

# Polish LFS person questionnaire design after IESS FR

Małgorzata Długołęcka, Hanna Strzelecka  
Statistics Poland

# Content

- Introduction
- General approach to outline the revised Polish LFS questionnaires
- Some examples of variables which made the process of questionnaire designing more problematic
  - NEEDCARE
  - WAYJFOUN and ad hoc module 2022 (on job skills) variables
  - ABSHOLID, ABSILLINJ and ABSOTHER
- Transformation table from national variables to the EU ones
- Concluding remarks

# General approach to outline the revised Polish LFS questionnaires

In the Polish LFS there are used two questionnaires (still 6.5% in paper):

- ZG Household File – concerning the household and basic demographic data of its members
- ZD questionnaire – for persons aged 15-89 years

Elaboration of the paper version is more difficult than the electronic one (especially when using complex filters), so we initiated developing a preliminary “paper” version of our forms

# General approach to outline the revised Polish LFS questionnaires

- 1) Assignment of variables required by the new regulations to appropriate forms:
  - Variables from the topic “Person and household characteristics” were placed in the ZG Household File
  - Variables from other topics (beside „technical items”) were placed in the ZD individual questionnaire
- 2) Appropriate arrangement of the variables in a given questionnaire so as to meet both the specified conditions regarding the order of variable blocks (flow-charts) and the filters used (so that the appropriate populations were covered by the required variables)

# General approach to outline the revised Polish LFS questionnaires

- 4) Placing of national variables to not violate EU variables, but in a consistent way within given thematic areas, addressed to appropriate populations (i.e. using appropriate filters that did not conflict with EU requirements)
- 5) Translation of the variables into appropriate questions and development of national LFS explanatory notes for interviewers
- 6) Complex variables => often more than one question
- 7) Complex filters for some variables => repeating the same question in several places to reach different groups
- 8) Some final EU variables derived secondarily on the basis of several questions
- 9) Testing by statistical interviewers
- 10) IT application design

# General approach to outline the revised Polish LFS questionnaires

The questionnaire ZD is structured according to the recommendations included in the Commission Implementing Regulation (EU) 2019/2240:

- the labour status module is placed at the beginning of the personal questionnaire (the questions on the demographic characteristics are generally included in the ZG Household File that is fulfilled as the first one)
- questions on labour status have been prepared according to the flowcharts and model questionnaire elaborated by Eurostat
- optional submodules – “small jobs” and questions connected with agricultural employment (production mostly for own consumption) included
- main activity status (self-perceived) is collected after the labour status module – it is placed in the last section of the core LFS questionnaire

# Employment adjustments due to agricultural production mostly for own consumption

The “agricultural employment recovery” submodule prepared by Eurostat not used due to it wasn’t included in the flowcharts and wasn’t covered even by the model questionnaire (while “small jobs” submodule was included in the model questionnaire).

Our own approach is presented below:

- Firstly go to submodule “At work” and, if necessary - “Absences from work” (for persons who declare not performing work in the reference week)
- If a person can be considered as employed >> submodule “Second or multiple jobs” and then ask about very general NACE activity categories:
  - 1) other activity than agricultural one
  - 2) agricultural activity on a private farm and
  - 3) agricultural activity outside private agriculture

# Employment adjustments due to agricultural production mostly for own consumption

- If a person declares that has only one job and works in agriculture on a private farm, we ask additional question what is the purpose of produced agricultural goods
- If the person's answer is "exclusively or mainly for own consumption", **she/he is excluded from the employed and is being redirected to the path for not employed people**, i.e. to submodules connected with searching for employment

The approach we use to "correct" the final employed population – i.e. excluding the identified people producing agricultural goods exclusively or mainly for own consumption is, in our opinion, a better solution than the opposite one – i.e. including people producing agricultural goods intended for sale or barter.



# Some examples of more problematic variables

## NEEDCARE

- In the Polish ZD questionnaire we have one question for this variable
- To fulfil the filter conditions, this question is placed in different places (with separate numbers) for each population included in the filter – questions S6A, S6B, S6C, P13A and P13B
- Since 2021 this variable has become very complicated by extending the filter, because now the same people will go through the question twice (we block it in the application, but in the paper version it needs additional notes like before question P13B, because some persons may have answered the P13A question)

# Some examples of more problematic variables

## WAYJFOUN and ad hoc module 2022 (on job skills)

Filter for the WAYJFOUN variable:

- $(EMPSTAT = 1) \text{ AND } ((YSTARTWK = REFYEAR) \text{ OR } (YSTARTWK = REFYEAR - 1 \text{ AND } 01 \leq MSTARTWK \leq 12 \text{ AND } MSTARTWK > REFMONTH))$

Filter for the AHM 2022 variables:

- $AGE = 15-74 \text{ AND } (EMPSTAT = 1 \text{ or } (EXISTPR = 2, 3 \text{ and } (REFYEAR-YEARPR = 0, 1 \text{ or } (REFYEAR-YEARPR = 2 \text{ and } 1 \leq MONTHPR \leq 12 \text{ and } MONTHPR \geq REFMONTH))))$

## Some examples of more problematic variables **WAYJFOUN and ad hoc module 2022 (on job skills)**

We implement a slightly wider filter relating to the whole years instead of checking precisely required number of months – i.e. the questions related to the variables are asked, respectively, to all respondents who started working in their current main job in the current or previous year (WAYJFOUN) or terminated their last work in the year of the survey and the two previous calendar years (AHM 2022).

In this way we obtain a slightly larger population than the one assumed for these exemplary variables, but this is made consistent with the filter at a later stage, after data collection.

The use of such detailed filtering may have another disadvantage – we can lose important information for a given variable for respondents for whom there is a lack of information on the exact month, because we can't calculate the exact period required in the filter.

# Some examples of more problematic variables

## **ABSHOLID, ABSILLINJ and ABSOTHER**

Variables regarding absenteeism from main work added in EU LFS from 2021 were a result of work of the Task force on the measurement of absences and working time in the LFS that found out that underreported absences had a marked impact on the reported hours actually worked.

The model questionnaire developed by the TF team for the LFS variables on working time (agreed by LAMAS WG in December 2015) was intended, mainly, to remind respondents about their absences in the reference week in order to improve the reporting of hours actually worked.

According to the model questionnaire, questions within these variables are directed to persons who worked during the reference week.

# Some examples of more problematic variables **ABSHOLID, ABSILLINJ and ABSOTHER**

But taking into account currently valid filter for these variables, they are ultimately have to be coded for all employed people (EMPSTAT=1).

This means that we have to code these variables also for persons who declared not performing any work during the reference week which is unnecessary - they were absent from work a whole reference week.

It seems that this filter is too wide – a reference to persons who worked during the reference week (what was applied in the model questionnaire: WSTATOR =1) would be more rational (it also limits potential errors arising when transcoding the set of domestic questions into the EU variables).

## Some examples of more problematic variables **ABSHOLID, ABSILLINJ and ABSOTHER**

It is worth to be noticed that these variables relate to the main job only.

Among persons who declare performing work during the reference week, there are also persons who worked in their additional job only, not in the main one.

This situation causes a problem, most of all during developing a questionnaire. The model questionnaire does not provide solutions for this case.

# Transformation table from national variables to the EU ones

- An appropriate transition key from domestic variables to be recoded into the required EU file
- One transition key for quarterly datasets and the second one for yearly and biennial frequencies as well as module variables
- The transition key - needed to transcode data for sending to Eurostat & to verify the correctness of a designed tools
- Simple variables with uncomplicated filters => coding like 1:1
- NEEDCARE - a variable with a few groups of target respondents included in its filter – coded using one question that was placed in different places in the questionnaire
- A complex variable – AGEPENSO from 2023 module on pension and labour market participation

# Transformation table from national variables to the EU ones - NEEDCARE

Variable identifier	Variable name	Code	Labels	Filter	Polish LFS – question numbers
NEEDCARE	Main reason why care for children or incapacitated relatives limits labour market participation	1	Relevant care services not available	SEEKREAS = 4 OR WANTREAS = 3 OR FTPTREAS = 3 OR AVAIREAS = 3 Persons whose reason for not searching for employment or for not wanting to work or for working part-time or for not being available to start work is having care responsibilities	S6A=1 or S6B=1 or S6C=1 or P13A=1 or P13B=1
		2	Relevant care services not affordable		S6A=2 or S6B=2 or S6C=2 or P13A=2 or P13B=2
		3	Want to provide care themselves		S6A=3 or S6B=3 or S6C=3 or P13A=3 or P13B=3
		4	Other factors were decisive		S6A=4 or S6B=4 or S6C=4 or P13A=4 or P13B=4
		Blank	Not stated		lack
		9	Not applicable		SEEKREAS ≠ 4 and WANTREAS ≠ 3 and FTPTREAS ≠ 3 and AVAIREAS ≠ 3



# Transformation table from national variables to the EU ones - AGEPENSO from 2023 module

Variable identifier	Variable name	Code	Labels	Filter	Polish LFS – question numbers
AGEPENSO	Age at which the person started receiving an old age pension	0-120	Age in completed years (3 digits)	PENSTYP1=1-7 Persons receiving an old age pension	1) If MP1=1 and MP2≠99 then AGEPENSO=MP2 2) If MP1≠1 and MP4=1 and MP6=1 and MP5≠99 and MP7≠99 then AGEPENSO=minimum (MP5, MP7) 3) If MP1≠1 and MP4=1 and MP6=1 and MP5≠99 and MP7=99 then AGEPENSO=MP5 4) If MP1≠1 and MP4=1 and MP6=1 and MP5=99 and MP7≠99 then AGEPENSO=MP7 5) If MP1≠1 and MP4=1 and MP6≠1 and MP5≠99 then AGEPENSO=MP5 6) If MP1≠1 and MP4≠1 and MP6=1 and MP7≠99 then AGEPENSO=MP7
		Blank	Not stated		(PENSTYP1=1-4 and MP2=99) or (PENSTYP1=5 and MP5=99 and MP7=99) or (PENSTYP1=6 and MP5=99) or (PENSTYP1=7 and MP7=99)
		999	Not applicable		PENSTYP1=8, blank, 9

# Concluding remarks

- Development of a national LFS questionnaire is not an easy task
- The better cohesion the greater harmonisation at the EU level
- All additional materials elaborated by Eurostat are very useful
- Easy communication between NSI and Eurostat needed and helpful
- Carrying out pilot tests of the introduced changes as well as preparation of a transition key to the EU LFS variables is a good practice
- Proper cooperation of different submodules (blocks of questions) is extremely important
- Complex filters for variables should be avoided as much as possible

**Thank you for your attention**

[H.Strzelecka@stat.gov.pl](mailto:H.Strzelecka@stat.gov.pl)