



EUROPEAN COMMISSION  
EUROSTAT

Directorate F: Social Statistics  
Unit F-4: Income and living conditions; Quality of life

ESTAT.F.4 (2021)

**EU-SILC**

**Guidelines for Data Submission**

**and Data Dissemination according to IESS legislation**

**version July 2021**

## Contents

I.	FOREWORD.....	4
II.	DATA SUBMISSION.....	4
A.	Transmission deadlines .....	4
1.	Regular data.....	4
2.	Provisional data .....	5
3.	Optional variables to measure impact of COVID-19 .....	5
B.	Quality data submission .....	5
C.	Submission structure .....	6
1.	Transmission format.....	6
2.	Means and filename structure of transmission .....	7
3.	Submission content.....	8
4.	Key variables .....	8
5.	Longitudinal data.....	9
D.	Variables.....	10
1.	New variables .....	10
2.	Household grid .....	10
3.	Split variables .....	11
4.	Change in variable identifier .....	11
5.	Change of flags of otherwise unchanged variables .....	12
6.	Changes in modalities of otherwise unchanged variables.....	12
7.	Removed variables .....	12
8.	Variables not pertaining to the annually collected nucleus variables.....	13
E.	Weights.....	13
F.	Transitional arrangements .....	13
1.	Derogations on data transmission.....	13
2.	Rotation schemes that include longitudinal data of operations before 2021 .....	14
III.	DATA DISSEMINATION.....	16
A.	Eurobase dissemination .....	16
B.	UDB data dissemination.....	16
IV.	ANNEXES .....	18
A.	Annex I: Comparison of nucleus variables before and after introducing the IESS-related legislation (operation 2021 vs. operation 2020).....	18
B.	Annex II: Example of data file including non-household respondents .....	19

1.	Extract of data file including non-household respondents .....	20
2.	Explanatory notes for first two example households .....	23
3.	Transmission of files .....	27
4.	Weights .....	29

## I. FOREWORD

The revised legal basis for the EU-Statistics on Income and Living Conditions (EU-SILC) has brought about changes to the way in which National Statistics Institutes (NSIs) of Member States report their microdata to Eurostat. The **METHODOLOGICAL GUIDELINES AND DESCRIPTION OF EU-SILC TARGET VARIABLES** (DocSILC 065), from the 2021 SILC operation onwards, put in place IESS-related legislation on the survey implementation.

These guidelines detail the legislative provisions further, with specific arrangements for data submission and data dissemination related to enhanced timeliness, precision requirements, transitional arrangements<sup>1</sup> and content modifications brought about by the IESS-related legislation. Modifications include the revision of variable identifiers or modalities, flag changes, addition of new variables, implementation of periodically recurring modules and of new policy needs (ad hoc) variables, deletion of variables or their shift from the SILC nucleus to periodically recurring or new policy needs modules.

These arrangements are a precondition for Eurostat to put in place validation rules for the data collection year 2021 and beyond.

Annex I shows a comparison of nucleus variables before and after introducing the IESS-related legislation (operation 2021 vs. operation 2020).

## II. DATA SUBMISSION

### A. Transmission deadlines<sup>2</sup>

#### 1. *Regular data*

Pursuing annex V of the IESS regulation, Member States shall submit for the Income and Living Conditions domain pre-checked microdata without direct identifiers, according to the following deadlines<sup>3</sup>:

- By the end of year N, submission of all cross-sectional and longitudinal target variables for the data collection of year N, including cross-sectional weights.
- In exceptional cases, microdata concerning income variables may be submitted as provisional data.
- By 28 February of year N+1, receipt of revised, final income data is expected.

---

<sup>1</sup> Transitional arrangements are warranted for countries granted derogations on deadlines for microdata delivery as well as for the period during which reported rotational schemes include data relating to years before the IESS implementation in 2021.

<sup>2</sup> Provisions for countries with temporary derogations are outlined in subsection “Transitional arrangements”.

<sup>3</sup> IESS [Regulation \(EU\) 2019/1700 establishing a common framework for European statistics relating to persons and households, based on data at individual level collected from samples](#)

- By the end of October of year N+1, at the latest, longitudinal weights are to be submitted to complete the data files.

## 2. *Provisional data*

- By the end of year N, provisional data **on income only** can be submitted. These will be processed and uploaded on Eurobase, flagged “p”, until regular, final data are released.
- Countries are invited to inform Eurostat until end of September of year N whether the data files of year N to be sent by the end of year N will include income data as provisional only.

## 3. *Optional variables to measure impact of COVID-19*

The optional variables to measure the impact of COVID-19 in year N of the EU-SILC data collection are to be submitted by the end of year N.

- Countries that comply with the IESS transmission deadline for regular data, should send these optional variables together with the regular data set of the survey year N.
- In case a country submits the regular data at a later date because it was granted a derogation from the data transmission deadlines, the optional variables should be sent twice: First at the end of year N, secondly included in the regular data set once these are transmitted.

For further details see the dedicated transmission guideline for the voluntary module on impact of COVID-19 in EU-SILC 2021<sup>4</sup> and, once agreed, such guidelines for the following year(s).

### **B. Quality data submission**

The data set sent by the countries is expected to be complete and in a quality that allows dissemination. To ensure this objective, Member States should pre-check the microdata against common validation rules, wherever available, before submitting these to Eurostat. To this end, Eurostat shares validation programs in SAS format with the Member States. These are updated yearly. Eurostat provides support in using them.

---

<sup>4</sup> Working Group “Statistics on Income and Living Conditions” meeting document (Nov. 2020), Item 2.1 Voluntary module on impact of COVID-19 in EU-SILC 2021 data collection and Item 2.2 Transmission guidelines of optional variables related to the COVID-19 crisis in EU-SILC 2021.

## C. Submission structure<sup>5 6</sup>

### 1. Transmission format

Data should be submitted as regular reconciled files, including a complete data set of cross-sectional variables of the year of operation as well as longitudinal data of previous waves of a rotation scheme, as

- Household Register (D)
- Household Data (H)
- Personal Data (P)
- Personal Register (R).

These **four** files (D-, H-, R- and P-data file) include all yearly nucleus variables, as well as all periodically recurring and new policy needs modules.

The file length for each given year is standard, equal across all data-submitting countries. In case that optional variables are not collected by a given country, the flag “-8” should be used.

The household grid variables should be transmitted with **22** entries which corresponds to the maximum household size detected in EU-SILC data.

The four files shall always be submitted together (even if only some of them changed) in csv format (comma separated values) via Edamis to Eurostat so to facilitate automatized structural validation in future.

The following rules apply:

Header row	first record containing the variable names
Delimiter of variables	comma (,)
Decimal separator	decimal point (.)
Character values	must NOT be enclosed by quotes
Blank variables	represented by nothing between the commas (...,...)
The first three variables	Year, Country and ID (for the rest of the variables no

---

<sup>5</sup> The legal basis is art. 8 “Formats for transmission of information” of IESS [Regulation \(EU\) 2019/1700 establishing a common framework for European statistics relating to persons and households, based on data at individual level collected from samples](#) and the delegated or implementing acts adopted pursuant thereto.

<sup>6</sup> For reference see art. 10 “Formats for transmitting information” of [Commission Implementing Regulation \(EU\) 2019/2242 of 16 December 2019 specifying the technical items of data sets, establishing the technical formats and specifying the detailed arrangements and content of the quality reports on the organisation of a sample survey in the income and living conditions domain](#)

fixed order is required)

e.g.

DB010,DB020,DB030,DB090,DB090\_F,...

2021,AT,12300,10.5,1,...

2021,AT,12301,4.1305,1,...

2021,AT,22400,1.1,1,...

Art. 10.3 of the IR 2019/2242 stipulates “*Member States shall annually transmit all the subsamples belonging to the rotational design of the given year, regardless of their duration. All the subsamples shall be transmitted together*”.<sup>7</sup>

A detailed example on how to submit a data file that also includes non-household members is provided in annex II. It is based on the new policy needs module 2021 “Living arrangements and conditions of children in separated and blended families”.<sup>8</sup>

Please note that after IESS implementation Eurostat will no longer process “early” submission data.

## 2. Means and filename structure of transmission

The means to transmit data is Edamis, Eurostat’s single data entry point.

The filename structure is

**SILC\_<Transmission type><File type>\_A\_<Country name>\_<Year>\_0000\_V<Sequence number>.CSV**

Where<sup>9</sup>:

< Transmission type >	:	R
<File type>	:	{D/H/P/R}
<Country name>	:	2 characters
<Year>	:	4 digits
<Sequence number>	:	4 digits

*Example:*     SILC\_RD\_A\_AT\_2021\_0000\_V0001.CSV  
                  SILC\_RH\_A\_AT\_2021\_0000\_V0001.CSV  
                  SILC\_RR\_A\_AT\_2021\_0000\_V0001.CSV  
                  SILC\_RP\_A\_AT\_2021\_0000\_V0001.CSV

<sup>7</sup> Separate datasets for survey years up to 2020 on the one hand, and for 2021 onwards on the other, as well as a separate submission of module data files cannot be accepted for the sake of consistency, to avoid errors and to keep administrative burden in validation and indicator programming, in dissemination and communication with interested parties at a minimum.

<sup>8</sup> For reference see LC-ILC/251/19/EN, Working Group “Statistics on Income and Living Conditions” meeting document (Dec. 2019), Item 2. Methodological guidelines for the SILC 2021.

<sup>9</sup> If the 2021 mini module with COVID-19 related variables is submitted separately, its filename structure will be the same as in the regular transmission, except one difference: 'RD', 'RH', 'RP' and 'RR' parts have to be substituted by 'M21D', 'M21H', 'M21P' and 'M21R'.

### 3. *Submission content*

Continuing the practice before IESS implementation, the household register file (D) must contain every household (selected + substituted + split off), also those where the address could not be contacted or which could not be interviewed.

The replacement of a sample household unit “x” against another household unit “y” is not required to be registered in the submitted data set. However, records of such household substitutions shall be kept in the national records and substitution arrangements duly described in the respective quality report.

In the other files, records related to a household will only exist if the household has been contacted (DB120 = 11 (or DB110 = 1)) AND has a completed household interview in the household data file (H) (DB130 = 11) AND at least one member has complete data in the personal data file (P) (RB250 = 11, 12, 13 or 14 => DB135 = 1). This member must be the selected respondent (RB245 = 2) if this mode of selection is used.

The personal register file (R) must contain a record for every person currently living in the household, even if the person is temporarily absent.<sup>10 11</sup>

The personal data file (P) must contain a record for every eligible person for whom the information could be completed from interview and/or registers (RB245 = 1, 2 or 3) AND (RB250 = 11, 12, 13 or 14).

### 4. *Key variables*

The four D, H, R and P files have to be adequately linked:

All observations from the P file must have a univocal link to the R file (and to D and H files via the R file).

All observations from the R file must have a univocal link to a D file observation (and to an H-file observation if the household was interviewed).

Also all observations from the H file must have a univocal link to a D file observation.

For that purpose, the variables:

- ‘Year’: DB010, RB010, HB010 and PB010;
- ‘Country’: DB020, RB020, HB020 and PB020;

---

<sup>10</sup> DocSILC 065 provides rules about the inclusion of former household members that died or moved out from household, namely on inclusion of all selected persons, including those where the address could not be contacted or persons which could not be interviewed.

<sup>11</sup> An example on how to submit a data file that also includes module data on non-household respondents is provided in annex III.



- ‘Household ID’: DB030 and HB030 and RB040;
- ‘Personal ID’: RB030 and PB030.

are used as key variables.

The Personal ID is constructed with the Household ID and two more digits.

The Personal ID does not change, not even if the person moves to another household. While the Household ID due to a split-off or fusion can get modified or changed.

The link between Household ID and Personal ID can be established using variable RB040: Current Household Identification (ID).

Each file must contain the list of variables with identifier values, *flags and imputation factors*, as defined in the methodological guidelines for the respective operation year.

For further details, see art. 10 of the Implementing Regulation (EU) 2019/2242 referenced above.

## 5. *Longitudinal data*

As a general rule, variables in the longitudinal data set shall be submitted with the variable value, as well as flag and imputation factor if applicable, corresponding to the methodological guidelines of the respective data collection year.

Eurostat is able to validate variables with different modalities, such as flags, in different years of the same SILC rotation scheme.

Detailed guidance on the various types of variables can be found in section D. “Variables”. Transitional arrangements are specified in subsection F.2.

### a) Nucleus variables

Nucleus variables shall be reported according to their data collection year, using the guidelines on variables, flags, imputation factors of the respective data collection year.

Example: **Data** on nucleus variables **collected in 2020** should be prepared and submitted to Eurostat according to the guidelines for the year 2020 in the data submission of 2020, 2021 and consecutive years.

### b) Modules variables

Periodically recurring and new policy needs modules variables should be reported in the regular file only if they are collected in a given operation year.

Within each rotation scheme lasting four years or more, at least one of the 3-yearly periodically recurring modules will be collected twice.

Collected module data should be transmitted to Eurostat in a way that allows linking all modules' data related to the same respondents over years.

## D. Variables

As a general rule, *variables should be prepared and submitted with the modalities and flags corresponding to the methodological guidelines of the respective data collection year.*

*The flag '-7' should be used for all variables that are not collected in a respective operation year, namely*

- New variables introduced in 2021 (change in identifier or introduced as new ones as from 2021 onwards);
- Modules variables (periodic 3-yearly or 6-yearly or new policy needs modules);
- Removed variables, decided by regulation not to be collected from 2021 onwards.

Transitional arrangements relating to the reporting of rotation schemes that include longitudinal data of operations before IESS implementation (before-2021 waves) are outlined in the subsections and summarised in subsection F.2 "Rotation schemes that include longitudinal data of operations before 2021".

### 1. New variables

The principle that *variables shall be submitted with the variable value, as well as flag and imputation factor if applicable, as defined in the methodological guidelines of the operation year* applies to variables that are newly added in the 2021 data collection to the annually collected core set of SILC ("nucleus"), namely

- DB076, HB110, HB120, HB130, HS022, HY121G/HY121N, PB205, PB230, PB240, PB260, PB265, PB270, PL145, PL271, RB032, RB065, RB066, RB081, RB082, RB083, RB280<sup>12</sup>, RB285, RB290<sup>13</sup>

For these variables the flag '-7' should be used for operation years before 2021 if the variables were not yet collected on a voluntary basis under the same identifier.

### 2. Household grid

The household grid is built as a matrix:

- RG\_Z#: GRID

---

<sup>12</sup> moved from P-file to R-file, instead of previously used PB210

<sup>13</sup> moved from P-file to R-file, instead of previously used PB220A

To ensure equal length of the data file across all countries, countries are asked to transmit household grid variables with exactly 22 entries. This number corresponds to the maximum household size detected in EU-SILC data.

For these variables the flag ‘-7’ should be used for operation years before 2021 if not yet collected on a voluntary basis.

### 3. Split variables

The principle that *variables shall be submitted according to the methodological guidelines of the operation year* should be understood for split variables PL040, PL051 and PL111 as follows:

- In the cross-sectional data set of 2021 and following data collection years the split variables should be reported as PL040A, PL040B, PL051A, PL051B, PL111A, PL111B.
- Longitudinal data for these variables, collected in 2020 or in earlier years, should be reported as PL040, PL051, PL111.

The flag ‘-7’ should be used for the variables that are no more / not yet collected in an operation year.

### 4. Change in variable identifier

The principle that *variables shall be submitted according to the methodological guidelines of the operation year* should be understood for changes in the variable identifier as follows:

- In the cross-sectional data set of 2021 and following data collection years, the variable identifiers should be reported as PE021, PE041, PL016, PL032, PL141, RB211.
- Longitudinal data, collected in 2020 or in earlier years, should be reported as PE020, PE040, PL015, PL031, PL140, RB210.

The flag ‘-7’ should be used for the variables that are no more / not yet collected in an operation year.

**Table 1:** Example submission of data and flags for variable PL031/PL032 in the data file of the year 2023 (e.g. 6 year panel)

	Survey year	PL031	PL031_F	PL032	PL032_F
Longitudinal data	2023	.	-7	X	1
Longitudinal data	2022	.	-7	X	1
Longitudinal data	2021	.	-7	X	1
Longitudinal data	2020	X	1	.	-7

Longitudinal data	2019	X	1	.	-7
Longitudinal data	2018	X	1	.	-7

5. *Change of flags of otherwise unchanged variables*

Variables shall be submitted with the flag that corresponds to the methodological guidelines of the operation year of their collection.

6. *Changes in modalities of otherwise unchanged variables*

Variables shall be submitted with the modalities according to the methodological guidelines of the operation year of their collection.

7. *Removed variables*

Nucleus variables removed from the annually collected SILC data files due to IESS implementation are the following:

- HB080, HB090, HH031, HS070, HS080, PB130, PB220B, PL180, PY031G, PY200G, RB031, RB070, RB140, RB150, RB160, RB170, RB180, RB190, RB260, RB270.

For information: There are variables that moved between the data files, namely from the R file to the P file and vice versa. Concerned identifiers are e.g. RB260 and RB270, which are replaced by newly collected variables in the P file (PB260, PB265, PB270).

The nucleus variables that are no longer collected in 2021 since they are moved from the nucleus to a module are HH040, HH061, HH081, HH091, HS130, HS140, HS160, HS170, HS180, HS190, HY030G/HY030N, PE030, PL020, PL025, PL035, PL120, PL130, PL160, PL170.

Applying the principle that *variables shall be submitted according to the methodological guidelines of the operation year*, pre-2021 longitudinal data should include the removed variables with modalities and flags corresponding to the methodological guidelines of the operation year. The flag value for removed variables for the years 2021 and following would be ‘-7’, meaning “Not applicable because the variable was not collected that year”.

8. *Variables not pertaining to the annually collected nucleus variables*

Periodically recurring and new policy needs (ad hoc) modules variables shall be reported only if these are described in the methodological guidelines of the operation year.

Modules variables shall be delivered in the same file as annual nucleus variables, so to ensure a consistent approach, to avoid errors and to keep administrative burden in validation and indicator programming, dissemination and communication with interested parties at a minimum.

Example: Variables of the 6-yearly recurring module “Quality of Life”, described in the methodological guidelines of the year 2022 will be reported in the submission files for the operation year 2022, but not in the ones of the survey years 2023 to 2027.

As detailed in subsection C.5.b), as a consequence that the Personal ID does not change, modules’ data related to the same respondents should be transmitted in a way that allows linking over years.

The flag ‘-7’ should be used for these variables which are no more collected.

**E. Weights**

Weights shall be submitted as outlined in section 1 “Submission deadlines”.

Each file must contain respective weights as defined in methodological guidelines for the respective operation year.

**F. Transitional arrangements**

*1. Derogations on data transmission*

Some countries were granted a derogation on deadlines for microdata delivery during a maximum of 3 years after the implementation of Regulation 2019/1700. Member States granted such extended submission deadline during an agreed period are invited to provide provisional data files as early as possible, before the agreed data submission deadline. These data will be flagged “p” on Eurobase until the regular, final data files are published.

2. *Rotation schemes that include longitudinal data of operations before 2021*

Variables shall be submitted with the variable value, as well as flag and imputation factor if applicable, corresponding to the methodological guidelines of the respective data collection year:

- Pre-2021 waves' longitudinal data shall be reported with the variable value, as well as flag, valid before 2021, of the respective collection year.
- 2021 and following years' longitudinal data should be reported with the variable value, as well as flag and imputation factor if applicable, valid from 2021, of the respective collection year.

Whenever variables are not specified in the methodological guidelines of a given operation year, these variables shall not be reported in the data set relating to this survey year.

As an example, the content of the regular data file 2021 that includes longitudinal data of operations before IESS implementation (pre-2021 waves) is detailed in the graph and table below.

**Graph 1:** Organisation of EU-SILC cross-sectional and longitudinal data, exemplified by data submission in the year 2021, based on a six waves' rotation scheme

Rotation schemes (W1 - W6)							
<b>All waves</b>							
2015	W6						
2016	W5	W6					
2017	W4	W5	W6				
2018	W3	W4	W5	W6			
2019	W2	W3	W4	W5	W6		
2020	W1	W2	W3	W4	W5	W6	
2021		W1	W2	W3	W4	W5	W6
2022			W1	W2	W3	W4	W5
2023				W1	W2	W3	W4
2024					W1	W2	W3
2025						W1	W2
2026							W1
<b>Actual wave</b>	<b>2020</b>	<b>2021</b>	<b>2022</b>	<b>2023</b>	<b>2024</b>	<b>2025</b>	<b>2026</b>

Note W1: sub-sample of people participating for the first year in the survey  
W2: sub-sample of people participating for the 2nd year in the survey

...

**Table 2:** Rotation scheme 2021, including longitudinal data of operations before IESS implementation (pre-2021 waves)

	Content of 2021 data file	Data collection year =
		Year of applicable methodological guidelines (doc65)
pre-IESS regulated data submission	1st wave of rotation scheme that started in survey year 2016	2016
	2nd wave of rotation scheme that started in survey year 2016	2017
	3rd wave of rotation scheme that started in survey year 2016	2018
	4th wave of rotation scheme that started in survey year 2016	2019
	5th wave of rotation scheme that started in survey year 2016	2020
	1st wave of rotation scheme that started in survey year 2017	2017
	2nd wave of rotation scheme that started in survey year 2017	2018
	3rd wave of rotation scheme that started in survey year 2017	2019
	4th wave of rotation scheme that started in survey year 2017	2020
	1 <sup>st</sup> wave of rotation scheme that started in survey year 2018	2018
	2 <sup>nd</sup> wave of rotation scheme that started in survey year 2018	2019
	3 <sup>rd</sup> wave of rotation scheme that started in survey year 2018	2020
	1 <sup>st</sup> wave of rotation scheme that started in survey year 2019	2019
	2 <sup>nd</sup> wave of rotation scheme that started in survey year 2019	2020
	1 <sup>st</sup> wave of rotation scheme that started in survey year 2020	2020
IESS regulated data	1 <sup>st</sup> wave of rotation scheme starting in survey year 2021	2021
	2 <sup>nd</sup> wave of rotation scheme that started in survey year 2020	2021
	3 <sup>rd</sup> wave of rotation scheme that started in survey year 2019	2021
	4th wave of rotation scheme that started in survey year 2018	2021
	5th wave of rotation scheme that started in survey year 2017	2021
	6th wave of rotation scheme that started in survey year 2016	2021

Continuing the practice before IESS implementation, R and D files are linked by the variables RB040 and DB030. In case of split-off households, the persons who leave the initial household will have two observations in the R file.

### **III. DATA DISSEMINATION**

#### **A. Eurobase dissemination**

Eurostat will release on Eurobase according to usual practice all indicators with available breakdowns. If longitudinal weights are submitted in the first submission step by end of year N, not only cross-sectional, but also longitudinal indicators are already calculated and published along cross-sectional ones. For provisional data, a “p” flag is added to the released data until delivery of final data.

#### **B. UDB data dissemination**

The user database (UDB) containing newly added cross-sectional microdata of year N is disseminated in March of year N+1, for all the country data received from in quality satisfactory for dissemination until end of January of year N at the latest.

The UDB containing newly added longitudinal microdata of year N is disseminated in November of year N+1 for all the longitudinal data received in quality satisfactory for dissemination until at least end of September of year N+1.

The dates can be adjusted, depending on the extent the data transmission deadline is respected and quality requirements are achieved, in order to increase the completeness (the number of countries) of the UDB.

Before the 2021 IESS implementation, the module data have been part of the cross-sectional part of the UDB. The files in the cross-sectional part:

- contain data of one year only.
- Household and person identifiers are randomized during the anonymization process for a number of countries to ensure agreed disclosure control.

As a consequence, the module data (including the 3-yearly periodically recurring modules' variables) cannot be linked together for more than one year in the cross-sectional UDB, in contrast to the specified under sect. II.C.5.b.

If such link is to be ensured, without linking the rest of the cross-sectional data for more than one year, the only possibility is to create separate files (with independent anonymization). These files would contain the module data (and some additional fields that have to be specified) only, for each periodically recurrent module of each survey year in which such module data are transmitted. Otherwise, all cross-sectional data of the different years will be possible to link together.

A possible solution could be to discontinue the differentiation between the longitudinal and cross-sectional parts of the UDB, starting from the survey year 2021.



The consequences of that are the following:

- Starting from survey year 2021, one (regular) UDB will be prepared only, for each survey year. It is suggested to do so in two steps: first step (March) without longitudinal variables (weights), second step (November) completed with longitudinal data.
- For each survey year, the regular UDB files will contain data (cross-sectional and longitudinal) of those year(s) that the country transmitted for that survey year (4 years or more, depending on the panel duration of the respective country).
- The identifiers will not be randomised anymore, so the researchers will be able to follow all (also the cross-sectional) variables of a household/person for 4 (or more) years.

#### IV. ANNEXES

##### **A. Annex I: Comparison of nucleus variables before and after introducing the IESS-related legislation (operation 2021 vs. operation 2020)**

- separate document -

**Annex II** (overleaf)

**B. Annex II: Example of data file including non-household members (children)**

For variables of the new policy needs module 2021 (Living arrangements and conditions of children in separated and blended families), the tables below exemplify how to submit data for different household situations, including non-household respondents, respecting one coherent data structure.

In the tables, each colour represents a household with a household ID (HH ID) attached. Household members are identified by simplified personal IDs (Pers ID). Variables HK010 and HK020 are reported at household level. They are essential for subsequent ones.

Whereas in the module “Intergenerational transmission of advantages and disadvantages”, a person could have only one mother and only one father, a condition that allows to record the answers in one row (link to the one person-respondent), in the 2021 new policy needs module a person could have more than one child living outside of a sampled household. To duplicate in such a case the parent in the data set would significantly complicate the validation of the data.

For the treatment of non-household members in variables RB010, RB020, RB030, RB040, RB110 (other nucleus variables will have missing values and flags) please see subsection 3. (Transmission of files) and consult the methodological guidelines DocSILC 065.

1. Extract of data file including non-household respondents

HH ID	Person's ID	Age of the person	Question used to construct HK010
			<i>Q1. Do any children (aged 0-17) living in your household (being a household member) have a parent who lives outside the household (who is not a household member)?</i> 1- Yes >> go to Q2 (Number) 2- No > Skip to HK020
1	1	50	No
1	2	55	
1	3	17	
2	1	65	Yes
2	2	25	
2	3	4	
3	1	40	No
3	2	38	
3	3	4	
3	4	12	
3	5	34	
4	1	26	No
5	1	34	Yes
5	2	16	
5	3	1	
5	4	36	
5	5	36	
5	6	5	

HH ID	HK010	HK010_F	Comments
	<i>Q2. How many children (aged 0-17) who are household members have the other parent lives outside the household?</i> _____ (number)		
1	0	1	There is no child in the household who has a parent outside the household
2	1	1	There is one child in the household who has a parent outside the household
3	0	1	There is no child in the household who has a parent outside the household
4	.	-2	There are no children 0-17 years old in the household
5	2	1	There are two children (ID 2 and 3) in the household who have a parent outside the household and 1 child with both parents inside the household

HH ID	Pers.ID	Age of the person	Q1 of HK010	HK010	HK010_F	Q1 of HK020	HK020 (Q2 of HK020)	HK020_F	PK010	PK010_F	RK010	RK010_F <sup>14</sup>	RK020	RK020_F <sup>15</sup>	PK020	PK020_F						
			<i>Do any children (aged 0-17) living in your household (HH member) have a parent who lives outside the household (not HH member)?</i> 1-Yes>>go to Q2 (Number) 2- No> Skip to HK020	<i>Number of children who are household members and whose other parent lives outside the household</i>		<i>Does any household member have a child (age 0-17) who lives outside the household?</i> 1-Yes>>go to Q2 (Number) 2- No> Skip to next question	<i>Number of children who are not household members and whose parent lives inside the household</i>		<i>Household member has children who are not household members</i>		<i>Parent ID and sequential number of the child who is not household member</i>		<i>Age of the child who is not household member</i>		<i>Main reason for not spending more time with children who are household members</i>							
1	1	50							.	-2	.	-2	.	-2	1-8	1						
1	2	55	No	0	1	No	0	1	.	-2	.	-2	.	-2	1-8	1						
1	3	17							.	-2	.	-2	.	-2		-2						
2	1	65	Yes	1	1	Yes	2	1	2	1	.	-4	.	-4	.	-2						
2	2	25							1	1	.	-4	.	-4	1-8	1						
2	201	6														201	1	6	1			
2	202	6														202	1	6	1			
2	3	4														.	-4	.	-4			
3	1	40	No	0	1	Yes	1	1	2	1	.	-4	.	-4	1-8	1						
3	2	38							2	1	.	-4	.	-4	1-8	1						
3	3	4														.	-4	.	-4			
3	4	12														.	-4	.	-4			
3	5	34													1	1	.	-4	.	-4	.	-2
3	501	8									501	1	8	1								
4	1	26	No	.	-2	No	0	1	.	-2	.	-2	.	-2	.	-2						
5	1	34	Yes	2	1	No	0	1	.	-2	.	-2	.	-2	1-8	1						
5	2	16							.	-2	.	-2	.	-2	.	-2	.	-2	.	-2		-2
5	3	1														.	-2	.	-2			
5	4	36													.	-2	.	-2	.	-2	1-8	1
5	5	36													.	-2	.	-2	.	-2	1-8	1
5	6	5							.	-2	.	-2	.	-2								

<sup>14</sup> When both flags '-2' and '-4' are applicable to the respondent, priority will be given to the flag '-2'.

<sup>15</sup> Idem

HH ID	Pers ID	Age of the person	PK030 Main reason for not spending more time with children who are not household members	PK030_F	RK030 Usual time the parent needs to get to the child who is not a household member	RK030_F	RK040 Frequency of contact (via phone, social media etc.) during the last 12 months with the child who is not a household member	RK040_F	RK050 Child having a bedroom in which to sleep (including bedroom shared with siblings)	RK050_F	RK060 Frequency of spending actively time with the child (meals, play, homework, walking, talking, etc.)	RK060_F	RK070 Number of nights per month the child spends in the household (average)	RK070_F	RK080 Legal child custody situation	RK080_F
1	1	50	.	-2	.	-2	.	-2	.	-5	.	-5	.	-2	.	-2
1	2	55	.	-2	.	-2	.	-2	.	-5	.	-5	.	-2	.	-2
1	3	17		-2	.	-2	.	-2	1-3	1	1-7	1	.	-2	.	-2
2	1	65	.	-2	.	-4	.	-4		-5		-5	.	-5	.	-5
2	2	25	1 - 8	1	.	-4	.	-4		-5		-5	.	-5	.	-5
2	201	6			minutes	1	1- 6	1	1-3	1	1-7	1	0-31	1	1- 4	1
2	202	6			minutes	1	1- 6	1	1-3	1	1-7	1	0-31	1	1- 4	1
2	3	4				-4		-4	1-3	1	1-7	1	0-31	1	1- 4	1
3	1	40	.	-2	.	-4	.	-4	.	-5	.	-5	.	-5	.	-5
3	2	38	.	-2	.	-4	.	-4	.	-5	.	-5	.	-5	.	-5
3	3	4			.	-4	.	-4	1-3	1	1-7	1	0-31	1	1- 4	1
3	4	12			.	-4	.	-4	1-3	1	1-7	1	0-31	1	1- 4	1
3	5	34	1- 8	1	.	-4	.	-4	.	-5	.	-5	.	-5	.	-5
3	501	8			minutes	1	1- 6	1	1-3	1	1-7	1	0-31	1	1- 4	1
4	1	26	.	-2	.	-2	.	-2	.	-5	.	-5	.	-2	.	-2
5	1	34	.	-2	.	-2	.	-2	.	-5	.	-5	.	-5	.	-5
5	2	16	.	-2	.	-2	.	-2	1-3	1	1-7	1	0-31	1	1- 4	1
5	3	1			.	-2	.	-2	1-3	1	1-7	1	0-31	1	1- 4	1
5	4	36	.	-2	.	-2	.	-2	.	-5	.	-5	.	-5	.	-5
5	5	36	.	-2	.	-2	.	-2	.	-5	.	-5	.	-5	.	-5
5	6	5			.	-2	.	-2	1-3	1	1-7	1	.	-2	.	-2

2. *Explanatory notes for first two example households*

**Household ID = 1**

<b>Variable</b>	<b>Modalities</b>	<b>Explanation / Flagging</b>
HHID	HHID = 1	A couple with a child aged 17 years who lives in the household.
HK010	HK010 = 0	The flag value should be '1' as it is a filled value.
HK020	HK020 = 0	No household member has a child between 0 and 17 years old who lives outside the household.
PK010	PK010 = missing	PK010 is applicable to all members of this household, as all of them are at least 16 years old. However, since this variable lists all members of the household having children who live outside the household, the value of PK010 will in this case be missing for all of the household members and the flag will be '-2', "not applicable (HK020 equal to 0)".
RK010	RK010 = missing	RK010, creating the ID of a child outside the household, would in its absence be blank. The flag value should be '-2', meaning "Not applicable (current household member)".
RK020	RK020 = missing	Since RK020 refers only to children who live outside the household, equal to RK010, the flag for this variable should be '-2', "Not applicable (HK020 equal to 0)".
PK020	PK020 = (1-8), missing	PK020 is asked to household members aged 16 and over, with at least one child of 0 to 17 years within the household. In this case, both parents would reply with one of the reasons. The flag value of this variable should be '1', "Filled". The 17-years' old child would be assigned the flag value '-2', "Not applicable (person does not have children 0-17 living in the household)".
PK030	PK030 = missing	PK030 is asked only to persons who have at least one child living outside the household. In the household at stake, no one is in such situation. Therefore, for this variable the flag value would be '-2', "Not applicable (PK010 not equal to 1)" for all household members, all aged over 16 years.
RK030, RK040	RK030 = missing RK040 = missing	These variables are filled only for households that have children outside the household, which is not the case here. A flag value of '-2' would be assigned in all cases, meaning "Not applicable (HK020 equal to 0)".
RK050	RK050 = (1-3), missing	RK050 should be filled for each child between 0 and 17 years old, including those who are not household members but have a parent inside the household. In this case, this concerns only the last member of the household. For household members 1 and 2, the flag value '-5', "Not applicable (person aged 18 and over)" should be chosen.

## Household ID = 1 continued

Variable	Modalities	Explanation / Flagging
RK060	RK060 = (1-7), missing	RK060 should be filled for each child between 0 and 17 years old, including those who are not household members but have a parent inside the household. In this case, this concerns only the last member of the household. For household members 1 and 2, the flag value '-5' "Not applicable (person aged 18 and over)" should be assigned.
RK070, RK080	RK070 = missing RK080 = missing	The variables are filled for households that have children aged 0-17 who <ul style="list-style-type: none"><li>• <b>are household members</b> but have parent <b>outside</b> the household OR</li><li>• <b>are not household members</b> but have a parent <b>inside</b> the household.</li></ul> None of these cases apply here. A flag value of '-2' is assigned in all cases, meaning "Not applicable (HK010 equal to 0 and HK020 equal to 0)".

## Household ID = 2

Variable	Modalities	Explanation / Flagging
HHID	HHID = 2	A person of 65 years, living with her child (25 years old) and the latter's child aged 4 years. The 25-year-old is also mother of twins who live with the other parent outside the household. The two children outside the household should also be attributed HHID = 2.
HK010	HK010 = 1	The flag value should be '1' as it is a filled value.
HK020	HK020 = 2	The flag value should be '1', as it is a filled value.
PK010	PK010 = 2, 1	PK010 is applicable to all household members aged 16+, that means to the persons aged 65 and 25 years respectively. For the 1st person (aged 65), the value would be '2', meaning "no" and the flag value should be '1', for "filled". For the 2nd person (aged 25), the value would be '1', meaning "yes", with a flag value of '1', for "filled".
RK010	RK010 = 201, 202, missing	RK010 creates the ID of the child(ren) outside the household. The ID code combines the value for the parent, in this case the 2 <sup>nd</sup> person in the household, represented by "2" and a sequential two-digit number for the child(ren) outside the household, meaning here "01" for the 1st child and "02" for the 2nd child. In this case, RK010 will be created for the two children outside of the household, represented by the values '201' and '202'. For both these children the flag values should be '1', for all other household members '-4'.



## Household ID = 2 continued

Variable	Modalities	Explanation / Flagging
RK020	RK020 = 6, missing	RK020 applies only to the children who live outside the household. Their age should be recorded. In the same vein as RK010, the corresponding flag will be '1'. Flag value '-4' will be attributed to all other household members.
PK020	PK020 = (1-8), missing	PK020 is asked to household members aged 16 and over, with at least one child of 0 to 17 years <b>within</b> the household. Therefore, for the 2 <sup>nd</sup> person, aged 25, one of the reasons (1-8) should be reported and the flag value should be '1' for "filled". The 1 <sup>st</sup> person, aged 65, would not be asked this variable and flag '-2' applies.
PK030	PK030 = (1-8), missing	PK030 is asked to persons aged 16 and over who have at least one child living <b>outside</b> the household. Therefore, for the 2 <sup>nd</sup> person one of the reasons (1-8) should be reported and the flag value should be '1' for "filled". The 1 <sup>st</sup> person would not be asked this variable and flag '-2' applies.
RK030	RK030 = minutes, missing	The variable RK030 is filled only for children outside the household. In this case, time in minutes related to the children with ID 201 and 202 would be recorded. For the other household members the flag value '-4', meaning "Not applicable (current household member)" applies.
RK040	RK040 = (1-6), missing	The variable RK040 is filled only for children outside the household. In this case, frequency values related to the children with ID 201 and 202 would be recorded. For the other household members the flag value '-4', meaning "Not applicable (current household member)" applies.
RK050	RK050 = (1-3), missing	RK050 should be filled for each child between 0 and 17 years old, including those not being household members but having a parent inside the household. In this case, a value should be provided for the 4-years old child within this household as well as for the two other children, who are not household members.  For the adult household members (Person IDs 1 and 2) the flag value '-5' should be provided: "Not applicable (person aged 18 and over)".
RK060	RK060 = (1-7), missing	RK060 should be filled for each child between 0 and 17 years old, including those who are not household members but have a parent inside the household. That means a value should be provided for all three children (IDs 201, 202, 3).  For household members 1 and 2, the flag value '-5' "Not applicable (person aged 18 and over)" should be selected.

## Household ID = 2 continued

Variable	Modalities	Explanation / Flagging
RK070	RK070 = (0-31), missing	<p>RK070 should be filled for each child between 0 and 17 years old, who</p> <ul style="list-style-type: none"><li>• <b>are household members</b> but have a parent <b>outside</b> the household</li><li>• <b>are not household members</b> but have a parent <b>inside</b> the household.</li></ul> <p>That means here the average number of nights should be entered for all three children (IDs 201, 202, 3).</p> <p>For household members 1 and 2, the flag value '-5' "Not applicable (person aged 18 and over)" should be selected.</p>
RK080	RK080 = (1-4), missing	<p>A value on RK080 should be filled for children aged 0-17 who</p> <ul style="list-style-type: none"><li>• <b>are household members</b> but have a parent <b>outside</b> the household</li><li>• <b>are not household members</b> but have a parent <b>inside</b> the household.</li></ul> <p>That means in this case a value should be entered for all three children (IDs 201, 202, 3).</p> <p>For household members 1 and 2, the flag value '-5' "Not applicable (person aged 18 and over)" should be selected.</p>

### 3. *Transmission of files*

Data on new policy need module variables should be submitted as part of the regular reconciled files (for details see section II.C. Submission structure). That implies that the variables of the new policy need module should be filled in for each of the transmitted files.

#### *Household Register (D)*

-

#### *Household Data (H)*

HK010: Number of children who are household members and whose other parent lives outside the household

HK020: Number of children who are not household members and whose parent lives inside the household

#### *Personal Data (P)*

PK010: Household member has children who are not household members

PK020: Main reason for not spending more time with children who are household members.

PK030: Main reason for not spending more time with children who are not household members.

#### *Personal Register (R)*

RB010: Year of the survey.

RB020: Country of residence.

RB030: Personal ID.

RB040: Current household identification (ID).

RB110: Membership status.

RK010: Parent ID and sequential number of the child who is not household member.

RK020: Age of the child who is not household member.

RK030: Usual time the parent needs to get to the child who is not a household member.

RK040: Frequency of contact (via phone, social media etc.) during the last 12 months with the child who is not a household member.

RK050: Child having a bedroom in which to sleep (including bedroom shared with siblings).

RK060: Frequency of spending actively time with the child (meals, play, homework, walking, talking etc.).

RK070: Number of nights per month the child spends in the household (average).

RK080: Legal child custody situation.

In addition, information about children whose parent(s), i.e. one or both, are members of the interviewed household but the child is not a household member, should be delivered

together with the information about the household where the parent of that children lives. This information should be included in the R-file.

The children who are not household members should have some attributes belonging to the household of their parent(s) in the EU-SILC nucleus variables, namely RB010, RB020, RB030, RB040, RB110 should be filled for children living outside of the household.

- RB010, RB020, RB040 should have the same values as the values of these variables for their parent(s).
- RB030 should be equal to RK010 (for more information see DocSILC 065).
- RB110 should be equal to 8 (for more information see DocSILC 065).

All other variables and the respective flags of R-file for all these records (except module variables) should be transmitted with missing values.

#### 4. *Weights*

Each file must contain respective weights as defined in methodological guidelines for the respective operation year. No additional weights will be used for the new policy needs module. Children living outside of the household could be analysed with the respective weight of their parent(s) which are members of the interviewed household.