

Methodological note

DATA COLLECTION FOR THE EU-LABOUR FORCE SURVEY IN THE CONTEXT OF THE COVID-19 CRISIS

EUROSTAT, DIRECTORATE F

UNIT F3 — LABOUR MARKET AND LIFELONG LEARNING

VERSION	DATE	CHANGES IN SECTION
1.0	8 APRIL 2020	
1.1	24 APRIL 2020	1 AND 4
1.2	30 JUNE 2020	1, 4 AND 5

1. Background

The spread of COVID-19 across Europe has led to the imposition of severe restrictions in the movement of people and the shut-down of economic activities. National Statistical Institutes (NSIs) have begun to close their premises, staff started to work from home, and field activities have been put to a halt in many countries. All this has negatively impacted the collection of household survey data starting as early as calendar week 5 in February 2020 in two countries, and affecting a majority of countries by calendar week 9 in March 2020. The situation is not expected to improve in the coming weeks or even months. This puts double pressure on the producers of EU-Labour Force Survey (LFS) data: data collection will be impacted by decreases in overall volume of attempted interviews, higher non-response and changes in the shares of interview modes, while at the same time the real impact of the crisis will lead to large shifts in the labour market. The explicit goal of Eurostat is to provide data that will allow users to evaluate these shifts, and compare data over time and between countries. This has implications on data production in the short and long term.

This paper discusses the following topics: Section 2 describes the most common problems associated with the disruption of regular LFS data collection due to the outbreak of COVID-19 based on the results of the survey on the topic. Section 3 discusses risks and options for the 2020 Quarter (Q) 1 (microdata) transmission, as well as production and transmission of monthly unemployment data from 2020 Month (M) 3 onwards. Section 4 gives a number of

recommendations for the data collection of 2020Q2 data. Section 5 presents the way forward agreed with the International Labour Organisation (ILO) on data publication and communication to users. The Eurostat LFS teams thanks the ILO Statistics Department for the useful comments to this paper and confirm their commitment to a common work to ensure high quality, meaningful and consistent data.

This paper will be continuously updated following new developments, and when requests for clarifications arise from countries. The current version (1.2) includes the following updates:

- Extension of the COVIDE variable filter to those who have lost their job (LAMAS WG agreement).
- Clarification on the coding of variable MODE when face-to-face interviewers switched to telephone interviews.
- Clarification on absent from work classification, voluntary vs. involuntary absences.
- Clarification on absent from work to take care of children due to the closure of schools.
- Update on new indicators

2. Preliminary analysis of survey results

At the time of writing, 31 countries, of which 25 Member States, had responded to the survey initialised by Eurostat on 18 March 2020. The results presented below are of a preliminary nature and do not give the full picture. A follow-up survey covering final, realised data collection for 2020Q1, and ongoing 2020Q2 data collection will be initiated in due time. At this point, only 5 countries (LU, NO, DK, SE, CH) reported no impact of COVID-19 measures on LFS data collection for 2020Q1, possibly due to the limited use of CAPI⁽¹⁾. Of 26 countries that did report an impact, 2 could not yet give any details. For the remaining 24, results are summarised below:

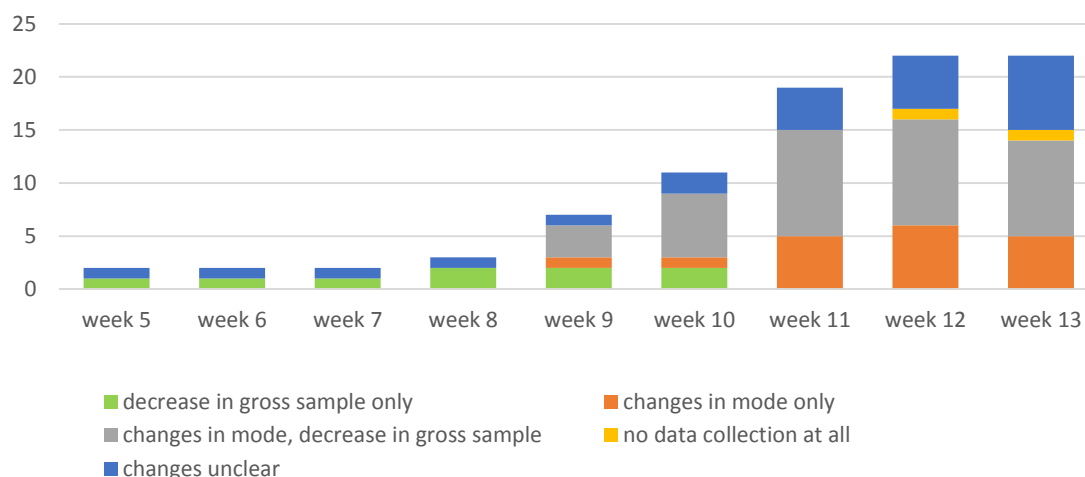


Figure 1: Number of countries reporting changes to gross sample and mode per calendar week 2020

⁽¹⁾ CAPI – Computer Assisted Personal Interviews.

Visibly, the main effect on data collection took place in calendar weeks 11 to 13 (results for weeks 14 to 16 still addressing 2020Q1 reference periods are incomplete and not yet reported). This stems to a very large degree (24 countries) from the ban of face to face interviews (f2f, CAPI mode).

Efforts have been made to collect more interviews by phone (CATI⁽²⁾). This is hampered in many countries by the fact that phone numbers are initially not known and collected with the first (f2f) interview. In many cases, households have been contacted by mail to collect the missing information. Further problems arise from closed call centers (7 countries), and call center staff having to work from home, or not showing up for work. Nevertheless, problems related to infrastructure breakdowns are rare. Overall, the effects on data for 2020Q1 which are likely to be observed are the following:

- Bias due to change in mode from CAPI to CATI: mode effect and possible self-selection bias in provision of phone numbers; unequal impact over the waves
- Loss of precision due to lower gross samples
- Loss of precision and potential bias due to lower response rates

Apart from changes in interview mode, several countries allow for a longer period to contact the sampled persons to increase the response rate. Some countries have introduced special questions (e.g. reason for absence from work), or plan to do so. Some countries plan to simplify or shorten their (national) surveys without impact for the EU-LFS. Only three countries of those who have answered have signalled major issues, mainly due to their reliance on PAPI⁽³⁾ mode.

At this point, not enough data is available nationally for countries to formulate strategies concerning data treatment to counter the likely effects. A number of countries have indicated that they cannot tell whether quarterly data for the first quarter 2020 can be delivered to Eurostat on time.

3. Data production and publication: monthly (2020M3 onwards) and quarterly (2020Q1 onwards) data

Quarterly LFS and Monthly Unemployment Rate (MUR) data are part of the domains prioritised by Eurostat in this period of COVID-19 outbreak. Priority is given to timeliness for the transmission of LFS quarterly microdata (as indicated in EU regulations) and for MUR data transmission in line with the agreed release calendar. As regards quality, a two-step approach can be used in which data submitted within the deadline can be revised when higher quality data become available.

In this context, the following points should be noted with regard to the production of monthly unemployment data:

Monthly series need to be benchmarked to quarterly LFS. In very exceptional cases and only if the quality of the monthly series is deemed higher than that of the underlying quarterly LFS, and after consultation and agreement with Eurostat, monthly series need not be benchmarked to quarterly LFS from 2020M3 onwards, until the quality issues in the quarterly LFS are resolved.

While transmission of all 8 input series (male, female; 15-24, 25-74; for unemployed and employed) is preferred, exceptionally estimates of total levels or rates are accepted, if this improves the quality of the estimates.

Seasonal adjustment should be done with care. Eurostat has opted to switch from partial

⁽²⁾ CATI – Computer Assisted Telephone Interview

⁽³⁾ PAPI – Paper Assisted Personal Interview

concurrent adjustment to controlled current adjustment for the time being. This is one strategy recommended by Eurostat. Other strategies are described in the note annexed to this document.

Depending on the overall situation in terms of country data availability and quality, decision will be taken by Eurostat to disseminate, MUR data according to the calendar release (data of March 2020 planned to be issued on 30 April 2020) for euro area and EU aggregates and for all countries that transmitted quality data on time.

Quarterly LFS Main Indicators as described in the draft manual annexed to this document are usually derived by Eurostat directly from the transmitted microdata, and this approach will be applied to 2020Q1 data as well. Exceptionally, if it is impossible for a country to transmit quarterly microdata on time, or if a country prefers direct modelling of the most important indicators to counter some of the likely effects on the data, it is possible to send estimates of these series. If this applies to your country, please get in contact with hannah.kiiver@ec.europa.eu to receive the templates for data transmission. While full sets of indicators including all breakdowns are preferred, exceptionally, partial transmissions can be accepted.

If the micro-data is not sent to Eurostat on time for processing and inclusion, or modelled time series are incomplete, Eurostat will rely on SARIMA models to nowcast the missing data for the LFS Main Indicators. These estimates will be included in the calculation of aggregates, but not published in the database, and revised accordingly once microdata have been sent.

Detailed quarterly tables will be derived from microdata as received. No aggregates will be derived and published unless data from all countries are available.

4. Data collection in 2020Q2 and further affected quarters

Several countries mentioned in the survey on COVID-19 measures that they have introduced or are planning to introduce additional questions, or further answer options to questions, to better capture the effect of the crisis on the labour market. One argument that has repeatedly been voiced is that of data users' expectations to see a strong and immediate rise in unemployment, and the dissonance of individuals self-identifying as unemployed but not being classified as such by ILO definitions; using the current variables and definitions, particularly the conditions of availability for work and active search for work seems inappropriate in the current context of COVID-19 restrictions, the argument goes. Eurostat has been in discussion with the ILO on the necessity and relevance of additional or changed questions and gives the following guidance:

NSIs should not change, delete or add to the set of questions used to derive the ILO labour market status. Comparability of data over time and between countries is only ensured if the variables used in the derivation of the ILO labour market status remain unchanged. Already now, there is a difference between ILO status and self-identification.⁽⁴⁾ The purpose of the LFS is not to mirror what data users might think should be going on in the labour market, but to show what is actually happening, according to fixed and objective standards. The essential concepts cannot be relaxed or changed: for example, unavailability because of closed child care and teaching institutions means that the concerned individuals are not available, and can thus not be counted as unemployed. A rise in involuntary exits from the labour market is foreseeable and unavoidable. The use of new, non-tested variables or definitions attempting to replace the

⁽⁴⁾ This can easily be derived by comparing the variables ILOSTAT with MAINSTAT: one quick data comparison found that during the last crisis in 2010Q1 for IE, only 70 % of those who reported themselves as being unemployed (in MAINSTAT) were, according to ILOSTAT, unemployed. Around 7% were employed, and 23% were outside the labour force.

current concepts to measure the impact of COVID-19 measure on the labour market could lead to unpredictable and undesirable results. However, this does not mean that additional information cannot and should not be collected; while it is important to keep the core questions, definitions and variables unchanged, additional questions can be posed to show the impact of COVID-19 measures.

The LFS already collects data with great detail. Instead of changing the core data collection itself, Eurostat and NSIs need to exploit what is already being collected, and complement rather than change existing variables;⁽⁵⁾ focus on indicators that give more or different details than the unemployment rate; and communicate the observed changes, or lack thereof with detailed and high-quality metadata. Clearly, this poses large challenges during a time, where the data collection itself has become more difficult. To improve the information Eurostat and NSIs can provide, the following recommendations are made:

- **Variables WSTATOR and NOWKREAS:** Eurostat asks NSIs to derive WSTATOR and NOWKREAS based on the same underlying questions and principles as has been done before the COVID-19 measures.
- **Variable COVIDE:** If NSIs have introduced, or plan to introduce additional questions, particularly to assess the extent of COVID-19 related measures on absences from work or on jobs lost, NSIs should ask and code them independently from NOWKREAS and WSTATOR for transmission to Eurostat. The new variable COVIDE is meant to be complementary to these variables.
From 2020Q3 it is recommended that the variable COVIDE covers also those who have recently lost their job (WSTATOR=5). The description of the variable COVIDE will become: absence from work or lost work due to COVID-19. The revised filter for the variable COVIDE should be:

WSTATOR=2,3 OR [WSTATOR=5 AND YEARPR=2020 AND 3<=MONTHPR<=12]

COVIDE		Absence from work or lost job due to COVID-19 measures	
Q	1	YES	WSTATOR=2,3 OR [WSTATOR=5 AND YEARPR=2020 AND 03<=MONTHPR<=12]
	2	NO	
	9	not applicable	
	blank	not stated	

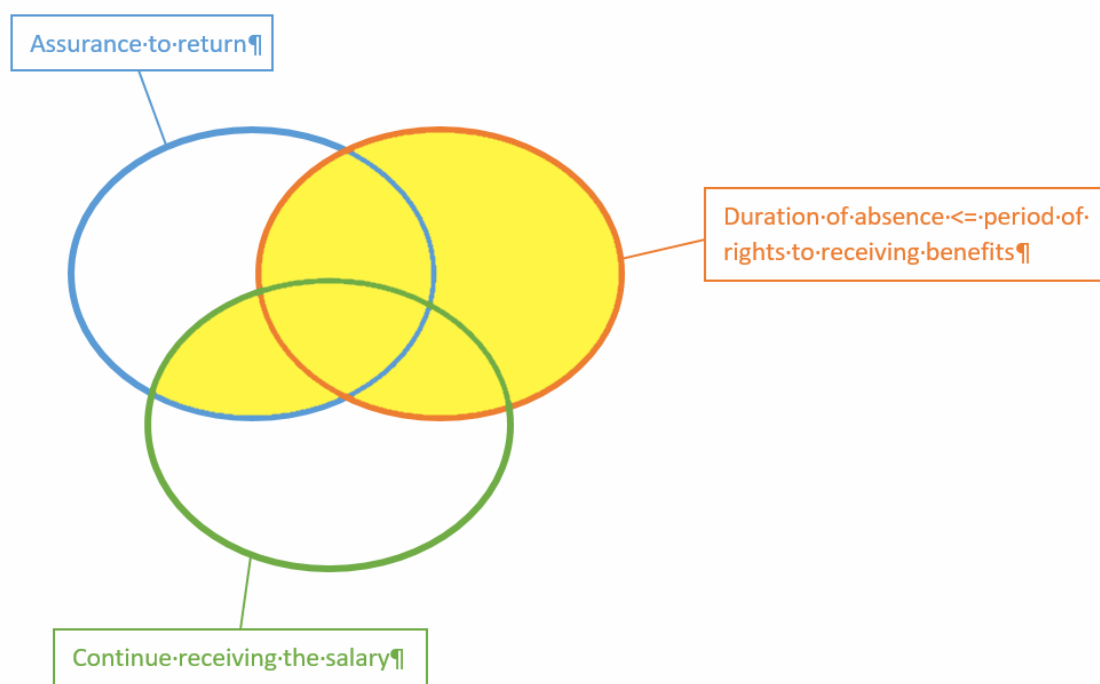
- **Variable COVIDA:** If NSIs have introduced, or plan to introduce additional questions, particularly to assess the extent of COVID-19 related measures on availability for work, NSIs should code them independently from AVAILBLE for transmission to Eurostat. The new variable COVIDA is meant to be complementary to this variable.

COVIDA		Non-availability due to COVID-19 measures	
Q	1	YES	AVAILBLE=2
	2	NO	
	9	not applicable	
	blank	not stated	

⁽⁵⁾ Eurostat has seen detailed questionnaires from Poland and Portugal; both are good examples of this strategy. A version of the questionnaire in English for the Polish LFS is annexed to this document.

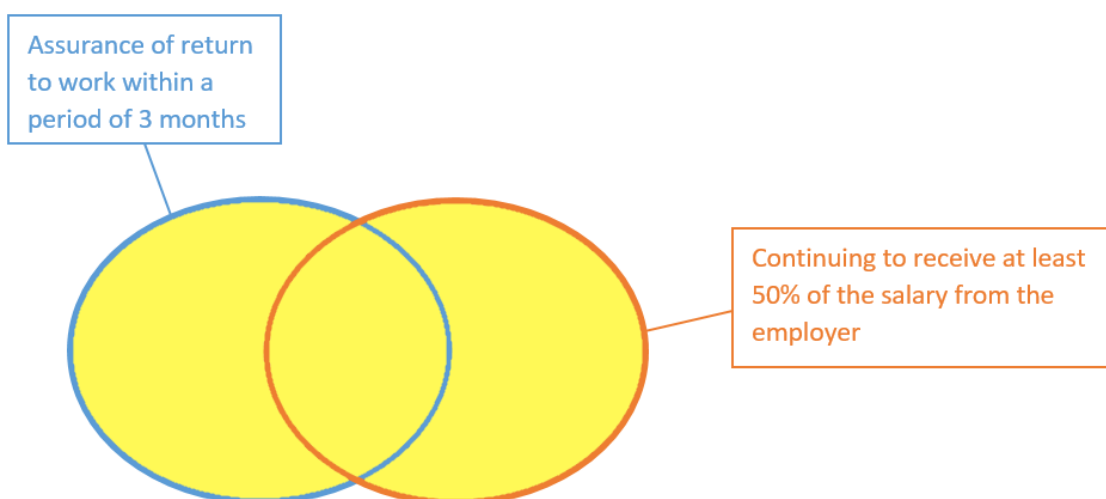
- **Variables AVAIREAS, SEEKREAS, LEAVREAS:** in most countries, these variables are collected for all respondents each quarter. In countries that use subsampling, Eurostat asks NSIs to consider moving these variables to the quarterly core.
- **Variable MAINSTAT:** Eurostat asks NSIs to start collecting and transmitting MAINSTAT quarterly, if not already done so.
- **Variable MODE:** Eurostat encourages countries to transmit the MODE for each interview using the definitions and codes in the new LFS implementing act 2019/2240 under the Integrated European Social Statistics Framework Regulation (IESS FR, regulation 2019/1700), already for 2020Q1 data. Some countries reported the situation that face-to-face interviews are banned, but also working in call centres is impossible. In such a case, a common solution was to switch face-to-face interviewers into telephone interviewers, sometimes using their usual computer based questionnaire designed for face-to-face interviews. In such cases the MODE variable should report a CATI interview (as perceived by the interviewees), but NSIs are invited to communicate such particular cases to Eurostat. This should be done by indicating the share of interviews initially assigned as CAPI but carried out as CATI (with details on CAPI or CATI questionnaire used) due to COVID-19 in the quarterly accuracy report in Section 3.1 – Comparability –Over time (Other comments), as percentage of total interviews of the quarter.
- **Explanatory notes for variable WSTATOR:** Eurostat asks NSIs to make sure that individuals doing any work from home, even if they are barred from their place of work, are categorised as WSTATOR=1. Working from home is equivalent to working at the work place.
Regarding those who are absent from work the current explanatory notes are complex and of only limited help in the current crisis situation. Eurostat proposes an interpretation framework to simplify the classification with more clear-cut rules. The main principle is that both WSTATOR = 2 and 5 cover voluntary absences, while WSTATOR = 3, layoff, covers involuntary absences for which the decision has not been taken by the worker:
 - **All voluntary absences** are either classified as WSTATOR=2 or as WSTATOR=5. These voluntary absences are dealt with in the explanatory notes for WSTATOR, code 2. For employees, the relevant text says:
In the case of employees, a person absent from work should be considered as employed if there is a formal attachment to the job, for example if at least one of the following criteria is fulfilled:
 - *the continued receipt of wage or salary, AND an assurance of a return to work (or an agreement as to the date of return) following the end of the contingency.*
 - *the elapsed duration of absence from the job which, wherever relevant, may be that duration for which workers can receive compensation benefits without obligation to accept other jobs.*

This can be described as it follows:



Those within the yellow area are to be classified WSTATOR=2, those outside the yellow area are to be classified WSTATOR=5.

- All **involuntary** absences decided by the employer are dealt with in the explanatory notes for WSTATOR, code 3, with a flow chart that can be similarly described as:



Those within the yellow area are to be classified WSTATOR=2 and NOWKREAS=01, those outside the yellow area are to be classified WSTATOR=3.

The following outlines some detailed examples relevant in the current situation. For individuals who are banned from work without wages due to government imposed COVID-19 measures, and these bans are linked to the duration of these measures, countries should check if WSTATOR=2 or 3 applies. In the EU-LFS, the rules applied to being classified as employed in the case of self-declared absences from work are stricter than those established in the 13th International Conference of Labour Statisticians resolution⁽⁶⁾. These deviations are justified by the local labour market peculiarities (of Europe) and it is important to stick to those stricter rules to assure comparability of data over time. The COVID-19 impact, and consequently the national COVID-19 measures, vary a lot among countries. In interpreting the rules, the specific national specific COVID-19 measures should be taken into account.

One important new issue is the treatment of the unknown duration of the absence due to specific COVID-19 government lock-down for employees: if employees who do not receive at least 50% of the salary from their employer, do not know when they can go back to work, Eurostat will continue to apply current rules where a “do not know” answer is equivalent to “longer than three months”. In this case, information on characteristics of the lost jobs is still collected in the EU-LFS by the variables covering the previous job (this does not apply in non-EU countries), and any loss in detail is outweighed by the gain in simplicity of the questionnaire and in early detection of lost jobs. For countries implementing the COVIDE variable, it will still be possible to identify the group of not employed because absent with unknown duration in a COVID-19 related absence. For self-employed, the rule will remain unchanged. Absent self-employed who expect to take up their previous activity will continue to be counted as self-employed.

Eurostat is aware that, in reaction to the COVID-19 crisis, ILO published a technical note, recommending to treat the “do not know” answer by either adding a new category among the existent ones, or adding a new question, or training the interviewers to detect the actual expectation to return to the same work once the measures will be over. Actually, the EU standard has always been slightly different from general ILO recommendations. For example, the 50% threshold is not in the ILO recommendations, and the two organisations have always agreed that the European specificities require a slightly differently tuned definition. The ILO definition also meets the need of developing countries that have less institutionalised labour markets. The different options have been discussed thoroughly between the two international organisations and ILO considers that the Eurostat proposal fits the international standards while respecting the local (i.e. European) labour market situation.

Reasons to favour the Eurostat proposal in Europe are numerous:

- Reduction of burden (generally the national EU-LFS questionnaires are longer than those of developing countries).
- Importance of time consistency of data to analyse the trends: it is thus important to keep the standards unchanged (even more so because they are going to change anyway soon with the entry into force of the IESS FR, and it would be meaningless to impose a double change within a short time period).
- A more formal labour market (like the European one) helps in simplifying rules: different contractual situations are more clearly perceived by the respondents

⁽⁶⁾ One example is the 50% salary threshold paid by the employer.

and require less contextualisation, simpler questions and sharp criteria can be applied.

- Early detection of loss of jobs (with a risk of overestimation of the fall in employment) is preferable to a conservative (with respect to changes in number of employed) criterion that risks to catch the fall in employment too late.
- The EU-LFS being quarterly, it is important to be fast in detecting all short-run changes. On the contrary, in several developing countries, the frequency of the LFS is not quarterly, and cross-sectional comparisons are more important than time-related analysis.

Eurostat hence suggests a simple rule, consistent with the current standard, and perfectly fit to quickly detect changes in the number of employed. Eurostat confirms that the differences with the ILO proposals are not the consequence of a disagreement between the two organisations; they have been discussed and agreed between Eurostat and ILO.

Another relevant example is the one of those absent from work to take care of their children due to the closure of schools, with unknown duration of absence (situation that differs from that of layoff). In this case, there are two possibilities:

1. [they continue receiving the wage or salary, AND have an assurance of a return to work], OR [they receive a compensation until they will be back at work and this is a proof that the duration is within the boundary stated in the Explanatory notes for WSTATOR, code 2, for employees]:

The correct classification is WSTATOR=2; NOWKREAS=09; COVIDE=1;

2. [they do not receive the salary, OR are not sure to go back to work] AND [the elapsed duration of the absence exceeds that duration for which workers can receive compensation benefits, for example they do not receive any benefits]:

The correct classification is WSTATOR=5; NOWKREAS=99; COVIDE=1 (COVIDE asked only starting 2020Q3, until Q2 COVIDE=9).

- **Some relevant additions to the questionnaire:** feasible changes to the questionnaire can include some wording changes that do not jeopardise the consistency with the underlying principles of the LFS as described in the LFS explanatory notes. This could be done with a twofold purpose: to adapt to the new situation with examples that are relevant, and to help respondents to answer showing that the survey is not just an abstract imposition but a useful tool anchored in reality. For example, when asking about at least one hour of work in the reference week it is possible to stress that the work can be executed "even from home" (ex. "have you done at least one hour of work even from home, during the reference week?"). Asking for the work search method, it can be specified that the search action can be done through internet, for example the question for the variable METHODE can explicitly quote "Inserted or answered advertisements in newspapers or journals or even on internet"; this can be specified also for METHODDC, METHODD, METHODDF and METHODDH. It is also possible to specify that availability (to start working within two weeks) can refer to working from home.

Concerning actual transmission, please submit the additional variables in the following positions:

240 – MODE (as from 2021) **and if possible, already include MODE in 2020Q1**

241 – COVIDE (as above)

242 – COVIDA (as above)

243 – LEAVREAS quarterly (for countries with yearly subsampling)

245 – SEEKREAS quarterly (for countries with yearly subsampling)

246 – AVAIREAS quarterly (for countries with yearly subsampling)

247 – MAINSTAT quarterly (for countries with yearly subsampling or not transmitting MAINSTAT yet)

It is important that all countries with yearly subsampling put the respective information in the columns above (i.e. columns 240 to 247). For technical reasons linked to Eurostat validation this is even needed for countries who survey some or all reason variables listed quarterly, as current Eurostat validation applies an all-or-nothing approach in case of subsampling.

For countries without yearly subsampling Eurostat will simply use the transmitted information in the normal reason variables. Those countries are asked to fill only MODE, COVIDE, COVIDA and MAINSTAT if not anyway transmitted already.

5. Publication of results

To guide data users, a methodological note based on the information collected from the survey summarised above, and regularly updated to reflect the situation, will be attached to the folder containing all LFS results in the data tree on Eurostat's website. It will contain the impact and timing of the COVID-19 measures in all countries, along with information taken from the open questions. It will also inform users if data has been estimated, by whom, and how.

As indicated in section 4, Eurostat, after discussion with the ILO, will analyse the impact of COVID-19 measures on the labour market by broadening the view from the unemployment rate as main measure to that of slack in the labour market in general, as well as focusing on experimental breakdowns of quarterly flow estimates to see where previously employed persons have transitioned, while distinguishing between unemployment and different types of inactivity.

Increases in unemployment will eventually be observed, but may depend on the end of lockdowns. Initially, large increases in inactivity are more likely, and the focus of analysis therefore needs to be on more detailed, meaningful breakdowns of that category. Eurostat has internally discussed and decided on its full communication strategy, including the indicators that will be put into the spotlight as well as the means of communication (press release, news item, Statistics explained page). The LABour MARKET Statistics Working Group (LAMAS WG) as well as the Directors of Social Statistics have been informed and supported the proposals.

The related new indicators are:

- Labour market slack, as endorsed by the LAMAS WG in June 2017;
- Changes in absences from work, with details by reason for not having worked at all although having a job;
- Development of total actual hours worked in the main job, using an index-like presentation to avoid direct comparison of levels among countries;
- Changes in recent job leavers and starters
- Weekly total absences
- Transitions out of employment and potentially also out of activity
- Depending on the additional variables received
 - Changes in mode
 - Inactivity broken down by LEAVREAS, AVAIREAS, SEEKREAS
 - Direct impact of COVID-19 measures on absences from work, availability.

As initial step, tables on labour market slack have been included in the online data base with the data release on 21 April, to be found in the LFS Main Indicator folder with table codes lfsi_sla_q and lfsi_sla_a. More new indicators will be published in the first half of July, while for other potential indicators, further analysis is necessary.

6. Next steps

Eurostat will continue evaluating the information already received from NSIs in more detail, and collect information on additional questions introduced nationally. This overview will be regularly shared with all Members of the LAMAS WG. Based on these inputs, further developments with regard to COVID-19 measures in the countries and feedback from NSIs on this paper, Eurostat will a) repeat the survey, trying to put the lowest possible additional burden on NSIs and b) continue to update all Members of the LAMAS WG on developments with regard to communication and analysis taken by Eurostat.