Main Characteristics of the Polish Labour Force Survey: Sampling Scheme, Sample Allocation, Rotation Pattern and Calculation of Weights

Department of Methodology, Standards and Registers

Central Statistical Office, Warsaw 2017

General Assumptions

- The Labour Force Survey is a probability sample survey. This allows generalization of the results over the whole population. The survey covers all persons at the age 15 years and more, living in the sampled dwellings.
- LFS is carried out as a continuous survey. This means that in each of 13 weeks in a quarter interviewers visit a specified number (currently it is 4260) of randomly sampled dwellings and collect data concerning economic activity during a preceding week.
- Weekly samples result from a random distribution of a quarterly sample into 13 parts. The quarterly sample currently amounts to 55380 dwellings. It was constructed in such a way that every one of 13 weekly samples is not only the same size but has also the same structure.

Sample Rotation Pattern

- The sample for each quarter consists of four independently drawn elementary samples (e-samples),
- Partial rotation of e-samples is carried out in every quarter: in a given quarter there are two e-samples surveyed in the previous quarter, one e-sample introduced into the survey for the first time and one e-sample which was not surveyed in the previous quarter and was introduced into the survey exactly a year before;
- As a result of this rotation system each sample is employed according to the 2-(2)-2 rule: two quarters in the survey, two quarters break, again two quarters in the survey and then out.

ROTATION PATTERN 1999 – 2004

(2–(2)–2)

Elementary Sample Number	YEARS and QUARTERS																						
	1999				2000				2001				2002				2003				2004		
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1				x																				
2				x	x																			
3				–	x	x																		
4				–	–	x	x																	
5				x	–	–	x	x																
6				x	x	–	–	x	x															
7					x	x	–	–	x	x														
8						x	x	–	–	x	x													
9							x	x	–	–	x	x												
10								x	x	–	–	x	x											
11									x	x	–	–	x	x										
12										x	x	–	–	x	x									
13											x	x	–	–	x	x								
14												x	x	–	–	x	x							
15													x	x	–	–	x	x						
16														x	x	–	–	x	x					
17															x	x	–	–	x	x				
18																x	x	–	–	x	x			
19																	x	x	–	–	x	x		
20																		x	x	–	–	x	x	
21																			x	x	–	–	x	x
22																				x	x	–	–	x
23																					x	x	–	–
24																						x	x	–
25																							x	x
26																								x

Sampling scheme

- Sampling for the LFS follows the stratified two-stage household sampling. The Primary Sampling Units (PSUs) subject to the first stage selection, are census units called census clusters in towns, while in rural areas they are enumeration districts (probability proportional to size (i.e. number of dwellings) sampling is applied: Hartley–Rao method).
- The second stage sampling units are dwellings (simple random sampling is applied).
- Division to weekly samples is performed at the Primary Sampling Units level.

Sampling of Primary Sampling Units (PSUs)

- The primary sampling units (PSUs) are sampled with the appliance of the so-called stratification. Division into strata is based on voivodships (i.e. 16 NUTS2 regions).
- Within voivodships (16 NUTS2 regions) there were additionally defined 3 to 7 strata according to the size of a place (town or village).
- The 1-2 largest (main) voivodship towns were distinguished as separate strata.
- The other strata within voivodships were created depending on the size of a place; rural areas were included into the smallest ones.

Sample allocation

- The numbers of PSUs and dwellings sampled from individual voivodships (16 NUTS2 regions) were determined in order to:
- combine organizational requirements
- with a need to obtain representative generalisation results for voivodship towns, as well as for voivodship areas outside of voivodship towns in 16 NUTS2 regions (voivodships).

Calculation of weights – *primary weights*

- The estimation process consists in defining the appropriate generalizing factors, further referred to as *weights*. This is achieved in three steps. The first step provides *primary weights* which basically are the reciprocals of selection probabilities for ultimate sampling units (i.e. dwellings):

$$w_h = M_h / m_h \quad , \qquad h = 1, 2, \ldots, H \quad ,$$

where

M_h - is the number of all dwellings in the *h*-th stratum,

m_h - is the number of dwellings sampled from the *h*-th stratum,

- Such primary weights are calculated for each of 4 elementary samples which together form a quarterly sample, and the quarterly sample weights are obtained as

$$w'_h = w_h / 4$$

Calculation of weights – second step

- Then, the so-called interview rates ***R*** are calculated by the formula ***R=(K-N)/K***, where ***K*** is the number of interviewed dwellings estimated using *primary weights* and ***N*** is the estimate of the number of dwellings that were qualified for the survey but were not interviewed regardless of the reasons
- The *secondary weights* are calculated in the next step by dividing the *primary weights* by ***R***, where ***R*** rate depends on a voivodship (NUTS2 region) and the category of a place of residence of a given dwelling (6 - rural area or one of the five town classes: 1- Warsaw, 2 - towns with 500 thousand up to 1 million inhabitants, 3 - towns with 100 thousand up to 500 thousand inhabitants, 4 - towns with 20 thousand up to 100 thousand inhabitants, 5 - other towns).

Calculation of weights – second step

- Hence, the *secondary weights* are calculated as follows

$$w_{hk} = w_h / R_k \quad , \quad h = 1,2,\ldots, H \quad , \quad k = 1,2,\ldots, 6 \quad ,$$

where the rate R_k depends on voivodship (NUTS2 region) and category of place of residence.

- Final weights for the results concerning population are calculated in the third step.
- The purpose of the final step is the adjustment of the LFS results to the current demographic estimates.
- It is obtained by calculation of the so-called modifiers for each of 48 categories defined by the place of residence (urban/rural), sex and 12 age groups, separately for each voivodship (NUTS2 region).

Calculation of weights – final step

- The above mentioned modifiers are calculated as :

$$C_l = P_l / \hat{P}_l \quad , \qquad l = 1, 2, \ldots, 48 \quad ,$$

separately for each voivodship and each of 48 categories defined by place of residence (urban/rural), sex, and 12 age groups (15-17 years, 18-19 years, 20-24 years, 25-29 years, 30-34 years, 35-39 years, 40-44 years, 45-49 years, 50-54 years, 55-59 years, 60-64 years, 65 and more years), where P_l is the number of persons in a given category according to demographic estimates, $\hat{P}_l$ is the analogous estimate based on quarterly LFS data and *secondary weights*.

- The final weights result from multiplication of the *secondary weights* by adequate modifiers :

$$w_{hkl} = w_{hk} * C_l = w_h * C_l / R_k$$