SILC_ESQRS_A_EE_2013_0000

National Reference Metadata in ESS Standard for Quality Reports Structure (ESQRSSI)
Compiling agency: Statistics Estonia
Time Dimension: 2013-A0
Data Provider: EE1
Data Flow: SILC_ESQRS_A

Eurostat metadata

Reference metadata

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

1. Contact

Top

1.1. Contact organisation	Statistics Estonia
1.2. Contact organisation unit	Population and social statistics department
1.5. Contact mail address	Statistikaamet. Tatari 51, 10134 Tallinn

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

Not requested by Reg. 28/2004

4. Relevance

Top

Not requested by Reg. 28/2004

4.1. Relevance - User Needs

Not requested by Reg. 28/2004

4.2. Relevance - User Satisfaction

Not requested by Reg. 28/2004

4.3. Completeness

Not requested by Reg. 28/2004

4.3.1. Data completeness - rate

Not requested by Reg. 28/2004

5. 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.

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 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.

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 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 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 countries we have used different variables to specify strata and cluster information. In particular, countries have been split into four 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 strata) for strata specification and DB060 (Primary Sampling Unit) for cluster 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) for cluster specification;

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.

5.2.1. Sampling error - indicators

	AROPE			At risk of poverty (60%)			Severe Material Deprivation			Very low work intensity		
	Ind. value	Stand. errors	Half CI (95%)	Ind. value	Stand. errors	Half CI (95%)	Ind. value	Stand. errors	Half CI (95%)	Ind. value	Stand. errors	Half CI (95%)
Total	23.5	0.69	1.35	18.6	0.62	1.21	7.6	0.46	0.91	8.4	0.49	0.96
Male	22.5	0.8	1.57	17.2	0.71	1.4	8.1	0.55	1.08	9.5	0.59	1.16
Female	24.4	0.76	1.5	19.9	0.7	1.38	7.1	0.48	0.94	7.3	0.54	1.06
Age0-17	22.3	1.29	2.54	18.1	1.2	2.36	7.0	0.83	1.63	6.6	0.78	1.53
Age18-64	22.7	0.77	1.5	17.3	0.67	1.31	8.0	0.54	1.06	8.9	0.48	0.94
Age 65+	28.0	1.16	2.27	24.4	1.1	2.17	6.3	0.64	1.26	-	-	-

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 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:

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

- population

Sampling frame for selection of the new part of the sample was the 2011 Population and Housing Census. Census involves the data of all permanent residents, households and dwellings in Estonia. Following persons are not considered permanent residents of Estonia: persons, who have already left Estonia (according to the statement of their relatives), persons permanently residing abroad (despite their wish to enumerate themselves) and persons, who reside in Estonia temporarily (3–12 months), but whose permanent residence is abroad.

In sampling, we use person-approach, i.e. we trace and interview individually people who were selected from the census, not the address nor dwelling. So, frame error is considered to be an over-coverage error if address-person did not actually belong to target population, i.e.

- was dead;
- had moved to another county;
- stayed in an institution permanently (had been there over half a year);
- was surveyed through one of his/her household members;

All households classified under DB120=23 are considered to constitute overcoverage error.

In the new part of the sample there were certain amount of address-persons those address in the census was definitely wrong and no information on new address could be obtained from neighbours nor other sources. According to national classification, this includes the following reasons for non-contact:

- Address-person does not live at given address, no information on new address available;
- Address-person has moved to another address, no information on new address available;
- Given address does not exist.

It does not seem reasonable to assume that these persons do not belong to target population nor constitute frame over-coverage. Above mentioned reasons for non-contact are currently classified under non-response reason DB120=21.

After the 2011 Census conducted in Estonia there have been reports of people, who were not enumerated. Some people, who were not enumerated at first, informed about themselves. However, the enumerators did not receive the information about all people, who were not enumerated – this group includes people, who were left out because they avoided the Census intentionally, or were just passive or disloyal. The Census data are presumably subject to at least 1% under-coverage.

5.3.1.1. Over-coverage - rate

	Main problems	Size of error
Cross sectional data	·Over-coverage ·Under-coverage ·Misclassification	Overcoverage rate is 2%

5.3.2. Measurement error

Cross sectional data

Source of measurement errors	Building process of questionnaire	Interview training	Quality control
The measurement errors can stem from the questionnaire (its wording, design etc), the interviewers and the data collection method. While it is impossible to avoid this type of errors completely, steps were taken to reduce them as much as possible.	The ESS questionnaire has been drafted following international experience in collecting income data. Where possible questions and wordings from Statistics Estonia's previous surveys, the reliability and validity of which had been checked in practise, were used. In 2007 the questionnaire was supplemented using the experience from the past three waves. The main corrections in the household questionnaire were adding in questions about production of foodstuffs for own consumption and questions allowing the calculation of savings from imputed rent. In the personal questionnaire the main developments in 2007 were adding questions about education obtained since the previous interview for the longitudinal panel, allowing the choice to report wage income as yearly or monthly and net or gross, adding questions about non-monetary income from wage labour and a separate block of income questions for entrepreneurs. The social benefit questions were also updated and additional checkpoints created to ask respondents the questions that concern their situation specifically. The questions on child-care, family benefits and unemployment benefits were also improved. In 2008 questions about managerial duties for current and last job were added and socio-economic statuses were prefilled for respondents who had answered the personal questionnaire the previous year for the months they had already provided answers for. An additional question was added regarding pensions paid by the local government and the conscript allowance paid to young men serving time in the armed forces. In 2009, the questions used to determine a respondent's level of education were improved. Previously a person had to choose their level of education from a long list of official names, resulting in considerable errors. In 2009 these questions were redesigned for more accuracy and less respondent-induced errors. In 2010, the questions used to determine using child care services reformulated to better meet Eurostat's guidelines. In 2011 during the first month questions about current costs were aked in two currencies (euros, kroons). Respondent was asked to choose which currency he would like to answer. Other notable modifications in 2011 concerned the following variables. The question about the number of rooms available to the household was reformulated according to the Eurostat's guidelines. 1) The question about the number of rooms available to the household was reformulated according to the Eurostat's guidelines. 2) The questions about intra-household sharing of resources were excluded 3) The questions about intergenerational transmission of disadvantages were added In 2012 questions about current costs were asked only in euros. Other notable modifications in 2012: 1) The questions about changing of dwelling were added. 2) The questions about formal childcare were a bit reformulated. 3) The questions about living conditions were reformulated and added. In 2013: *The module of living conditions have been left (M-questions). Two module has been added: 1. Well-being (all questions in personal questionnaire) 2. Material deprivation (1 question in the hosehold q and the rest in the personal q) *The most of the questions about family/children related allowances have been left (the data come from the register) *The most of the questions about old-age, survivor', unemployment and disability benefiits, additional contributions to the income tax and income tax returns have been left (the data come from the registers)	To reduce interviewer-induced measurement errors, the training programm was conducted in 4 smaller groups of about 15 people, with emphasis on practical work and discussion of mistakes from previous years. All returning interviewers attended a day long training session. During the training, the EU-SILC survey manager briefed the interviewers on all updates in the questionnaires, discussed previous years' errors, tracing rules and specifics of assigning household and person numbers in the longitudinal survey. Practical work sessions were conducted in groups of five and each interviewer had to conduct a model interview in a simulated situation using their laptop. At the end of the training session, each interviewer received personal feedback about their mistakes in the previous wave. Interviewers new to EU-SILC attended a 2 day training session, which included a thorough overview of questionnaires and practical exercises as well as all the topics covered with returning interviewers. Overall, 59 interviewers were responsible for conducting the interviews. The household–interviewer ratio was 85 households per interviewer.	

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 28/2004:

- **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.

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-response rate (NRp)*	
A*	B*	A*	B*	A*	B*	A*	B*	A*	B*	A*	B*
93.33	87.76	83.66	69.15	98.38	96.95	21.91	39.0	1.62	3.05	23.18	40.86

* 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-response rate is provided for the main income variables both at household and personal 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-age and survivors benefits (HY022)	Total disposable hh income before all social transfers (HY023)
% of household having received an amount				
% of household with missing values (before imputation)				
% of household with partial information (before imputation)				

	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)
% of household having received an amount							
% of household with missing values (before imputation)							
% of household with partial information (before imputation)							

	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												
% of household with missing values (before												

	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
imputation)												
% of household with partial information (before imputation)												

5.3.4. Processing error

Data entry and coding

Checking the data was done in three stages: data-entry checks during the interview, additional in-office checks during fieldwork and lastly data cleaning.

The data for 2013 operation was collected using CAPI. The data-entry program was written in Blaise and contained most of the consistency checks. In Statistics Estonia, interviewers are required to react in some form to all error messages that occur during interviewing.

The solution is either to correct an erroneous situation or if the situation is unusual but correct, add a remark to the data entry-program explaining this error. When assessing the quality of an interviewer's work, not adding a remark to an actually correct situation tat prompts an error message is also counted as an error. These logical checks allow to correct most of the errors already during an interview.

The primary data-entry consistency controls were of 6 major types:

1) Checks of consistency between different answers. These included, but were not limited to following instances:

1. whether a household or a person who according to other data should/should not have received a certain type of income reported it or not (e.g. whether households with children received family benefits, retired people (or people below retirement age) received pensions, employed persons received wages and so on);

2. whether benefits reported to have been received were logical in the age and gender dimensions. For instance student benefits for over 50 year-olds, income taxes for under 15 year-olds, maternity leave and childbirth allowances for men etc;

3. Whether an educational level attained was possible below a certain age, or educational levels were possible in said combinations for given years;

4. whether answers provided to different non-monetary deprivation items agreed with each other;

5. whether the relationships in the household matrix were consistent with each other as well as with the age and sex of the household members;

6. whether the difference between the

Editing controls

All mistakes found during the secondary in-office data editing were put up in a shared excel table, and had to be clarified with the interviewer or interviewee by the end of the fieldwork period. This was done in co-operation of the EU-SILC team and the interviewers' supervisors.

In 2007, there was a dramatic drop of the number of primary consistency errors. In 2006 there had been a total of 5654 errors, in 2007 the number had fallen to 1677. In 2008 the total number of errors was 1779, in 2009 – 1939, in 2010 - 1856, in 2011- 2102. In 2012 the number of errors decreased to 1883, in 2013 -- 986

Out of all the errors in 2013 38% (375 cases) required callback and clarification with the interviewer or interviewee, in 2012, 56% of cases had required callback.

Table. Interviewer errors and their processing, 2013

Type of error	Number of errors detected	Share of errors requiring a call-back
No remark explaining unusual situation	236	37,71%
Interviewer made an error, but did not correct it	213	38,03%
Interviewer's remark does not explain unusual situation	0	0,00%
Data not sufficient for coding	23	95,65%
Starting and finishing times recorded incorrectly	2	50,00%
Use of category Other, while a suitable category exists	436	32,57%
In-office checks	48	81,25%
Interviewer has misunderstood a question	21	4,76%
Data entry mistake	0	0,00%
Not interviewers error	7	0,00%
Total	986	38,03%

The third and final stage of data checks involved later in-office data cleaning. The controls implemented at this stage involved further checks of data consistency, consistency across time, and of extreme income values and as a final step the Eurostat data-checks. Extreme values for all income components as well as total income were checked and handled on a case-by-case basis.

Data entry and coding	Editing controls
starting and finishing time of the interview was too short or too long and so on. 7. whether reported taxes or medical benefits received were consistent with income. 8. membership in pension plans checked by year of birth to see if legally bound to have joined pension pillar. 9. checks for correct survey area, interviewer code and personal numbers matching household numbers. 2) Lower and upper bounds of income variables (incl. benefits). These checks were developed with regard to data collected in the previous wave as well as administrative information. 3) Tracing checks. These controls were implemented to ensure that all split-off households and new household members were assigned correct split numbers and person numbers respectively. 4) Checks not allowing for occupations to be written on too general a scale for coding. (e.g. salesperson, cleaner) 5) Checks for goods produced for own consumption, for instance quantities; 6) Checks with information from the previous year. These controls concerned demographic data, information on educational level and labour status as well as the calendar of activities. The in-office staff promptly checked the questionnaires that were electronically transmitted to the central office. This stage included the following controls: 1) All the errors suppressed by interviewers were activated and checked; 2) All remarks made by interviewers in the data entry-program were read through and where necessary, relevant corrections were made. 3) All split-off households as well as all households from which at least one member had left were scrutinized one by one. 4) All category 'other' answers were gone through to see if they could be classified under one of the given options. 5) Additionally paid income tax was checked in-household to check for double-reporting. 6) Errors in coding were gone through. 7) Study benefits were checked by possibility of obtaining them in the school the respondent attended and legally set amounts. 8) Consistency between time reported working under socio-economic status and months that salary was received. 9) Reported amounts of family benefits were checked compared with eligibility based on the structure of the family and benefit levels set out in legislation. Demographic information in the interviewers' reports was compared to the data recorded in the electronic questionnaires.	

5.3.4.1. Imputation - rate

Not requested by Reg. 28/2004

5.3.4.2. Common units - proportion

Not requested by Reg. 28/2004

5.3.5. Model assumption error

Not requested by Reg. 28/2004

5.3.6. Data revision

Not requested by Reg. 28/2004

5.3.6.1. Data revision - policy

Not requested by Reg. 28/2004

5.3.6.2. Data revision - practice

Not requested by Reg. 28/2004

5.3.6.3. Data revision - average size

Not requested by Reg. 28/2004

5.3.7. Seasonal adjustment

Not requested by Reg. 28/2004

6. Timeliness and punctuality

Top

Not requested by Reg. 28/2004

6.1. Timeliness

Not requested by Reg. 28/2004

6.1.1. Time lag - first result

Not requested by Reg. 28/2004

6.1.2. Time lag - final result

Not requested by Reg. 28/2004

6.2. Punctuality

Not requested by Reg. 28/2004

6.2.1. Punctuality - delivery and publication

Not requested by Reg. 28/2004

7. Accessibility and clarity

Top

Not requested by Reg. 28/2004

7.1. Dissemination format - News release

Not requested by Reg. 28/2004

7.2. Dissemination format - Publications

Not requested by Reg. 28/2004

7.3. Dissemination format - online database

Not requested by Reg. 28/2004

7.3.1. Data tables - consultations

Not requested by Reg. 28/2004

7.4. Dissemination format - microdata access

Not requested by Reg. 28/2004

7.5. Documentation on methodology

Not requested by Reg. 28/2004

7.5.1. Metadata completeness - rate

Not requested by Reg. 28/2004

7.5.2. Metadata - consultations

Not requested by Reg. 28/2004

7.6. Quality management - documentation

Not requested by Reg. 28/2004

7.7. Dissemination format - other

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
Persons living in collective households are included in the reference population. The share of persons who are living in collective households and who are not at the same time members of some other private household is likely to be very low. Additionally, there is no feasible way to estimate their share in the total population. Thus, the exclusion of these persons is unlikely to affect the comparability and reliability of the estimates.	There were no divergences from the common definition.	There were no divergences from the 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
There were no divergences from the common definition. Tax on income and social insurance contributions, as well as tax repayments and receipts refer to the income received during the income reference period (previous calendar year).	There were no divergences from the common definition. The income reference period was the previous calendar year (2012).	There were no divergences from the common definition. Taxes on wealth paid during the income reference period (previous calendar year) were recorded.	The lag between the income reference period and current variables ranges from 3 to 7 months, thus not exceeding 8 months stipulated in the regulation.

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)
F	F	F	F	F	F	F	F	F	F	F

Imputed rent:

User cost method was employed, as the share of market rents is very small. External data used for modelling refers to survey year and not income year. As sale prices have been rising quickly, imputed rent value may consequently be overestimated compared to other income variables.

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	NC

PY050: Profits or losses reported in annual accounts for tax purposes were recorded. In the case of unregistered self-employment, the respondents were asked to estimate the income received this way.

PY070: Most quantities were imputed from answers provided by respondents and unit costs were taken from the Household Budget Survey. Production costs were deducted from the total price thus obtained for own-consumption goods, and the profits were transferred to the personal level. The transfer was done by dividing the household aggregate characteristic by all members of the household aged 16 or over who answered the personal questionnaire. This value was added to their record as variable PY070N.

PY110: If more than one household member is eligible for survivors' benefits, the individual benefits are, by default, combined and paid as a single sum to one household member. Due to infeasibility of dividing the survivors' benefit received between household members, the whole benefit is recorded only for the household member to whose account it was transferred. This can marginally affect variable HY110 (income received by those under 16), but has no effect on total household income.

PY200: Variable was not recorded, as EU-SILC is not used to calculate the gender pay gap.

Income data from survey-year 2013 onwards are partially registry data (Tax and Customs Board, Unemployment Insurance Fund, Health Insurance Fund, Social Insurance Board) -- PY090, PY100, PY110, PY130, HY050, HY145

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
Income variables were collected via face-to-face interviews at component or where applicable at sub-component level or came from the registers.	Table summarizes mode in which different income variables were collected. It should be noted that where collection of only gross values is indicated designate in fact income components, which are not taxable (HY060, HY070, HY080, HY100, HY120, HY130, PY035, PY140), i.e. where gross equals net. Variables HY040, HY110, PY010 and PY050 were collected as either net or gross, depending on which was easier for the respondent to report. The remaining variables were collected only in net. Income data from survey-year 2013 onwards are basically registry data – gross; net is calculated (Tax and Customs Board, Unemployment Insurance Fund, Health Insurance Fund, Social Insurance Board) -- PY090, PY100, PY110, PY130, HY050, HY145	Where only net values were collected or only net or gross value was recorded, the corresponding net and gross values were calculated on the basis of recorded values. Conversion algorithms were created on the basis of the local tax system. Information as to which taxes were paid on income components were also collected and taken into account in conversions.

Table. Mode of collection for gross income variables in Estonian EU-SILC 2013 operation

Income component	**Collected gross**	**Collected net of tax and social contributions**	**Mixed mode net/gross**

	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
HY040				X
HY050				
HY060		X		
HY070		X		
HY080		X		
HY090			X	
HY100		X		
HY110				X
HY120		X		
HY130		X		
HY140			X	
HY145				
PY010				X
PY020			X	
PY035		X		
PY050				X
PY080			X	
PY090				
PY100				
PY110				
PY120			X	
PY130				
PY140		X		

8.2. Comparability - over time

Table compares the mean and number of recipients of most income components in EU-SILC 2013 to the estimates from the 2012 operation. Changes that emerge are, in general, in line with what could be expected. It should be noted that the fieldwork period ended in June and the 2013 data actually refers to the incomes of 2012.
EU-SILC in Estonia collects the respondent's annual income from the previous calendar year. Income data from survey-year 2013 onwards are partially registry data (Tax and Customs Board, Unemployment Insurance Fund, Health Insurance Fund, Social Insurance Board) -- PY090, PY100, PY110, PY130, HY050, HY145.

Within a year the average salary increased by 7% and the number of wage receivers increased 2%. Administrative data confirms the survey results.

At the same time, the number of people receiving unemployment benefits has increased – this is due to transition to registry data.

Benefits from self-employment increased and the number of entrepreneurs decreased. The number of entrepreneurs seems to fluctuate between survey years, which also hint to a relatively big pool of short-lived businesses.

Table. Mean and number of recipients of income components in EU-SILC 2012 and 2013

	Mean (EUR)		Number of recipients	
	2012	2013	2012	2013
Individual level components				
PY010N	7086	7591	644544	657207
PY020N	695	968	173891	181273
PY035N	391	413	77187	71261
PY050N	1752	1894	64432	61196
PY090N	955	1076	29878	45301
PY100N	3615	3831	288091	288823
PY110N	1620	1144	7328	7417
PY120N	261	304	95690	95826
PY130N	1991	2243	88279	101407
PY140N	883	851	36243	34047
Household level components				
HY040N	664	1359	13326	17494
HY050N	1656	1615	174316	169758
HY070N	814	812	12342	10877
HY080N	1722	1629	26323	23831
HY090N	131	342	199401	188449
HY110N	331	485	4359	4499
HY120N	51	62	334553	241689
HY130N	1480	1725	29576	29343
HY145N	-247	-267	232560	340039
HY010	13051	14409		
HY020	10888	12014		
HY022	10215	11155		
HY023	10471	10381		

Household level variables reflect changes in line with personal level variables.

HY040N increased -- more people are receiving income from renting your property or land. Family allowances have decreased due to the decrease of the number of recipients. The average amount of housing allowances has decreased and the amount of recipient households has decreased.

The number of households receiving and paying transfers from other households has decreased and the sums paid have increased. This might have something to do with people having greater financial possibilities for helping their relatives with larger sums than before.

Less people had to pay taxes on wealth but the amount went up.

Total household income increased by approximately 10% in 2013.

8.2.1. Length of comparable time series

Not requested by Reg. 28/2004

8.3. Comparability - domain

Not requested by Reg. 28/2004

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

This section will compare the EU-SILC 2013 data to various external sources, including EU-SILC 2012, National Accounts (NA), the Labour Force Survey (LFS), wage statistics and social protection statistics.

The LFS is a continuous survey, which is carried out according to the common EU methodology since 1995. The yearly sample size is about 12,000 working aged persons. From 2006, LFS is carried out using CAPI. LFS is the main source for labour market information.

Wage statistics have in their current form been continuously calculated since 1992. All enterprises employing 50 persons or more are obliged to provide data. A sample is drawn from smaller enterprises. Wage data is used to calculate hourly and monthly wages, both gross and net, as well as labour costs. All figures have been converted into full-time units.

To receive statistics on social protection, all institutions are enumerated whose fields of activity are related to pensions, social insurance benefits, health insurance, social welfare (social welfare institutions, children left without parental care) and other services.

Annexes:

Annex- Coherence

9.1.1 Coherence - sub annual and annual statistics

Not requested by Reg. 28/2004

9.1.2. Coherence - National Accounts

In section 9.1.

9.2. Coherence - internal

Not requested by Reg. 28/2004

10. Cost and Burden

Top

Not requested by Reg. 28/2004

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. 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

Sampling frame for selection of new part of the sample was the 2011 Population and Housing Census. Census involves the data of all permanent residents, households and dwellings in Estonia. Following persons are not considered permanent residents of Estonia: persons, who have already left Estonia (according to the statement of their relatives), persons permanently residing abroad (despite their wish to enumerate themselves) and persons, who reside in Estonia temporarily (3–12 months), but whose permanent residence is abroad. Census was conducted by Statistics Estonia.

12.1.1. Sampling design and procedure

Type of sampling design

The design used is one-stage stratified unequal probability sampling of households, with a household selected with probability proportional to the number of persons aged 14+ in it. It is because a sample of persons aged 14+ (so called address-persons) is selected first with equal probabilities within strata, and then the household of the selected person is identified, and all eligible persons in the household are interviewed. Stratification is done by geographical region.

Stratification and sub stratification criteria

Geographical stratification was used. The counties (and capital Tallinn) were grouped into three strata by the population size:

1. big counties:Tallinn, Harju (excludingTallinn), Ida-Viru, Lääne-Viru, Pärnu,Tartu;
2. small counties: Jõgeva, Järva, Lääne, Põlva,Rapla, Saare, Valga, Viljandi, Võru;
3. Hiiu County formed a separate stratum as the smallest county with the population size times smaller of the next smallest.

Sample selection schemes

Systematic sampling of address-persons in each stratum. For households this procedure results in unequal probability sampling with inclusion probabilities proportional to household size (number of persons aged 14+ in it).

Sample distribution over time

Fixed income reference period was used and therefore the sample was not principally divided into months or weeks. The fieldwork period was from February to May 2013. For the convenience of fieldwork administration, the sample was allocated into the four months with proportions approx. 3:3:3:1. When allocating households into the months of fieldwork period, uniform workload of interviewers was targeted. Actual month of interview is nevertheless different from the month the household was allocated to. Due to lack of interviewers in some areas, ca 5% of households was interviewed after the official end of fieldwork period in June 2013.

12.1.2. Sampling unit

One stage sampling design was used. Households are regarded as sampling units although selection was made using the sample of address-persons.

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 - 7566 households

- the achieved sample size which is the number of observed sampling units (household or individual) with an accepted interview - 5775 households and 12348 persons (203 persons formed within-household non-response and their income was fully imputed)

- the effective sample size is thus 4511 households and 9646 persons. (According to Commission regulation we use here the design effect of the at-risk-of-poverty rate, which was 1.28). Minimum requirements are thus satisfied (3500 households and 7750 persons).

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.

12.2. Frequency of data collection

Data is collected every year.

The sample was not principally divided into months or weeks. The fieldwork period was from February to May 2013. For the convenience of fieldwork administration, the sample was allocated into the four months with proportions approx. 3:3:3:1. When allocating households into the months of fieldwork period, uniform workload of interviewers was targeted. Actual month of interview is nevertheless different from the month the household was allocated to. Due to lack of interviewers in some areas, ca 5% of households was interviewed after the official end of fieldwork period in June 2013.

12.3. Data collection

Mode of data collection

A description of the mode of data collection used in your country. Please mention if you use mixed mode of data collection.

1-PAPI (% of total)	2-CAPI (% of total)	3-CATI (% of total)	4-Self administrated (% of total)
0.6	98.5	0.9	0

The mean interview duration

The mean interview duration per household is calculated as the sum of the duration of all household interviews plus the sum of the duration of all personal interviews, divided by the number of household questionnaires completed. Only households accepted for the database have to be considered.

Average interview duration = 50 min

12.4. Data validation

Not requested by Reg. 28/2004

12.5. Data compilation

Not requested by Reg. 28/2004

12.5.1. Weighting procedure

Design factor Non-response adjustments Adjustment to external data Final cross sectional weights

12.5.2. Estimation and imputation

Imputation procedure used Imputed rent Company car

12.6. Adjustment

Not requested by Reg. 28/2004

13. Comment

Top

No comments

Related metadata

Top

Annexes

Top

Annex on data collection
Annex on imputation and estimation
Annex on weighting
Annex on item non-response
Annex on sampling errors