Labour force survey (LFS) - sampling design, sample size and sampling errors

Statistics Explained

This article focuses on the significance of changes in employment and unemployment rates between two consecutive years, using the European Union Labour Force Survey (EU-LFS). While the first part of the article presents the sample size and the main characteristics of the sample designs of the EU-LFS participating countries, the second part is dedicated to the obtained precision and variance estimates of net changes for the employment and unemployment rates.

Who decides on the sampling methodology ?

The European Union Labour Force Survey (EU-LFS) is a large sample survey of people living in private households, which is conducted in 35 countries: the EU Member States, the United Kingdom, three EFTA countries (Iceland, Norway, Switzerland) and four Candidates Countries (Montenegro, North Macedonia, Serbia and Turkey).

The National Statistical Institutes (NSIs) of the participating countries are responsible for: designing the sample, developing the national questionnaires, conducting interviews, and sending results to the Commission (Eurostat) in accordance with a common coding scheme set by Commission Regulation (EC) No 377/20083. On the other hand, Eurostat is in charge of: monitoring the implementation of Regulation (EC) No 577/98, providing assistance to NSIs, promoting harmonised concepts and methods, and disseminating comparable national and European labour market statistics.

Each NSI is responsible for the sample design and applies the design which fits the most its needs and quality criteria, according to national specificities.

Which sample design is used in each country ?

First, distinction can be made between probability and non-probability sampling. Non-probability sampling uses a subjective method of selecting units from a population. In order to make inferences about the population, one must make the (often false) assumption that the sample is representative. On the other hand, probability sampling is based on three basic principles that make up the statistical framework. First, it is based on randomization, i.e. the units in the sample are selected at random. Second, all survey population units have a known positive probability of being selected in the sample, and third, we can calculate those probabilities, which are then used to calculate estimates along with estimates of the sampling error. The ability to make reliable inferences about the entire population and to quantify the error in the estimates makes probability sampling the best choice for most statistical programs.

In all 35 participating countries, the EU-LFS is based on a probability sampling (random sampling).

Second, the sampling can be stratified or not. In stratified sampling, the population is partitioned into non-overlapping groups, called strata, and a sample is selected within each stratum. Sampling units are independently sampled from each stratum. Stratification is used to ensure that the sample represents different groups in the population, and a stratum aims to represent a homogenous group of the population.

As shown in Table 1, which presents the sample design used in each country for the 2018 exercise, all countries

except Lithuania, Luxembourg, Malta and Iceland use a stratified random sampling method for the LFS. The stratification is mainly done on geographi-
cal areas (NUTS 2 or NUTS 3 or NUTS 4 levels). The degree of urbanisation is also a common stratification variable.

Countries' sample design

Sample Design	Country
With stratification	
Stratified one-stage cluster sampling	DE, DK, EE, CY, AT, SI, FI, SE, UK and NO
Stratified two-stage cluster sampling	BE, BG, CZ, IE, EL, ES, FR, HR, IT, LV, HU, NL, PT, PL, RO, SK, CH, ME, MK, RS and TR
Without stratification	
Simple Random Sampling	LT, LU and IS
Systematic Random sampling	MT

Source: 2018 National Quality reports

eurostat ◼

Table 1: Countries' sampling design

In addition, cluster sampling (also called clustering) can be used or not. With cluster sampling, the population is split into groups or clusters, one or a few clusters are selected at random, and a census is performed within each of them (which means that all units in the selected clusters are included in the sample). Cluster sampling is used to make sampling more practical or affordable. Multistage sampling (also known as multistage cluster sampling) is a more complex form of cluster sampling which contains two or more stages in sample selection, using smaller and smaller clusters at each stage.

On the other hand, the simple random sampling is the sampling method where each unit has the same probability to be included in the sample (no stratification and no clustering). The systematic random sampling corresponds to the situation where the sample is selected from an ordered sampling frame.

Table 1 shows that the most common sample design for the EU-LFS is the stratified two-stage cluster sampling (21 countries). The stratified one-stage cluster sampling is also frequently used (10 countries). On the contrary, the systematic random sampling (without stratification) is only implemented in Malta.

The Eurostat publication Labour Force Survey in the EU, candidate and EFTA countries MAIN CHARACTERISTICS OF NATIONAL SURVEYS, 2018 — 2019 edition provides more information on countries' sampling designs.

What is the sample size in each country ?

The achieved sample size is the number of observed sampling units successfully sampled. For the 2018 EU-LFS, the achieved quarterly sample (average computed over the 4 quarters of the year) is 1.743 million individuals for all participating countries (EU-28: 1.508 million), of which 1.333 million are in the age group of 15–74 years (EU-28: 1.152 million). The achieved sample in the EU-LFS is thus approximately 0.29 % of the total population. The achieved sample size in each country participating in the EU-LFS in 2018 is presented in Table 2.

Achieved sample and sampling rate in the EU-LFS by country 2018

	Overall sampling rate per quarter (%)	Achieved sample (average quarter 2018)	
		Households	Persons 15-74
Belgium	0.14	15 700	28 800
Bulgaria	0.65	14 500	25 500
Czechia	0.60	23 800	39 100
Denmark	0.73	-	22 000
Germany	0.25	92 600	140 200
Estonia	0.75	3 100	5 500
Ireland	1.70	14 200	25 300
Greece	0.89	27 000	43 600
Spain	0.40	64 900	121 000
France	0.25	58 100	93 200
Croatia	0.48	3 600	7 100
Italy	0.28	64 600	104 500
Cyprus	1.40	3 800	7 500
Latvia	0.96	4 800	7 300
Lithuania	0.90	6 700	12 200
Luxembourg	1.60	4 700	6 000
Hungary	0.92	23 200	40 600
Malta	1.70	2 200	4 600
Netherlands	0.50	43 100	83 500
Austria	0.60	20 100	33 400
Poland	0.40	28 500	53 700
Portugal	0.60	15 900	29 100
Romania	0.38	23 500	46 500
Slovenia	1.00	6 200	12 400
Slovakia	0.60	8 700	17 400
Finland	0.88	-	28 000
Sweden	1.20	-	47 300
United Kingdom	0.33	37 300	66 800
Iceland	1.70	-	3 200
Norway	0.60	13 600	20 600
Switzerland	0.54	-	27 900
Montenegro	1.54	2 100	4 900
North Macedonia	0.90	4 200	10 500
Serbia	0.77	12 300	25 500
Turkey	0.22	37 200	88 400

Source: 2018 National Quality reports

eurostat

Table 2: Achieved sample and sampling rate in the EU-LFS by country, 2018

The number of households and persons surveyed is generally proportional to the size of the population, i.e. bigger countries have bigger sample size. For the 2018 EU-LFS, the largest achieved sample can be found in Germany (140 200 persons aged 15-74), followed by Spain (121 000 persons aged 15-74) and Italy (104 500 persons aged 15-74). All other countries have an achieved sample smaller than 100 000 persons. However, to have a good indication of the representativeness of the sample, the sample size must be combined with the sampling rate. It also allows to make comparison between countries. The sampling rate is the ratio between the sample size and the population size. Malta, Ireland and Iceland have the highest sampling rate (all three 1.70 % per quarter), followed by Luxembourg (1.60 %). Three of these four countries are the ones with the smallest achieved sample size (3 200 persons for Iceland, 4 600 persons for Malta and 6 000 persons for Luxembourg), also because of their smallest population size. For the 2018 EU-LFS, most countries have a sampling rate lower than 1 %.

Which changes are significant ?

The use of a probability sample allows making reliable inferences about the entire population and quantifying precision (which is the degree of closeness between the estimation and the true value) of estimates based on variance estimation.

The EU-LFS, like all surveys, is based upon a sample of the population and this sample is used to make estimates that reflect the whole population. Estimations from the EU-LFS are therefore subject to the usual types of errors associated with sampling techniques, including sampling errors. This type of errors only affects sample surveys (like the EU-LFS) and arise because not all units of the target population are surveyed. Sampling errors can be quantified with the variance.

Eurostat has computed variance estimates of net changes between two consecutive years for 23 labour market indicators. The objective is to measure the significance of year to year changes for this set of indicators. The change in the annual average (average on the 4 quarters of the year) compares the average of one year with the average of the previous year. This evolution can be significant or can occur by statistical accident. For this exercise, the evolution is considered as significant when the confidence interval of the difference between the two annual averages does not contain zero, or in others words, when there is statistically 95 % of chance that the difference between the two annual averages is not equal to zero.

These variance estimates have been either produced by the National Statistical Institutes (NSIs) willing to do so or by Eurostat for the other countries. Eurostat has developed a method to estimate the variance for annual net changes, taking into account countries sampling schemes and overlaps between quarters and years, based on Taylor linearisation. Namely, information on the stratum and primary sampling unit ("first-stage" cluster) of each sampling unit is used in the computation of the variance estimates, as well as information on the final (or ultimate) sampling unit which is the household, the dwelling, the address or the person, depending on the country.

Tables 3, 4 and 5 present results on the significance of the net change for a couple of selected indicators. The significance 'yes' or 'no' means that the change between 2017 and 2018 is statistically significant or not (at a given level of confidence).

As shown in Table 3, the employment rate of persons aged 20-64 increased between 2017 and 2018 in all EU Member States. Iceland is the only EU-LFS participating country recording a decrease. The largest increase occurred in Cyprus, where the employment rate of persons aged 20-64 increased by 3.1 percentage point (p.p.). Results show that this large increase is significant. Smaller increases can also have an impact as seen in the case of Germany (+ 0.7 p.p.) and the United Kingdom (+ 0.5 p.p.). On the other hand, employment rate significantly decreased in Iceland (- 1.1 p.p.). In fact, changes between 2017 and 2018 were not significant in only two EU Member States (Estonia and Luxembourg) and in one EFTA country (Switzerland). For these three countries, the difference between the 2017 and
the 2018 employment rate is lower than 1 p.p. and the recorded difference is due to statistical sampling random effect.

Net change significance of the employment rate 20-64, total, in %

	2017	2018	Difference	Significance	Method	Estimated by
Belgium	68.5	69.7	1.2	Yes	Taylor	Eurostat
Bulgaria	71.3	72.4	1.1	Yes	Taylor	Eurostat
Czechia	78.5	79.9	1.4	Yes	Taylor	Eurostat
Denmark	76.9	78.2	1.3	Yes	Taylor	Eurostat
Germany	79.2	79.9	0.7	Yes	Taylor	Eurostat
Estonia	78.7	79.5	0.8	No	Taylor	Eurostat
Ireland	73.0	74.1	1.1	Yes	Taylor	Eurostat
Greece	57.8	59.5	1.7	Yes	Taylor	Eurostat
Spain	65.6	67.0	1.4	Yes	Jackknife	NSI
France	71.0	71.8	0.8	Yes	Taylor	NSI
Croatia	63.6	65.2	1.6	Yes	Taylor	Eurostat
Italy	62.3	63.0	0.7	Yes	Taylor	NSI
Cyprus	70.8	73.9	3.1	Yes	Taylor	Eurostat
Latvia	74.8	76.8	2.0	Yes	Taylor	NSI
Lithuania	76.0	77.8	1.8	Yes	Taylor	Eurostat
Luxembourg	71.5	72.1	0.6	No	Taylor	Eurostat
Hungary	73.3	74.4	1.1	Yes	Bootstrap	NSI
Malta	73.0	75.0	2.0	Yes	Taylor	Eurostat
Netherlands	78.0	79.2	1.2	Yes	Taylor	Eurostat
Austria	75.4	76.2	0.8	Yes	Bootstrap	NSI
Poland	70.9	72.2	1.3	Yes	Taylor	Eurostat
Portugal	73.4	75.4	2.0	Yes	Taylor	Eurostat
Romania	68.8	69.9	1.1	Yes	Taylor	Eurostat
Slovenia	73.4	75.4	2.0	Yes	Taylor	Eurostat
Slovakia	71.1	72.4	1.3	Yes	Taylor	Eurostat
Finland	74.2	76.3	2.1	Yes	Taylor	Eurostat
Sweden	81.8	82.6	0.8	Yes	Greg Estimation - CLAN	NSI
United Kingdom	78.2	78.7	0.5	Yes	Taylor	Eurostat
Iceland	87.6	86.5	-1.1	Yes	Taylor	Eurostat
Norway	78.3	79.2	0.9	Yes	Linearization	NSI
Switzerland	82.1	82.5	0.4	No	Taylor	Eurostat

France: data corresponds to France metropolitan
Due to resampling indicators obtained using Bootstrap can be different from Eurobase

Source: Eurostat

eurostat ◼

Table 3: Net change significance of the Employment rate 20-64, total, 2018. (%)

Table 4 presents the significance of the net change between 2017 and 2018 for the employment rate of low skilled persons aged 20-64. The concerned samples are therefore smaller (sub-samples) and more subject to variation (larger variance). The employment rate of low skilled persons aged 20-64 increased in most countries, especially in Cyprus (+ 4.6 p.p.), in Lithuania (+ 2.8 p.p.) and in Croatia, Luxembourg and Malta (all three + 2.5 p.p.). Nevertheless, the increase is only significant in Cyprus and Malta. Only three countries recorded a decrease of their employment rate of low skilled persons: Slovakia (- 0.9 p.p.), Belgium (- 0.3 p.p.) and Iceland (- 0.9 p.p.). Moreover, the decrease is not significant in any of these countries. To summarize, changes are significant for 10 countries, where it always corresponds to an increase of the employment rate.

Net change significance of the employment rate of persons with low level of education, age group 20-64, in %

	2017	2018	Difference	Significance	Method	Estimated by
Belgium	45.9	45.6	-0.3	No	Taylor	Eurostat
Bulgaria	44.4	46.2	1.8	Yes	Taylor	Eurostat
Czechia	49.2	50.9	1.7	No	Taylor	Eurostat
Denmark	60.2	60.6	0.4	No	Taylor	Eurostat
Germany	59.6	60.7	1.1	Yes	Taylor	Eurostat
Estonia	65.7	65.7	0.0	No	Taylor	Eurostat
Ireland	50.5	51.5	1.0	No	Taylor	Eurostat
Greece	49.2	49.9	0.7	No	Taylor	Eurostat
Spain	54.7	56.6	1.9	Yes	Jackknife	NSI
France	52.6	53.1	0.5	No	Taylor	NSI
Croatia	34.7	37.2	2.5	No	Taylor	Eurostat
Italy	51.0	51.7	0.7	Yes	Taylor	NSI
Cyprus	57.0	61.6	4.6	Yes	Taylor	Eurostat
Latvia	56.7	57.5	0.8	No	Taylor	NSI
Lithuania	44.1	46.9	2.8	No	Taylor	Eurostat
Luxembourg	56.4	58.9	2.5	No	Taylor	Eurostat
Hungary	53.9	55.7	1.8	Yes	Bootstrap	NSI
Malta	59.2	61.7	2.5	Yes	Taylor	Eurostat
Netherlands	61.2	62.6	1.4	Yes	Taylor	Eurostat
Austria	53.9	55.6	1.7	No	Bootstrap	NSI
Poland	40.8	42.0	1.2	No	Taylor	Eurostat
Portugal	67.7	69.6	1.9	Yes	Taylor	Eurostat
Romania	54.7	55.2	0.5	No	Taylor	Eurostat
Slovenia	49.3	50.8	1.5	No	Taylor	Eurostat
Slovakia	37.3	36.4	-0.9	No	Taylor	Eurostat
Finland	51.3	52.9	1.6	No	Taylor	Eurostat
Sweden	61.8	62.0	0.2	No	Greg Estimation - CLAN	NSI
United Kingdom	63.6	65.1	1.5	Yes	Taylor	Eurostat
Iceland	77.9	77.0	-0.9	No	Taylor	Eurostat
Norway	59.9	60.5	0.6	No	Linearization	NSI
Switzerland	67.7	69.2	1.5	No	Taylor	Eurostat

France: data corresponds to France metropolitan
Due to resampling indicators obtained using Bootstrap can be different from Eurobase

Source: Eurostat

eurostat 🇪🇺

Table 4: Net change significance of the Employment rate of low skilled persons, age group 20-64, 2018. (%)

Table 5 provides the estimates and the significance of the net change for the unemployment rate of persons aged 15-74. The unemployment rate decreases between 2017 and 2018 in all countries, except in Luxembourg (+ 0.1 p.p.) and Iceland (0.0 p.p). The largest decrease is observed in Croatia and Cyprus (both - 2.7 p.p.). Observed decreases are significant for all countries, except for Estonia (- 0.4 p.p.), Malta (- 0.3 p.p.) and Switzerland (- 0.1 p.p.).

Net change significance of the unemployment rate, persons aged 15-74, percentage.

	2017	2018	Difference	Significance	Method	Estimated by
Belgium	7.1	6.0	-1.1	Yes	Taylor	Eurostat
Bulgaria	6.2	5.2	-1.0	Yes	Taylor	Eurostat
Czechia	2.9	2.2	-0.7	Yes	Taylor	Eurostat
Denmark	5.7	5.0	-0.7	Yes	Taylor	Eurostat
Germany	3.8	3.4	-0.4	Yes	Taylor	Eurostat
Estonia	5.8	5.4	-0.4	No	Taylor	Eurostat
Ireland	6.7	5.8	-0.9	Yes	Taylor	Eurostat
Greece	21.5	19.3	-2.2	Yes	Taylor	Eurostat
Spain	17.2	15.3	-1.9	Yes	Jackknife	NSI
France	9.1	8.7	-0.4	Yes	Taylor	NSI
Croatia	11.2	8.5	-2.7	Yes	Taylor	Eurostat
Italy	11.2	10.6	-0.6	Yes	Taylor	NSI
Cyprus	11.1	8.4	-2.7	Yes	Taylor	Eurostat
Latvia	8.7	7.4	-1.3	Yes	Taylor	NSI
Lithuania	7.1	6.2	-0.9	Yes	Taylor	Eurostat
Luxembourg	5.5	5.6	0.1	No	Taylor	Eurostat
Hungary	4.2	3.7	-0.5	Yes	Bootstrap	NSI
Malta	4.0	3.7	-0.3	No	Taylor	Eurostat
Netherlands	4.9	3.8	-1.0	Yes	Taylor	Eurostat
Austria	5.5	4.9	-0.6	Yes	Bootstrap	NSI
Poland	4.9	3.9	-1.0	Yes	Taylor	Eurostat
Portugal	9.0	7.1	-1.9	Yes	Taylor	Eurostat
Romania	4.9	4.2	-0.7	Yes	Taylor	Eurostat
Slovenia	6.6	5.1	-1.5	Yes	Taylor	Eurostat
Slovakia	8.1	6.5	-1.6	Yes	Taylor	Eurostat
Finland	8.6	7.4	-1.3	Yes	Taylor	Eurostat
Sweden	6.7	6.4	-0.3	Yes	Greg Estimation - CLAN	NSI
United Kingdom	4.4	4.0	-0.4	Yes	Taylor	Eurostat
Iceland	2.7	2.7	0.0	No	Taylor	Eurostat
Norway	4.2	3.8	-0.4	Yes	Linearization	NSI
Switzerland	4.8	4.7	-0.1	No	Taylor	Eurostat

FR data corresponds to France metropolitan
Due to resampling indicators obtained using Bootstrap can be different from Eurobase

Source: Eurostat

eurostat

Table 5: Net change significance of the Unemployment rate, persons aged 15-74, 2018. (%)

Source data for tables and graphs

- Labour force survey (LFS) - sampling design, sample size and sampling errors - Tables

- Net changes – Results for 23 indicators

Data sources

Source: The European Union Labour Force Survey (EU-LFS) is a large sample survey providing results for the population in private households in the EU, the United Kingdom, EFTA and the candidate countries. Each quarter around 1.8 million interviews are conducted throughout the participating countries to obtain statistical information for some 100 variables. The EU-LFS provides both quarterly and annual labour market statistics on employment and unemployment, as well as on people outside the labour force. It also collects multi-annual information from modules, and provides input for model-based monthly estimates of unemployment and unemployment rates. It is consequently an important source of information about the situation and trends in the EU labour market.

Coverage: The data for France cover the metropolitan territory (excluding overseas regions). Country codes: Belgium (BE), Bulgaria (BG), Czechia (CZ), Denmark (DK), Germany (DE), Estonia (EE), Ireland (IE), Greece (EL), Spain (ES), France (FR), Croatia (HR), Italy (IT), Cyprus (CY), Latvia (LV), Lithuania (LT), Luxembourg (LU), Hungary (HU), Malta (MT), the Netherlands (NL), Austria (AT), Poland (PL), Portugal (PT), Romania (RO), Slovenia (SI), Slovakia (SK), Finland (FI), Sweden (SE), the United Kingdom (UK), Iceland (IS), Norway (NO), Switzerland (CH), Montenegro (ME), North Macedonia (MK), Serbia (RS) and Turkey (TR).

Definitions: Sampling is a means of selecting a subset of units – a sample - from a target population for the purpose of collecting information. The sample information is used to draw inferences about the population as a

whole.

Other articles

- All articles on labour market
- European Union Labour Force Survey (EU-LFS)

Dedicated section

- Labour Market

Publications

- Labour Force Survey in the EU, candidate and EFTA countries — Main characteristics of national surveys, 2017 , 2019 edition
- Quality Report of the European Union Labour Force Survey 2018 , 2020 edition