</li></ul></li></ul> The target population of the survey was persons living in private households, the demographic statistics aggregate data were adjusted by subtracting institutionalised population (from social security administrative data and Ministry of Justice) and the persons living outside dwellings as based on the 2011 Census. In longitudinal data files was calibration done for DB095 in case of 2017 data. In the first wave, the longitudinal base weights (RB060) are identical to the cross-sectional weights. The longitudinal weight RB062 was derived from RB060. Because the sum of RB062 weights should be equal the size of the longitudinal population of individuals in scope for the four last waves, the weights was multiplied by ratio of longitudinal and cross-sectional population. The longitudinal population 2014-2017 differs from 2014, 2015, 2016 and 2017 population for died and moved abroad people.												
<b>Final cross sectional weights</b> Final household cross-sectional weight was result of Calmar calibration. <table><tr><td></td><td>N</td><td>Minimum</td><td>Maximum</td><td>Mean</td><td>Std. Dev.</td></tr><tr><td>Weights DB090</td><td>8 701</td><td>226.0</td><td>2 189.7</td><td>502.5</td><td>211.0</td></tr></table> The number of cross-sectional weights (number of DB090 > 0 is 8 507) is the same as the number of successfully interviewed households (DB130 = 11 is 8 701).		N	Minimum	Maximum	Mean	Std. Dev.	Weights DB090	8 701	226.0	2 189.7	502.5	211.0
	N	Minimum	Maximum	Mean	Std. Dev.							
Weights DB090	8 701	226.0	2 189.7	502.5	211.0							
<b>3.5.2. Estimation and imputation</b>												
<b>Imputation procedure used</b> Situation of missing income data for one of the household members was rare (18 cases) in 2017. For these persons, the income was imputed by the simple hot-deck method (using randomly chosen person with similar characteristics from another household). Access to administrative register information on individual level is not possible. We use our developed model for gross/net conversion, which was developed with regard to the Czech tax laws. Deductive imputation took place within the frame of logical checks. Regional staff is responsible for checking of the data for their respective region, using a special software application containing a set of logical checks, captured data and linked images of the questionnaires. The item non-response of non-income-variables is rare, so model approach development is useless. We use hot-deck method for new households and information from last year for households in next waves of survey.												
<b>Imputed rent</b> Since 2011 there was administrative transition from rent regulated by government to market rent. In 2017, share of households paying market rent was 15.2 % and share of households paying rent at a reduced rate was 1.8 %. Regulated rent was definitely cancelled in 2013. In the dataset 2017 the rent at a reduced rate means rent based on legislation; it is used only for apartments with low quality of accommodation. At the beginning of transition, the prices of rent that changed from regulated to market rent were almost the same but during the time increased. We tested 3 methods (subjective method, stratification method, Heckman model) for computing rent and finally we decided for subjective method, because it seemed best in the Czech conditions. Respondents were asked to estimate the price for which their dwelling could be sold. Subsequently, the market rent is derived. The advantage of this method lies in its simplicity but this is substantially outweighed by its drawback - the fallibility of responded values due to lack of knowledge of housing market of the respondents. The values can be overestimated as well as underestimated, depending on how the household is informed about the current market prices.												
<b>Company car</b> The amount of CZK 3000 was added to income in kind of an employee for each month of using a company car.												
<b>3.6. Adjustment</b>												
Not requested by Reg. 28/2004												

<b>4. Quality management</b>	<a href="#">Top</a>
<b>4.1. Quality assurance</b>	

Not available.  
 New concept added with the migration to SIMS 2.0.  
 Information (content) will be available after the next collection.

#### 4.2. Quality management - assessment

Not requested by Reg. 28/2004

### 5. Relevance

[Top](#)

#### 5.1. Relevance - User Needs

Not requested by Reg. 28/2004

#### 5.2. Relevance - User Satisfaction

Not requested by Reg. 28/2004

#### 5.3. Completeness

Not requested by Reg. 28/2004

##### 5.3.1. Data completeness - rate

Not requested by Reg. 28/2004

### 6. Accuracy and reliability

[Top](#)

The concept of accuracy refers to the precision of estimates computed from a sample rather than from the entire population. Accuracy depends on sample size, sampling design effects and structure of the population under study. In addition to that, sampling errors and non-sampling errors need to be taken into account. Sampling error refers to the variability that occurs at random because of the use of a sample rather than a census and non-sampling errors are errors that occur in all phases of the data collection and production process.

#### 6.1. Accuracy - overall

In terms of precision requirements, the EU-SILC framework regulation as well the Commission Regulation on sampling and tracing rules refers respectively, to the effective sample size to be achieved and to representativeness of the sample. The effective sample size combines sample size and sampling design effect which depends on sampling design, population structure and non-response rate.

#### 6.2. Sampling error

EU-SILC is a complex survey involving different sampling design in different countries. In order to harmonize and make sampling errors comparable among countries, Eurostat (with the substantial methodological support of Net-SILC2) has chosen to apply the "linearization" technique coupled with the "ultimate cluster" approach for variance estimation. Linearization is a technique based on the use of linear approximation to reduce non-linear statistics to a linear form, justified by asymptotic properties of the estimator. This technique can encompass a wide variety of indicators, including EU-SILC indicators. The "ultimate cluster" approach is a simplification consisting in calculating the variance taking into account only variation among Primary Sampling Unit (PSU) totals. This method requires first stage sampling fractions to be small which is nearly always the case. This method allows a great flexibility and simplifies the calculations of variances. It can also be generalized to calculate variance of the differences of one year to another.

The sampling design was assimilated to a two stage stratified type we used DB050 (primary strata) for strata specification and DB060 (Primary Sampling Unit) for cluster specification.

##### 6.2.1. Sampling error - indicators

See Annex 1 which shows the estimates for the leading indicator People at-risk of poverty or social exclusion (AROPE) and its components namely: At-risk of poverty rate, People living in household with very low work intensity and severe material deprivation rate.

The computations were done in SAS programs for variance estimation of the measures. All indicators were calculated at individual level.

See Annex 2 Income components, which show the mean, number of observations (before and after imputation) and standard error.

Comment for variable HY090G - Since 2016 there has been one household with very high stocks income, without this household the average is 18 000 CZK.

#### Annexes:

[Annex 1 Sampling error](#)

[Annex 1 Sampling error - income components](#)

#### 6.3. Non-sampling error

Non-sampling errors are basically of 4 types:

- Coverage errors: errors due to divergences existing between the target population and the sampling frame.
- Measurement errors: errors that occur at the time of data collection. There are a number of sources for these errors such as the survey instrument, the information system, the interviewer and the mode of collection
- Processing errors: errors in post-data-collection processes such as data entry, keying, editing and weighting
- Non-response errors: errors due to an unsuccessful attempt to obtain the desired information from an eligible unit. Two main types of non-response errors are considered:

– Unit non-response: refers to absence of information of the whole units (households and/or persons) selected into the sample

– Item non-response: refers to the situation where a sample unit has been successfully enumerated, but not all required information has been obtained

##### 6.3.1. Coverage error

Coverage errors include over-coverage, under-coverage and misclassification:

- Over-coverage: relates either to wrongly classified units that are in fact out of scope, or to units that do not exist in practice
- Under-coverage: refers to units not included in the sampling frame
- Misclassification: refers to incorrect classification of units that belong to the target population

##### 6.3.1.1. Over-coverage - rate

	Main problems	Size of error
Cross sectional data	<ul style="list-style-type: none"> <li>• Over-coverage</li> </ul>	9.8 %

#### 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 4 750 newly sampled dwelling unit records (in the first wave), 468 were found to be ineligible for the survey (9.8 %). 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.

##### 6.3.1.2. Common units - proportion

Not requested by Reg. 28/2004

### 6.3.2. Measurement error

#### Source of measurement errors

It might occur as a consequence of many reasons, mostly of inaccurate methodological instructions, not respecting them by interviewers, wrong wording of questions, processing mistakes, unwillingness to participate in the survey or giving purposely biased answers.

#### Building process of questionnaire

Data collection had the form of an interview and interviewers filled in the answers into paper questionnaires (PAPI data collection) or into electronic questionnaires (CAPI data collection). Data from paper questionnaires were then transcribed into CAPI questionnaire by specially trained transcribers.

The survey was conducted using electronic questionnaires with the assistance of programmatic system BLAISE. It is developed Statistics Netherlands and it is standard for questionnaire survey. Since 2008 will be a gradual transition to CAPI data collection.

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 characteristics, information on sharing of expenses to determine household units 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, 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) and childcare.

**Questionnaire C** (personal questionnaire): filled in by each household member aged 16+ as of 31 December 2016 (i.e. persons born in 2000 and earlier). This questionnaire contained information on labour status and employment, personal income, participation in private pension plans, health and selected biographical information.

**Questionnaire BM** (household questionnaire): filled in for each household. Module questions relate to the information about access to services.

See annex 3 Questionnaire 2017.

#### Interview training

Workers from regional departments conducted methodical training of fieldworkers.

A total of 702 interviewers participated in the survey; each interviewer performed an interview in 12,4 households on average. 413 interviewers used electronic questionnaires (CAPI) and each of them successfully examined about 10,7 households. Each of 503 interviewers with paper questionnaires (PAPI) examined about 8,5 households on average. PAPI questionnaires were then transcribed into CAPI questionnaire by specially trained transcribers.

#### Quality control

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).

#### Annexes:

[Annex 3 CZ, Questionnaire 2017](#)

### 6.3.3. Non response error

Non-response errors are errors due to an unsuccessful attempt to obtain the desired information from an eligible unit. Two main types of non-response errors are considered:

1) **Unit non-response** which refers to the absence of information of the whole units (households and/or persons) selected into the sample. According the Commission Regulation 28/2004:

- **Household non-response rates** ( $NRh$ ) is computed as follows:

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

Where  $Ra$  is the address contact rate defined as:

$Ra$  = Number of address successfully contacted/Number of valid addresses selected

and  $Rh$  is the proportion of complete household interviews accepted for the database

$Rh$  = Number of household interviews completed and accepted for database/Number of eligible households at contacted addresses

- **Individual non-response rates** ( $NRp$ ) will be computed as follows:

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

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

$Rp$  = Number of personal interview completed/Number of eligible individuals in the households whose interviews were completed and accepted for the database

- **Overall individual non-response rates** ( $*NRp$ ) will be computed as follows:

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

For those Members States where a sample of persons rather than a sample of households (addresses) was selected, the individual non-response rates will be calculated for 'the selected respondent', for all individuals aged 16 years or older and for the non-selected respondent.

2) **Item non-response** which refers to the situation where a sample unit has been successfully enumerated, but not all the required information has been obtained.

#### 6.3.3.1. Unit non-response - rate

##### Cross sectional data

Address contact rate (Ra)*		Complete household interviews (Rh)*		Complete personal interviews (Rp)*		Household Non-response rate (NRh)*		Individual non-response rate (NRp)*		Overall individual non-response rate (NRp)*	
A*	B*	A*	B*	A*	B*	A*	B*	A*	B*	A*	B*
90.47	90.17	79.80	57.75	100.00	100.00	27.81	47.93	0.00	0.00	27.81	47.93

\* All the formulas are defined in the Commission Regulation 28/2004, Annex II

A\* = Total sample; B\* = New sub-sample

See Annex 4 Distribution of households by 'record of contact at address' (DB120), by 'household questionnaire result' (DB130) and by 'household interview acceptance' (DB135); cross sectional sample 2017

See Annex 5 Distribution of households by 'household status', by 'record of contact at address', by 'household questionnaire result' and by 'household interview acceptance'; Distribution of persons for membership status (RB110); longitudinal sample 2014-2017

#### Annexes:

[Annex 4 Distribution of households C17](#)

[Annex 5 Distribution of households persons L17](#)

#### 6.3.3.2. Item non-response - rate

The computation of item non-response is essential to fulfil the precision requirements concerning publication as stated in the Commission Regulation No 1982/2003. Item non-response rate is provided for the main income variables both at household and personal level.

##### 6.3.3.2.1. Item non-response rate by indicator

See Annex 7 Overview of the non-response for the income variables; cross sectional sample 2017, longitudinal sample 2014-2017

#### Annexes:

[Annex 7 Non-response for income variables C17, L17](#)

##### 6.3.4. Processing error

###### Data processing

PAPI data were transcribed into CAPI by specially trained transcribers. In case of CAPI data were collected into electronic questionnaire with the aid of programming system BLAISE in application eDomset. After the data collection in the field, the regional fieldwork staff takes data file from the questionnaire material. While accepting the data file gathers the questionnaire material from each interviewers, the initial check is performed - the way, how the questionnaires are filled, completeness of the questionnaires, basic consistence checks. After this preparatory phase, data from questionnaires are co-ordinate to general database CZSO.

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).

###### Data entry and coding

In case of CAPI the basic checks are included in questionnaire. The electronic questionnaire has three types of control - longitudinal, warning and binding. Longitudinal control is engaged for checking information between years of the survey. Warning control warn of the suspect value. Binding control point out wrong value and interweavers must fix the value.

In case of PAPI checks on values are made during the transcription into CAPI.

###### Editing controls

No information about the rates of failed edits for income variables and other relevant variables.

##### 6.3.4.1. Imputation - rate

Not requested by Reg. 28/2004

##### 6.3.5. Model assumption error

Not requested by Reg. 28/2004

##### 6.4. Seasonal adjustment

Not requested by Reg. 28/2004

##### 6.5. Data revision - policy

Not requested by Reg. 28/2004

##### 6.6. Data revision - practice

Not requested by Reg. 28/2004

##### 6.6.1. Data revision - average size

Not requested by Reg. 28/2004

## 7. Timeliness and punctuality

[Top](#)

Not requested by Reg. 28/2004

### 7.1. Timeliness

Not requested by Reg. 28/2004

#### 7.1.1. Time lag - first result

Not requested by Reg. 28/2004

#### 7.1.2. Time lag - final result

Not requested by Reg. 28/2004

### 7.2. Punctuality

Not requested by Reg. 28/2004

#### 7.2.1. Punctuality - delivery and publication

Not requested by Reg. 28/2004

## 8. Coherence and comparability

[Top](#)

The coherence of two or more statistical outputs refers to the degree to which the statistical processes, by which they were generated, used the same concepts and harmonised methods. A comparison with external sources for all income target variables and the number of persons who receive income from each 'income component' will be provided, where the Member States concerned consider such external data to be sufficiently reliable.

According to the Regulation (EC) No 1177/2003 of the European Parliament and of the Council concerning EU-SILC: "Comparability of data between Member States shall be a fundamental objective and shall be pursued through the development of methodological studies from the outset of EU-SILC data collection, carried out in close collaboration between the Member States and Eurostat".

Although the best way for keeping the comparability of data is to apply the same methods and definitions of variables, small departures of the definitions given by Eurostat are allowed in EU-SILC. In this way, the mentioned Regulation in its article 16th says: "Small departures from common definitions, such as those relating to private household definition and income reference period, shall be allowed, provided they affect comparability only marginally. The impact of comparability shall be reported in the quality reports."

#### 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 (2016)
- 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: two to five months (the survey took place from February to the end of May 2017)
- The total duration of the data collection of the sample: 14 weeks (PAPI), 17 weeks (CAPI)
- Basic information on activity status during the income reference period: no differences

### 8.1. Comparability - geographical

Not requested by Reg. 28/2004

#### 8.1.1. Asymmetry for mirror flow statistics - coefficient

Not requested by Reg. 28/2004			
<b>8.1.2. Reference population</b>			
Reference population - fully comparable			
Private household definition - fully comparable			
Household membership - fully comparable			
<b>8.1.3. Reference Period</b>			
Period for taxes on income and social insurance contributions - calendar year 2016			
Income reference periods used - calendar year 2016			
Reference period for taxes on wealth - calendar year 2016			
Lag between the income ref. period and current variables - 2 to 5 months			
<b>8.2. Comparability - over time</b>			
It is not relevant.			
<b>8.2.1. Length of comparable time series</b>			
Not requested by Reg. 28/2004			
<b>8.3. Coherence - cross domain</b>			
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.			
Social income – comparison with administrative sources (Ministry of Labour and Social Affairs) – in million CZK			
	EU-SILC 2017	Administrative source	Ratio*
<b>Total social income</b>	<b>464 526</b>	<b>501 378</b>	<b>92.6</b>
Sickness benefits	17 265	26 284	65.7
Pensions (all)	386 804	394 467	98.1
Unemployment benefits (PY090G)	4 548	8 254	55.1
Child benefits	2 887	2 817	102.5
Parental allowances	22 853	22 625	101.0
Housing allowances (HY070G)	7 534	9 261	81.4
* (EU-SILC/Administrative source)*100			
<b>8.4. Coherence - sub annual and annual statistics</b>			
Not requested by Reg. 28/2004			
<b>8.5. Coherence - National Accounts</b>			
The 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 2017 are in full agreement with data from the previous year and with related statistics from developed nations of the European Union.			
Income – comparison with national accounts – in million CZK			
	EU-SILC 2017	National Accounts	Ratio*
Income of employees	1 290 888	1 494 977	86.3
Income of self-employed	276 397	308 521	89.6
<b>Total gross income</b>	<b>2 093 494</b>	<b>2 249 811**</b>	<b>93.1</b>
Total net income	1 783 475	2 115 596**	84.3
* (EU-SILC/National Accounts)*100			
** Excluding imputed rent			
<b>8.6. Coherence - internal</b>			
Not requested by Reg. 28/2004			

<b>9. Accessibility and clarity</b>		<a href="#">Top</a>
Not requested by Reg. 28/2004		
<b>9.1. Dissemination format - News release</b>		
Not requested by Reg. 28/2004		
<b>9.2. Dissemination format - Publications</b>		
Not requested by Reg. 28/2004		
<b>9.3. Dissemination format - online database</b>		
Not requested by Reg. 28/2004		
<b>9.3.1. Data tables - consultations</b>		
Not requested by Reg. 28/2004		
<b>9.4. Dissemination format - microdata access</b>		
Not requested by Reg. 28/2004		
<b>9.5. Dissemination format - other</b>		
Not requested by Reg. 28/2004		
<b>9.6. Documentation on methodology</b>		
Not requested by Reg. 28/2004		
<b>9.7. Quality management - documentation</b>		
Not requested by Reg. 28/2004		
<b>9.7.1. Metadata completeness - rate</b>		
Not requested by Reg. 28/2004		
<b>9.7.2. Metadata - consultations</b>		

Not requested by Reg. 28/2004

**10. Cost and Burden**[Top](#)**Interview duration**

The average interview duration in successfully interviewed households (the whole interview time: household + all personal questionnaires combined) was 25.7 minutes. The average interview duration we can divide between paper questionnaire interview (PAPI) 39.2 minutes and computer questionnaire interview (CAPI) 29.3 minutes.

The following tables presents the mean interview duration in minutes calculated as the sum of the duration of all household interviews (HB100) plus the sum of the duration of all personal interviews (PB120), divided by the number of household members aged 16 and over whose household questionnaire is completed and accepted for the database (PB030).

Average interview durations in minutes (2008 to 2017), cross sectional sample 2017

2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
36.6	30.3	31.3	29.8	28.7	30.5	27.9	29.5	26.7	25.7

Average interview durations in minutes by type of interview, cross sectional sample 2017

Method	Total duration	Duration of personal questionnaire
Face to face interview - PAPI	39.1	17.9
Face to face interview – CAPI	29.3	8.6
Self-administered by respondent	-	-
Proxy interview - PAPI	37.2	15.7
Proxy interview - CAPI	28.9	8.6
<b>Total</b>	<b>25.7</b>	<b>13.1</b>

**11. Confidentiality**[Top](#)

Not requested by Reg. 28/2004

**11.1. Confidentiality - policy**

Not requested by Reg. 28/2004

**11.2. Confidentiality - data treatment**

Not requested by Reg. 28/2004

**12. Comment**[Top](#)

--

**Related metadata**[Top](#)**Annexes**[Top](#)