eurostat

SILC_ESQRS_A_HU_2014_0000

National Reference Metadata in ESS Standard for Quality Reports Structure (ESQRSSI)
Compiling agency: Hungarian Central Statsitical Office
Time Dimension: 2014-A0
Data Provider: HU1
Data Flow: SILC_ESQRS_A:1.1

Eurostat metadata

Reference metadata

For any question on data and metadata, please contact: EUROPEAN STATISTICAL DATA SUPPORT

1. Contact

Top

1.1. Contact organisation	Hungarian Central Statsitical Office
1.2. Contact organisation unit	Living Standard-, Labour- and Education Statistics Department Living Standard Statistics Survey Section
1.5. Contact mail address	H-1024 Budapest Keleti Károly u. 5-7.

2. Introduction

Top

The production of quality reports is part of the implementation of the EU-SILC instrument. In order to assess the quality of data at national level and to make a comparison among countries, the National Statistics Institutes are asked to report detailed information mainly on: the entire statistical process, sampling and non-sampling errors, and potential deviations from standard definition and concepts.

This document follows the ESS standard for quality reports structure (ESQRS), which is the main report structure for reference metadata related to data quality in the European Statistical System. It is a metadata template, based on 13 main concepts, which can be used across several statistical domains with the purpose of a better harmonisation of the quality reporting requirements in the ESS.

For that reason the template of this document differs from that one stated in the Commission Reg. 28/2004.

Finally it is the combination of the previous intermediate and final quality reports therefore it is worth mentioning that it refers to both the cross sectional and the longitudinal data.

3. Quality management - assessment

Top

This item is not requested by Reg 28/2004

4. Relevance

Top

This item is not requested by Reg. 28/2004

4.1. Relevance - User Needs

This item is not requested by Reg. 28/2004.

4.2. Relevance - User Satisfaction

This ietm is not requested by Reg. 28/2004

4.3. Completeness

This item is not requested by Reg. 28/2004

4.3.1. Data completeness - rate

This item is not requested by Reg. 28/2004

5. Accuracy and reliability

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 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-sampl process.

5.1. 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 si size combines sample size and sampling design effect which depends on sampling design, population structure and non-response rate.

5.2. 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 "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-line 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 tak 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 alsc

The main hypothesis on which the calculations are based is that the "at risk of poverty" threshold is fixed. According to the characteristics and availability of data for different countrie particular, countries have been split into three groups:

1)BE, BG, CZ, IE, EL, ES, FR, IT, LV, HU, NL, PL, PT, RO, SI, UK and HR whose sampling design could be assimilated to a two stage stratified type we used DB050 (primary strat specification;

2) DE, EE, CY, LT, LU, AT, SK, FI, CH whose sampling design could be assimilated to a one stage stratified type we used DB050 for strata specification and DB030 (household ID)

3) DK, MT, SE, IS, NO, whose sampling design could be assimilated to a simple random sampling, we used DB030 for cluster specification and no strata;

In case Eurostat methodology is not accepted by your country, please describe the methodology used at national level for computing the estimates. The Eurostat methodology is accep

5.2.1. Sampling error - indicators

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 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-sampl process.

2014:

Indicator	Breakdown	Indicator value	SE %	CI95% lower bound	CI95% upper bound
AROPE	Total	31.8	0.68	30.5	33.1
	Male	31.4	0.77	29.9	32.9
	Female	32.3	0.71	30.9	33.7
	Age 0-17	41.8	1.52	38.8	44.8
	Age 18-64	32.4	0.64	31.1	33.7
	Age 65+	19.0	1.16	16.7	21.3
ARPT60	Total	15.0	0.57	13.9	16.1
	Male	15.5	0.67	14.2	16.8
	Female	14.5	0.55	13.4	15.6
	Age 0-17	25.0	1.42	22.2	27.8
	Age 18-64	14.9	0.55	13.8	16.0
	Age 65+	4.5	0.43	3.7	5.3
SMD	Total	24.0	0.73	22.6	25.4
	Male	23.7	0.79	22.2	25.2
	Female	24.4	0.77	22.9	25.9
	Age 0-17	31.9	1.5	29.0	34.8
	Age 18-64	23.8	0.68	22.5	25.1
	Age 65+	16.5	1.19	14.2	18.8
LWI	Total	12.8	0.63	11.6	14.0
	Male	12.3	0.75	10.8	13.8
	Female	13.3	0.63	12.1	14.5
	Age 0-17	15.2	1.34	12.6	17.8
	Age 18-59	12.1	0.5	11.1	13.1

5.3. 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 inte
- 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:

1. – Unit non-response: refers to absence of information of the whole units (households and/or persons) selected into the sample
1. – Item non-response: refers to the situation where a sample unit has been successfully enumerated, but not all required information has been obtained

5.3.1. 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

5.3.1.1. Over-coverage - rate

	Main problems	Size of error
Cross sectional data	·Over-coverage	not present
	·Under-coverage	not present
	·Misclassification	not present

5.3.2. Measurement error

Cross sectional data

Source of measurement errors	Building process of questionnaire	Interview training
	Based on the experiences of the previous waves (HU-SILC2005 -2013) the following steps were done: • The questionnaire was formed according to Eurostat recommendations. • To avoid non-response of respondents because of personal data-protections reasons we have kept the separated data sheet for the names and birth date of the respondents. It was called address sheet (Címkártya).	We used computer assisted personal (CAPI) interviews during the data collection A detailed manual was complied for interviewers to deepen their knowledge about the structure of the questionnaire and the management of the interview. Training was organized for the colleagues working in the Regional offices by the experts of the Central office. Detailed interviewers manual and presentations were prepared on the questions of all the questionnaires (household, personal, and data-sheet questionnaire), possible problems and respondent approach as well. The training for interviewers was organized by the 7 Regional offices using the supporting document and presentations supplied for the central training. Uniformed training schedule and script were used for the regional trainings. Programs used for data capturing was tested in central office and regional offices as well.

5.3.3. 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 28/2

- **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.

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.

5.3.3.1. 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 (NRp)*	
A*	B*	A*	B*	A*	B*	A*	B*	A*	B*	A*	
0.9978	0.9997	0.8481	0.7873	0.9604	0.9493	15.38	21.29	3.9637	5.0655	24.41	25.8

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

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

5.3.3.2. 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-respons level.

5.3.3.2.1. Item non-response rate by indicator

	Total hh gross income (HY010)	Total disposable hh income (HY020)	Total disposable hh income before social transfers other than old-a (HY022)
% of household having received an amount	100.0	100	99.1
% of household with missing values (before imputation)	0	0	0
% of household with partial information (before imputation)	0	0	0

	Imputed rent (HY030)	Income from rental of property or land (HY040)	Family/ Children related allowances (HY050)	Social exclusion payments not elsewhere classified (HY060)	Housing allowances (HY070)	Reg tra
% of household having received an amount	96.5	3.1	30.2	6.2	11.9	19.8
% of household with missing values (before imputation)	96.5	0	0	0	0	0
% of household with partial information (before imputation)	0	0	0	0	0	0

	Cash or near-cash employee income (PY010)	Other non-cash employee income (PY020)	Income from private use of company car (PY021)	Employers social insurance contributions (PY030)	Cash profits or losses from self-employment (PY050)	Value of goods produced for own consumption (PY070)	Unemployment benefits (PY090)	Old-age benefits (PY100)	Survivors benefits (PY110)	Sickness benefits (PY120)	Disability benefits (PY130)	Education-related allowances (PY140
% of household having received an amount	49.3	4.2	0.2	49.3	7.2	0	5.3	28.5	1.2	3.3	5.7	1.3
% of household with missing values (before imputation)	11.2	0	0	100	0.2	0	0	0.3	0	0	0	0
% of household with partial information (before imputation)	28.1	2.9	54.8	0	56.3	0	0	31.4	0	0	0	0

FOR THE LONGITUDINAL ELEMENT:		Wave 1 - year 2011				Wave 2 - year 2012			Wave 3 - year 2(		
INCOME GROSS VARIABLES		Mean	No. Of observations Before Imputation	No. Of observations After Imputation	Standard error	Mean	No. Of observations Before Imputation	No. Of observations After Imputation	Standard error	Mean	No. Of observa Before Imputat
Total hh gross income	(HY010)	2841744	4683	4753	343003009995	4118	4210	397682996446		3608	
Total disposable hh income	(HY020)	2333551	4683	4753	250122469264	4118	4210	294112400978		3608	
Total disposable hh income before social transfers other than old-age and survivors benefits	(HY022)	2041701	4625	4694	254052187992	4055	4147	294222149246		3570	
	(HY023)	1649764	4162	4225	289151767895	3613	3699	344741656559		3240	

Total disposable hh income before all social transfers									
Imputed rent	(HY030)	510371	0	4546	5067502654	0	4059	5067491551	0
Income from rental of property or land	(HY040)	445234	54	54	8575797976	46	46	306322565350	26
Family/ Children related allowances	(HY050)	459776	1641	1641	10450445830	1383	1383	11898411143	1127
Social exclusion payments not elsewhere classified	(HY060)	156921	350	350	11002179138	267	267	18159194235	231
Housing allowances	(HY070)	70519	368	368	248960466	349	349	182253028	336
Regular inter-hh cash transfers received	(HY080)	225470	900	900	9151218485	860	860	9597213271	719
Interest, dividends, profit from capital investments in incorporated businesses	(HY090)	653548	83	83	170200602248	101	101	188706433962	71
Interest repayments on mortgage	(HY100)	261828	842	842	3644374278	664	664	10859371279	547
Income received by people aged under 16	(HY110)	67594	16	16	083907	15	15	0257167	3
Regular taxes on wealth	(HY120)	17109	2876	2876	33816955	2572	2572	29916865	2297
Regular inter household cash transfer paid	(HY130)	153940	684	684	7552149898	676	676	7802153528	625
Tax on income and social contributions	(HY140)	685714	3300	3300	11856834080	2400	2400	17060884234	2310
Cash or near-cash employee income	(PY010)	1660308	4942	4967	235201733446	4413	4461	268591760845	3784
Other non-cash employee income	(PY020)	90566	470	470	4692101585	312	312	4983105249	345
Income from private use of company car	(PY021)	251196	24	24	445247265200	20	20	11756291667	12
Employers social insurance contributions	(PY030)	474316	0	4393	7118502699	0	4461	7789475460	0
Cash profits or losses from self-employment	(PY050)	974056	811	865	53483955738	758	827	45726993545	672
Unemployment benefits	(PY090)	259340	767	767	5317255868	657	657	8123194386	478
Old-age benefits	(PY100)	1094970	2312	2320	111351124081	2185	2197	114151187165	1916
Survivors benefits	(PY110)	475066	154	154	22147481746	134	134	15780532006	105
Sickness benefits	(PY120)		0	0	86240	407	407	605291450	219
Disability benefits	(PY130)	637121	694	695	12762662575	620	620	13545663824	510
Education-related allowances	(PY140)	179293	176	176	10273160273	134	134	7291190364	123

5.3.4. Processing error

Data entry and coding

Blaise was used as data entry program. The data entry program was tested by colleagues of Regional offices and Central office experts. After the testing the data entry program was corrected.
The data entry program was loaded to each computer of each interviewer before the satring of the field work. The program contained checks to ensure the basic data consistency.

Editing controls

After building up the raw database the data were controlled i
- Identification numbers controlling
- Outlier controlling
- Data consistency checking (for instance, basic demogra

status; economic status under the income reference period
- Controlling of the amount of social transfers

5.3.4.1. Imputation - rate

Nor requested by Reg. 28/2004

5.3.4.2. Common units - proportion

Not rquested by Reg. 28/2004

5.3.5. Model assumption error

Not requested by Reg. 28/2004

5.3.6. Data revision

This item is not requested by Reg. 28/2004

5.3.6.1. Data revision - policy

This item is not requested by Reg. 28/2004

5.3.6.2. Data revision - practice

This item is not requested by Reg. 28/2004

5.3.6.3. Data revision - average size

This item is not requested by Reg. 28/2004

5.3.7. Seasonal adjustment

This item is not requested by Reg. 28/2004

6. Timeliness and punctuality

Top

6.1. Timeliness

According to Reg. 28/2004:
Timeliness of information' reflects the length of time between its availability and the event or phenomenon it describes
Punctuality refers to the time lag existing between the actual delivery date of data and the target date when it should have been delivered, for instance, with reference to dates announced in some official release calendar, laid down by regulations or previously agreed among partners

6.1.1. Time lag - first result

The data collection was carried out in March April and May of 2014 with the income reference year of 2013.
First publication of HU-SILC were available after 7 months after the data collection in 26.11.2014. It takes 11 months compared to the income reference period (2013).
We did not prepare any preliminary relase.

6.1.2. Time lag - final result

The same as described in 6.1.1.

6.2. Punctuality

Punctuality refers to the time lag existing between the actual delivery date of data and the target date when it should have been delivered, for instance, with reference to dates announced in some official release calendar, laid down by regulations or previously agreed among partners

6.2.1. Punctuality - delivery and publication

The data base release was done according to schedule.
The first publication based on HU-SILC 2014 module data (material deprivation) was done according to schedule.
The first comprehensive study on social exclusion indicators based on HU-SILC 2014 was published on 30th November 2014.
Indicators are available on Hungarian Central Statistica office website since that time.
http://www.ksh.hu/docs/hun/xstadat/xstadat_eves/i_zaa007.html

7. Accessibility and clarity

Top

7.1. Dissemination format - News release

This item is not requested by Reg. 28/2004

7.2. Dissemination format - Publications

The publication based on HU-SILC 2014 was published on 26.11.2014 and available on this link (only in Hungarian):
http://www.ksh.hu/apps/shop.kiadvany?p_kiadvany_id=44869&p_temakor_kod=KSH&p_session_id=289643876659611&p_lang=HU

7.3. Dissemination format - online database

Most important national poverty and social exclusion indicators are avaiable in our website:
http://www.ksh.hu/docs/eng/xstadat/xstadat_annual/i_zaa007.html
For internation comparison poverty related figures available in our website based on Eurostat data.
http://www.ksh.hu/docs/hun/eurostat_tablak/tabl/tsdsc320.html

7.3.1. Data tables - consultations

This item is not requested by Reg. 28/2004

7.4. Dissemination format - microdata access

This item is not requested by Reg. 28/2004

7.5. Documentation on methodology

Meta informations are available on the followig link:
http://www.ksh.hu/apps/meta.objektum?p_lang=HU&p_menu_id=110&p_ot_id=100&p_obj_id=ZAA&p_session_id=87774435

7.5.1. Metadata completeness - rate

This item is not requested by Reg. 28/2004

7.5.2. Metadata - consultations

This item is not requested by Reg. 28/2004

7.6. Quality management - documentation

This item is not requested by Reg. 28/2004

7.7. Dissemination format - other

This item is not requested by Reg. 28/2004

8. Comparability

Top

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."

8.1. Comparability - geographical

This item is not requested by Reg. 28/2004

8.1.1. Asymmetry for mirror flow statistics - coefficient

This item is not requested by Reg. 28/2004

8.1.2. Reference population

Reference population	Private household definition	Household membership
No difference to common definition	No difference to common definition	No difference to common definition

8.1.3. Reference Period

Period for taxes on income and social insurance contributions	Income reference periods used	Reference period for taxes on wealth	Lag between the income ref period and current variables
Fixed twelwe month period was used, which was the previous calendar year:2013.	2013	2013	The lag between income reference period and the current variables is 3 months since the reference time of interview was 1st of March 2014.

8.1.4. Statistical concepts and definitions

Total hh gross income (HY010)	Total disposable hh income (HY020)	Total disposable hh income before social transfers other than old-age and survivors' benefits (HY022)	Total disposable hh income before all social transfers (HY023)
F	F	F	F

Imputed rent (HY030)	Income from rental of property or land (HY040)	Family/ Children related allowances (HY050)	Social exclusion payments not elsewhere classified (HY060)	Housing allowances (HY070)	Regular inter-hh cash transfers received (HY080)	Interest, dividends, profit from capital investments in incorporated businesses (HY090)	Interest paid on mortgage (HY100)	Income received by people aged under 16 (HY110)	Regular taxes on wealth (HY120)	Regular inter-hh transfers paid (HY130)
L Estimated by a regression model	F	F	F	F	F	F	F	F	F	F

Cash or near-cash employee income (PY010)	Other non-cash employee income (PY020)	Income from private use of company car (PY021)	Employers social insurance contributions (PY030)	Cash profits or losses from self-employment (PY050)	Value of goods produced for own consumption (PY070)	Unemployment benefits (PY090)	Old-age benefits (PY100)	Survivors benefits (PY110)	Sickness benefits (PY120)	Disability benefits (PY130)	Education-related allowances (PY140)	Gross monthly earnings for employees (PY200)
F	F	F	F	F	F	F	F	F	F	F	F	F

The source or procedure used for the collection of income variables	The form in which income variables at component level have been obtained	The method used for obtaining target variables in the required form
All the income variables were collected from the respondents. The income target variables were grouped into more detailed sub-components according to Hungarian tax and benefit system	Gross income data were collected for the income items but in case of certain benefits according to tax law which were not considered to be belonging to the taxable income net value were asked, like old-age pension or family allowance	The income items were divided into sub-components according to the Hungarian tax regulations and benefit practice in the questionnaire. The personal and household incomes were separated. Gross income and net items were asked for work related incomes and other incomes belonging to the personal tax system and net income items were asked for benefits and other allowances. The following steps were taken to obtain income target variables in the required form. 1. The subcomponents were summed up to obtain the income items on personal income level. 2. While Hungary has a personal income tax system, the household type incomes had to be connected to household members. It was done on the basis of the income type, eg. Agricultural income was connected to the household member(s) reporting agricultural activity. Obviously just adult members were involved. 3. The value of taxable income was calculated for each household member. 4. The total household gross income was calculated for the household including all income types on basis of the process listed at i. and ii. 5. On the basis of value of taxable income for each household member, the value of personal income tax and social insurance fee was calculated. The deductions were summed up for total of the household. The total disposable income on household level was calculated as difference between the total household gross income and the total tax deductions.

8.2. Comparability - over time

HU-SILC survey was introduced to the Hungarian data collection sytem in 2005. The survey follows EU-regulations from the beginning and produces annually comparable data till present.

8.2.1. Length of comparable time series

Comparable time series are available from 2005 till 2014.

8.3. Comparability - domain

Cross sectional and longitudinal data sets are also comparable.

9. Coherence

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.

9.1. Coherence - cross domain

Number of persons aged 16-74 by self-classification and by gender in HU-LFS and in HU-SILC, 2014

Age-group	HU-LFS			HU-SILC		
	Men	Women	Total	Men	Women	Total
Persons (thousand)						
Working	2201.1	1865.3	4066.4	2053.5	1790.9	3844.5
Unemployed	285.3	256.9	542.2	389.5	277.5	666.9
Pupil, student, further training, unpaid work experience	316.9	323.3	640.2	360.4	346.8	707.2
In retirement or in early retirement or permanently disabled	751.8	1071.0	1822.8	741.3	1086.8	1828.1
Fulfilling domestic tasks and care responsibilities	11.5	314.0	325.6	3.4	103.8	107.3
Other inactive person	36.1	44.0	80.1	44.0	261.2	305.2
Total	3602.6	3874.6	7477.3	3592.2	3867.1	7459.2
Distribution (%)						
Working	61.1	48.1	54.4	57.2	46.3	51.5
Unemployed	7.9	6.6	7.3	10.8	7.2	8.9
Pupil, student, further training, unpaid work experience	8.8	8.3	8.6	10.0	9.0	9.5
In retirement or in early retirement or permanently disabled	20.9	27.6	24.4	20.6	28.1	24.5
Fulfilling domestic tasks and care responsibilities	0.3	8.1	4.4	0.1	2.7	1.4
Other inactive person	1.0	1.1	1.1	1.2	6.8	4.1
Total	100.0	100.0	100.0	100.0	100.0	100.0

9.1.1. Coherence - sub annual and annual statistics

This item is not requested by Reg. 28/2004

9.1.2. Coherence - National Accounts

An analysis was carried out to compare HU-SILC income data on item level to National accounts correspondingfigures. Altough the trends of the total aggreagates of HY020G and Real Gross Household Disposable Income moving into the same direction but National Accounts uses of differrent reference population (institutional and private households living in Hungary) than HU-SILC (private households living in Hungary) we can not compare data on item level, so this comparison would not be provided for this quality report.

9.2. Coherence - internal

This item is not requested by Reg. 28/2004

10. Cost and Burden

Top

This item is not requested by Reg. 28/2004

11. Confidentiality

Top

11.1. Confidentiality - policy

This item is not requested by Reg. 28/2004

11.2. Confidentiality - data treatment

This item is not requested by Reg. 28/2004

12. 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.

12.1. Source data

In 2014 the sampling frame was the list of occupied dwelling units in census 2011 dataset.

12.1.1. Sampling design and procedure

Type of sampling design

In 2014 a new rotational group (number 13) with 2894 selected households was introduced. As opposed to previous years the new sub-sample is a one-phase sample with two-stage selection.

At stage 1 we have a stratified sample of localities with pps selection.

At stage 2 (in the sampled localities) we have a stratified simple random sample of households.

Stratification and sub stratification criteria

At stage 1 the population of localities is stratified. Each of the larger localities is a stratum of its own. These are the self-representing localities, the number of which is 91. Smaller localities are stratified by NUTS3 regions and the size of locality.

At stage 2 the households are stratified by the charachteristic of the head of household within each locality.

Sample selection schemes

At stage 1 localities were selected with pps without replacement.

At stage 2 households within each strata were selected with srs without replacement.

Sample distribution over time

Size of rotational groups (selected sample)

Size of rotational groups (selected sample)

	2011	2012	2013	2014
Rotational group1	-	-	-	-
Rotational group2	-	-	-	-
Rotational group3	-	-	-	-
Rotational group4	-	-	-	-
Rotational group6	-	-	-	-
Rotational group7	2158	-	-	-
Rotational group8	3016	2672	-	-
Rotational group 9	2779	2478	2274	-
Rotational group 10	5495	4788	4374	3678
Rotational group 11	-	3055	2571	2169
Rotational group 12	-	-	3191	2522
Rotational group 13	-	-	-	2894
Total sample	13448	12993	12410	11263

12.1.2. Sampling unit

In the population of smaller localities PSUs are the localities and SSUs are households.

In the population of larger (self-representing) localities PSUs are households.

12.1.3. 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
- the achieved sample size which is the number of observed sampling units (household or individual) with an accepted interview
- 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

Given that the effective sample size has been already treated in the section dealing with sampling errors, in this section the attention focuses mainly on the achieved sample size.

Achieved sample size for the Cross sectional data

	No of households	No of persons 16+
1st rotational group	2225	4582
2nd rotational group	3139	6584
3rd rotational group	1801	3623
4th rotational group	2046	4307

Achieved sample size for the Longitudinal data

	No of households	No of persons 16+
1st wave 2011	4754	
2nd wave 2012	4211	
3rd wave 2013	3650	
4th wave 2014	3139	

12.2. Frequency of data collection

Data collection of HU-SILC has a fixed period and duration in the fieldwork timetable in the Hungarian social data collection system.It is annual dtata collection. The reference data is 1st Of March. The data collection started at the first working day of March and took 10 weeks.

Fieldwork timing and sample development over time

Weeks of interview	Achieved sample size	Distribution of achieved sample
Weeks		
1 March - 2 March	310	3.4%
3 March - 9 March	1822	19.8%
10 March - 16 March	2297	24.9%
17 March - 23 March	2483	27.0%
24 March - 30 March	2050	22.3%
31 March - 6 April	218	2.4%
7 April - 13 April	15	0.2%
14 April - 20 April	8	0.1%
21 April - 27 April	1	0.0%
28 April - 30 April	7	0.1%
	9211	100%

12.3. Data collection

RB260-Contact address	Total	%
PAPI	0	0.0
CAPI	16955	88.8
Proxy	2141	11.2
Total	19096	100.0

Distribution of household members aged 16 and over by RB250 (total and rotational groups breakdown)

HOUSEHOLD MEMBERS 16+ (RB245 = 1 to 3)

	Total	RB250 = 11	RB250 = 12	RB250=13	RB250=14	RB250 = 23	RB250 = 31	RB250 = 33
Total	19096	19096	0	0	0	0	0	0
%	100	100.0%						
R1	4582	4582						
%	100	100.0%						
R2	6584	6584						
%		100.0%						
R3	3623	3623						
%		100.0%						
R4	4307	4307						
%		100.0%						

HOUSEHOLD MEMBERS 16+ (RB245 = 2)

Total 0

%

HOUSEHOLD MEMBERS 16+ (RB245 = 3)

Total 0

%

Distribution of household members aged 16 and over by RB260 (total and rotational groups breakdown)

HOUSEHOLD MEMBERS 16+ (RB245 = 1 to 3) and RB250 = 11 or 13

	Total	RB260 = 1	RB260 = 2	RB260 = 3	RB260 = 4	RB260 = 7	Missing
Total	19096		16955			2141	
%	100		88.8%			11.2%	
R1	8645		4063			4582	
%	100		47.0%			53.0%	
R2	12415		5831			6584	
%	100		47.0%			53.0%	
R3	6857		3234			3623	
%	100		47.2%			52.8%	
R4	8134		3827			4307	
%	100		47.0%			53.0%	

HOUSEHOLD MEMBERS 16+ (RB245 = 2) and RB250 = 11 or 13
Total 0
%
HOUSEHOLD MEMBERS 16+ (RB245 = 3) and RB250 = 11 or 13
Total 0
%

12.4. Data validation

This item is not requested by Reg. 28/2004

12.5. Data compilation

This item is not requested by Reg. 28/2004

12.5.1. Weighting procedure

Design factor	Non-response adjustments	Adjustment to external data	Final cross sectional weights
By definition design weight is the reciprocal of the inclusion probability. However, thanks to the fact that three rotational groups still have a rather complex sample design and selection scheme, only an approximation of design weight was calculated. Weighting classes were defined by NUTS2 regions, category of size of localities and household strata. Within each weighting classes the 'design' weight is equal to then ratio of the overall number of households to the number of selected households.	In the new rotational group non-response weights were introduced to reduce bias caused by unit non-response on household level. Non-response adjustment was a simple expansion applied by the same classes as design factors were calculated by.	The aim of this adjustment was to improve the accuracy of data using socio-economical information available from the constantly updated Census 2011 and other surveys. Iterative raking scale methods were applied. For the integrative calibration the following controls were used: • Population totals for sex * age * region groups defined by ages 0-14, 15-29, 30-59, 60 or more; • Population totals for sex * age * type of locality groups defined by ages 0-14, 15-29, 30-59, 60 or more; • Population totals for activity status * type of locality groups • Population totals of the actives for education level * type of locality groups • Total number of households for household* type of locality groups. Calibration was carried out with a self made SAS program. Calibration was carried out in each rotational group. For the new rotational group the input weight for calibration is the one described previously. For the rest of rotational groups the input weight for calibration is previous year's final cross- sectional weight.	After calibrating the new and former rotational groups separately, those adjusted weights were reduced proportional to the group size. Finally, one more calibration was applied for the overall sample with a small number of iterations. Final cross-sectional weights for the whole sample are in the interval [90,1250].

12.5.2. Estimation and imputation

Imputation procedure used	Imputed rent	Company car
Incase of partial data on income items PY010G, PY050G, PY100G regression method is used for income imputation	Hungaryhas got a special housing market situation in the aspect of imputed rental calculation. The share of market rental sector is 3 %. Owner occupiers constitute 97 % of the total housing market. Personal attitudes and social circumstances make stronger the role of private property in the housing market. Geographical and physical attributes and mainly the location of the dwelling within the country determines mostly the value of a dwelling, and possibility to let it on the rental market. Comparison of standard of living on the basis of EU-SILC survey between different social groups is not affected by the minor groups of market renters. The calculation of imputed rent is reasoned by international comparison of data within EU. Regression method was used to calculate the value of imputed rent on household level. We asked the value of subjective rent on household level. The following question was asked in the questionnaire: “How much you should pay as a rent for a dwelling similar to your current one either in size, number of rooms and conditions in your close neighborhood?” The value of the subjective rent was used as a dependent variable in the regression calculation. Wide set of explaining variable and linear regression models were tested as well. The one with the highest R2 was chosen.	A question was used to determine the value of private use of company car in on the questionnaire. It was answered by the respondents reporting use of company cars. The respondent had to estimate this value and this estimation was used in the database

12.6. Adjustment

This item is not requested by Reg. 28/2004

13. Comment

Top

Related metadata

Top

Annexes

Top