

► COVID-19: Guidance for labour statistics data collection

29 / April / 2020

Guidance to data producers to maintain labour force survey data collection

Key points

- COVID-19 is impacting the collection of labour data in various ways.
- Data compilers across many countries are attempting to deal with the impacts and maintain the flow of data. However, breaks in series or disruption in data availability cannot be ruled out.
- The typical impact is the need to suspend collection of data through face-to-face interviews.
- The most common response is to attempt to move all face-to-face interviewing to telephone interviewing, with varying challenges by country.
- One challenge faced by countries is the creation of a database of contact telephone numbers for respondents. Difference sources need to be reviewed and assessed for coverage.
- Countries with panel surveys or already using computer assisted data collection will have advantages over those with independent surveys or those using paper and pencil interviewing. Those with panels are considering re-using panels as a way to maintain response levels.
- Technological and logistical challenges will be faced to enable telephone interviewing, linked to the resources and systems available.
- Questionnaire content may also need to be reviewed. Changes in content or any other changes, for example in sampling, need to be planned and implemented carefully.
- In many cases interviewing may have to be partially or fully suspended for some period of time. Different options can be considered to handle this.
- Countries will also face challenges after collection has been completed, for example to assess and adjust for non-response, weighting the data, assessing quality and reporting transparently.



Contents

1 OVERVIEW	3
1.1 General points of note	4
2 SCENARIO 1: MOVING FROM FULLY FACE-TO-FACE INTERVIEWING TO TELEPHONE BASED INTERVIEWING	5
2.1 Challenge 1: Creating a database of contact details	6
2.1.1 Re-using old samples/panels	7
2.1.2 Assessing the result	8
2.2 Challenge 2: Implementing telephone interviewing	8
2.3 Assessing suitability of PAPI/CAPI questionnaire for telephone interviewing	9
2.4 Other options and challenges	10
3 SCENARIO 2: MULTI-MODE DATA COLLECTION WITH PARTIAL FACE-TO-FACE INTERVIEWING	11
3.1 Creating a database of contact details	12
3.2 Implementing the new interviewing approach	12
3.3 Reviewing questionnaire content	13
3.4 Other options and challenges	13
4 TEMPORARILY SUSPENDING INTERVIEWING	13
4.1 Choosing the appropriate reference period to use when the survey recommences	14
4.1.1 Retaining current content and reference periods	14
4.1.2 Changing reference period and/or using retrospective questions	15
4.2 Covering missed interviews at a later date in case of low response rates	15
4.3 Periodic survey	15
4.4 Timing of the disruption (during a cycle or covering a full cycle)	16
5 OTHER MODE-RELATED SCENARIOS	16
6 OTHER OPTIONS OR CHANGES - SAMPLING	17
6.1 Using the same frame but selecting a new (reduced) sample	17
6.1.1 Other sub-sampling approaches	17
6.2 Using a new frame	18
7 EX-POST (AFTER DATA COLLECTION) CHALLENGES OR ISSUES TO CONSIDER	19
7.1 Response rates and bias	19
7.2 Non-response adjustment and weighting systems	20
7.2.1 Non-response adjustments	21
7.2.2 Weighting/calibration	21
7.3 Summary on bias, non-response and weighting	22
8 ASSESSING DATA QUALITY AND REPORTING THE DATA	23
9 SUMMARY AND CONCLUSIONS	25

1 Overview

This note is part of a set of guidance and support documents being developed by the ILO Department of Statistics covering the impact of COVID-19 on labour market statistics.

The focus of this particular note is to provide guidance on approaches to maintain labour force survey (LFS) data collection. A separate note [“COVID-19 impact on the collection of labour market statistics”](#) has summarized the many impacts on data collection being experienced by countries as of the end of March 2020, mainly as a result of restrictions being imposed to slow the spread of COVID-19. While the impacts will inevitably evolve over time, several common patterns emerge highlighting the most substantial challenges being faced, along with the main responses being considered or implemented to date.

This latest note summarizes the key challenges and some of the options countries can consider to sustain their LFS operations to the extent possible. Given the variety of country contexts, appropriate solutions will ultimately need to be assessed at the country level and all challenges faced cannot be fully anticipated or prescribed for. In addition, there may be no good solution or perfect fit for the many challenges now faced, nonetheless options can be considered to reduce the impact on the statistics generated.

The note does not address the use of other data sources or additional independent surveys (for example rapid assessments covering labour and other issues), which may be valuable at the time but are an additional consideration.

While there are many different sets of circumstances across countries, the main challenges facing most countries can be broadly captured under two scenarios.

- [Scenario 1](#): Moving from fully face-to-face interviewing to telephone based interviewing. This scenario is something being assessed or implemented by many countries, some of whom have panel designs, others having independent samples for each survey instance. These countries rely either on Computer Assisted Personal Interviewing (CAPI) or Paper and Pencil Interviewing (PAPI)
- [Scenario 2](#): Multi-mode data collection is being used (e.g. CAPI plus Computer Assisted Telephone Interviewing (CATI)). In these cases face-to-face CAPI is generally being used for the first interview of a panel design, plus a limited number of households for subsequent interviews. The challenge is to move these interviews to telephone or Computer Assisted Web Interviewing (CAWI).

Either scenario 1 or 2 (or others) could involve temporary disruptions (short or long) in data collection, which in reality creates a number of variations within each scenario. Indeed, some length of temporary suspension of interviewing is being reported by the majority of countries who have provided information to the ILO. The issue of dealing with these suspensions is covered in [section 4](#).

The two scenarios are not exhaustive of all the circumstances in countries and some other scenarios will be more briefly discussed later in the note. However, the range of variations in country practices and circumstances make an exhaustive and prescriptive set of guidance impossible. The correct solution for any country will ultimately require a national assessment of various options in light of the national context, including data sources, resources etc.

Under each scenario there will be proposals to guide countries on how to assess the validity/robustness of different approaches, as this needs to be an important consideration regardless of scenario or approach to change. Quality assessment and evaluation will be discussed in further depth in the final part of the note.

A starting assumption for each scenario is that the preference is to maintain the current sample frame and design. This should be the preferred approach where possible as it is assumed that the original sample has been selected using a robust sampling frame and a sound sampling design. It is further assumed that other methodologies will be retained as they are, such as approaches to non-response adjustment or weighting.

Recognizing that these assumptions may not necessarily hold for all countries (e.g. it may be considered to change the sample design) other types of possible methodological changes are referenced later in the note. However, care will be

needed in introducing any changes beyond those necessary, as any changes can increase burden on the NSO and introduce additional discontinuity in series.

Another starting assumption is that countries are looking to implement changes quickly, in an attempt to maintain previously identified field collection periods, or as close to them as possible. This is a consideration countries will need to bear in mind when evaluating their options.

The structure of the remainder of this note is as follows:

- [Section 2](#): Scenario 1: Moving from fully face-to-face interviewing to telephone based interviewing
- [Section 3](#): Scenario 2: Multi-mode data collection with partial face-to-face interviewing
- [Section 4](#): Temporarily suspending interviewing
- [Section 5](#): Other mode related scenarios
- [Section 6](#): Other options or changes - sampling
- [Section 7](#): Ex-post (after data collection) challenges or issues to consider
- [Section 8](#): Assessing data quality and using the data
- [Section 9](#): Summary and conclusions

As a general comment, ideally, any changes are brought in through a substantial process of development and testing. Even in the current pressurized circumstances, the need for some form of testing before implementing changes, even on a small scale, needs to be reflected in plans in order to reduce difficulties during implementation.

1.1 General points of note

The range of challenges and responses reported by countries to attempt to maintain the operation of their LFS is very wide.

The severity of the challenge also varies heavily. Countries already fully using CATI/CAWI are generally facing minimal disruption to data collection (although potentially still some). By contrast, some countries are losing their capacity to collect data through existing methods entirely. The challenge facing these countries cannot be underestimated and it is feasible, if not likely, that data collection may need to be suspended, either briefly or for a longer period of time, and indeed this has already happened in many countries.

Furthermore, even with flexibility of approach and the best efforts of all involved, the changes to mode, or other methodologies that need to be made, can lead to breaks in series or bias in various ways. This is quite apart from the inevitable shifts in series due to the economic impacts of COVID-19. The process of distinguishing genuine shifts from breaks due to changes in methodology, response rates, etc, will be difficult.

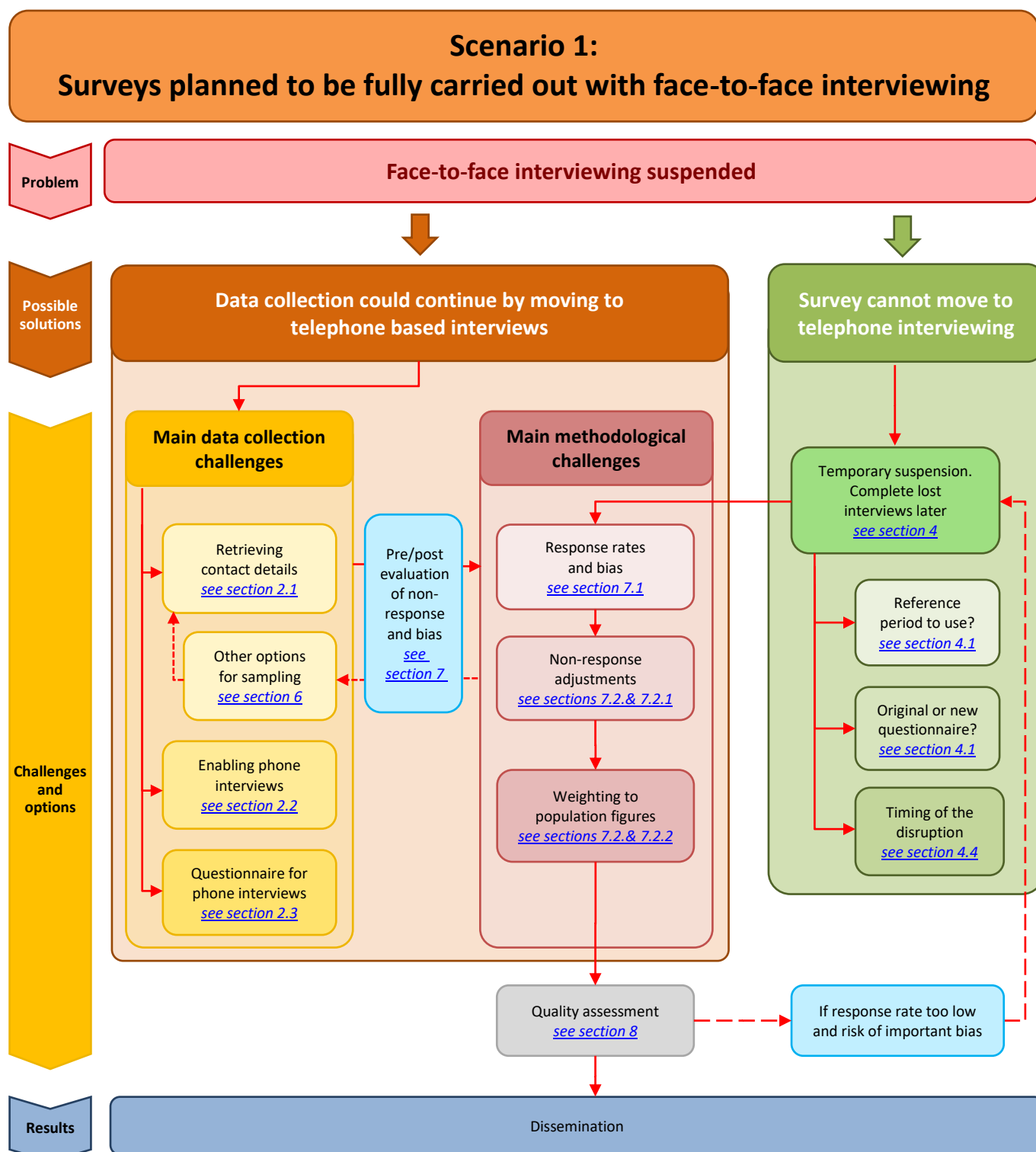
With all this in mind, it must be acknowledged that maintaining any continuity in the system of labour statistics at this time will be an important achievement, not to be taken for granted. It may ultimately be impossible to avoid breaks in series, or entire loss of periods of data. The guidance presented below should be understood a selection of options for countries to assess, and related considerations to attempt to minimize impacts. The options range in complexity to implement and their impact on the data.

As already noted this will mean that the appropriate solution for any country will ultimately require a consideration of the specific circumstances at the national level, meaning that a first step in any process, quite likely already completed in many countries, is an assessment of possibilities to narrow down to realistic options for further assessment/development.

In general the note refers to national statistical offices (NSOs) as for most countries they will be the institute with responsibility to carry out the LFS. However, in some countries other agencies, such as central banks or ministries of labour may be responsible for the LFS.

2 Scenario 1: Moving from fully face-to-face interviewing to telephone based interviewing

This is a scenario reported by a large number of countries, whereby restrictions on movement have caused all face-to-face interviews to be halted, and face-to-face interviewing was the only method in use (CAPI or PAPI). Under this scenario there can be different periodicities (quarterly continuous collection, quarterly periodic, annual continuous, annual periodic, etc.). Some of the countries in this situation use independent samples for each survey instance, while others are using a panel design. The challenges are mostly the same but those with a panel design can potentially benefit from it in different ways.



The guidance below covers both cases (independent samples or panels), noting differences where relevant.

This scenario may well be accompanied by a temporary suspension in data collection and more guidance on options to deal with temporary suspension of data collection (of different lengths) is presented in [section 4](#). Another possibility, is that countries may decide after assessing the options that a move to telephone interviewing is not feasible in the immediate term, for example if the necessary technology or systems are unavailable. In this case a temporary suspension (either partial or complete) will again be the outcome, with data collection recommencing whenever conditions allow. While this may need to be face-to-face in the immediate term, it nonetheless remains recommended to consider ways to begin the move to telephone interviewing as it is less disrupted in case of restrictions of movement, like now being experienced.

The main challenges to be faced under this scenario are the following: a) retrieving contact details for the final sampling units (FSUs – generally households); b) enabling capacity to carry out interviews by phone; and c) assessing the suitability of PAPI/CAPI questionnaire for telephone interviews.

2.1 Challenge 1: Creating a database of contact details

As already stated, we are assuming the desire is to maintain the current sample design and any selected sample, therefore the challenge is to get telephone contact details for the selected households in the sample. A related consideration for many countries is to re-use old samples/panels. This possibility is discussed further under [section 2.1.1](#). For the moment, we simply refer to the task of adding contact details to an already selected sample.

For countries with independent samples this will refer to the entire sample. For those with a panel survey this may only refer to newly introduced households plus some previously interviewed households for which contact details are not available – however, this assumes the contact details were captured during a previous interview.

Countries are approaching this challenge in various ways that can be broadly summarized as following the steps below:

- **Step 1: Assess information already available for the selected sample (within the NSO):** For countries with a panel design, contact details may already be available for many of the households which have previously been interviewed. For those with independent samples evidently this is less likely. The starting point will be to assess if there is any internally available information, for example from the sample frame, that can allow phone numbers to be retrieved, either landline or mobile. Even for panel designs it cannot necessarily be assumed that the contact details have been captured systematically. For example, they may be held by the interviewers but not entered on any database, requiring an exercise to capture this information in a more accessible manner to enable a central assessment of coverage. Any other potential sources of information, within or already accessible to the NSO, should also be considered, for example, if administrative data is accessible that may contain contact information.
- **Step 2: Assess available large scale sources of information for telephone numbers:** This can be public telephone registers, electoral registers etc. The possible sources will vary by country. Some countries are considering linking multiple sources to obtain one set of contact details that is as comprehensive as possible. An important issue with be the ability to legally access the data. This may not be possible if it refers to a privately held database containing customer information, or an administrative data source with legal restrictions on access.
- **Step 3: Link the contact information** to the sample file. An underlying assumption is that some identification information is available to allow some form of systematic/automated approach to matching (e.g. by name and elements of the address). This assumption may not hold in reality. This also assumes that some identification information is available with regard to the selected sample. This may not necessarily be the case, for example if the selected sample only identifies geographic areas and household listing is necessary as a first step of any field collection process (in other words the FSUs have not yet been identified).
- **Step 4: Manually add contact details** to the sample file. This can be done either using the data from step 1 or independent sources (for example web searches using name and address details). This may either supplement a systematic approach (step 2) or be the only approach if systematic linking is not possible. Some countries are using the existing face-to-face interviewers to complete this task for the sample of households they are assigned, typically

using web searches or perhaps using local networks to retrieve contact information.

An additional option being implemented by some countries is to use introductory letters, either sent by mail or delivered to households in the sample, describing the survey and requesting the household to provide contact details for a telephone interview. In these cases households may be given a number of options to make contact with the NSO such as returning a form by post, calling a central telephone line to provide details, calling the designated interviewer to either complete the interview or make an appointment, providing information by email or online etc. This approach can have the added benefit of promoting response by providing respondents with information about the survey, but this evidently has resource and cost implications. In countries with low levels of unique addresses or limited available contact information (e.g. area sampling was being used with household listing by interviewers as a first step of the field collection) this process will evidently be more difficult or impossible. At least one country has reported using face-to-face interviewers to deliver the letters but this may not be possible in many countries due to COVID-19 restrictions and the risks to interviewers.

Countries may combine the above options or steps in different ways, but a starting point will be the initial assessment of available sources for contact details including their comprehensiveness, and the possibility to link them to the sample file. This initial assessment may include a test based on a subset of the sample to evaluate the likely level of matching which can be achieved - e.g. take a random sub-sample of a number of households from the sample and assess how many matches can be achieved. If a stratified sample is being used, this should also be applied in the test to assess any differences in match levels by strata (for instance better coverage in rural or urban areas) and some distinction should be made between new households and those previously interviewed (if relevant). The scale of the test could be higher or lower depending on available resources but should be sufficient to draw reasonable conclusions on coverage and potential bias across any strata used. Based on the results of the test or the initial assessment a more informed decision can be taken on the feasibility of obtaining a high level of contact details to support the survey process.

2.1.1 Re-using old samples/panels

For several reasons, an option being seriously considered, or already pursued, by many countries is the re-use of old samples/panels. Some of the motivations for this include:

- Contact details may already be available for the sample.
- Even if phone numbers were not captured, the available information may make the sourcing of contact details more straightforward.
- The households are already familiar with the survey, potentially increasing co-operation levels versus those who would be interviewed for the first time.

This option may be feasible with independent samples with regular periodicity (e.g. quarterly), but will be a more likely option where panel designs are in use. The higher frequency the survey is the greater the chance that this approach can be viable.

Regarding the challenge of contact details, as already stated it may be the case that the details have already been captured for many of the households involved. However, this is unlikely to have been fully comprehensive, for example as some households do not have telephones (either landline or mobile) or the information was simply not captured. Therefore, the challenge of creating a comprehensive set of contact details will remain, and perhaps be almost as substantial as those trying to introduce new samples. Nonetheless, it is likely that some information will be available from the previous interview to support the process, for example name and address of respondents, which may not exist for the new sample. Depending on the level of coverage of existing contact details manual searches to supplement the available details may be sufficient to improve partial undercoverage, or the larger scale approaches described under [section 2.1](#) may be more suitable.

The manner of implementation of this option will depend on the current panel design, with various approaches currently in use across countries. The most common way to apply this approach will be to add an additional visit/interview to those households that were due to rotate out of the sample. This is done in place of introducing a new panel. Taking the example

of a country using a design which involves participation in 5 consecutive quarters, this would mean re-interviewing the households in the 5th wave an additional (6th) time, but not introducing a new group of wave 1 households. The full sample then could potentially remain fixed (a full panel) for the period of time needed to enable new panels to be introduced by whatever means.

A key concern with this approach will be panel attrition as the level of response will drop with each consecutive interview. The rate of attrition will also be likely to increase due to the move from face-to-face to telephone interviewing as some households may not have provided contact details, may not have a telephone, or may have a preference for face-to-face interviewing. Some alteration to weighting/calibration may be needed to adjust for any biases arising (discussed further in [section 7](#)). Nonetheless, this may be a preferable option to attempting to introduce a new panel of households for which contact information may be limited, or where initial co-operation may be more difficult to obtain without the benefit of a face-to-face interview.

2.1.2 Assessing the result

The guiding principle in this case will be the possibility to achieve a high level of coverage of contact details (expressed as a percentage of sampled households by strata). In addition to evaluating the overall percentage coverage, NSOs should assess any biases in coverage, for example some regions covered well, others poorly. If a stratified sample is being used then this could be seen by differences in level of coverage across the strata. Any important differences in levels of coverage need to be followed up to assess the possibility to achieve more even coverage, for example by manual searches just for selected strata.

No single threshold for viability can be suggested given the many varied approaches and possible outcomes. However, countries could evaluate likely impact on precision of estimates through a number of assumptions. Firstly, apply an expected success rate to the known level of contact (e.g. contact details available for 60 per cent of households and an assumed success rate of 80 per cent giving a level of completion of 48 per cent of the original selected sample). Secondly, based on the estimated level of completion (or expected number of completed interviews) calculate projected precision estimates on key indicators. Any biases in coverage also need to be considered in interpreting any such set of projected precision estimates. A conclusion may be that it remains possible to generate sufficiently unbiased and precise estimates for some indicators or levels of disaggregation, but not others, for example state level being ok, but geographic disaggregation being impossible.

Ultimately, the objective of this process must be the achievement of a relatively unbiased and representative sample to support the generation of estimates, at the national level as a minimum. Even if this is considered feasible, an ex-post evaluation of the results achieved will also be required to support transparent reporting of any results (discussed further later in the note).

Once all efforts have been made to minimize biases during the data collection stage, the remaining options to reduce the impact of biased samples will relate to non-response adjustment and weighting. These are discussed further in [section 7](#).

2.2 Challenge 2: Implementing telephone interviewing

Countries facing scenario 1 will have the added logistical and technological challenge of putting in place a telephone interviewing process and related systems. In general this involves enabling the interviewer to complete the interviews from home. Those countries already using a CAPI solution will have advantages over those using PAPI, but both will face challenges. Some points to note or options include:

- The interviewer will need to have access to a suitable space to complete interviews and have the necessary technology, for example a mobile phone (belonging to the interviewer or provided by the NSO), perhaps with a service that is prepaid by the NSO.
- Interviewers will need additional training and guidance to complete telephone interviews.
- Of particular note is the fact that it is no longer possible to use the initial face-to-face contact to secure co-operation with the survey. NSOs are concerned about the impact of this on non-response rates (in addition to the problem of

obtaining contact details). They are dealing with this challenge in various ways:

- To increase response some countries are using publicity (local or national) to advertise the survey in advance
 - As already mentioned a posted out or delivered introductory letter can be used to provide information about the survey and/or introduce the interviewer.
 - Countries are considering the introduction of incentives for participation. This option needs careful consideration of the type/level of incentive to avoid the introduction of bias, and can be very sensitive to the country or local context.
 - The ex-post treatment of non-response once data collection has been completed is discussed further in [section 7](#).
- Arrangements will need to be considered for data entry or transmission. If CAPI is already being used this will relate to ensuring secure internet access is available to facilitate transmission of data using the existing systems in place, but it will need to function from the interviewers home or other designated location. Any changes to existing systems or locations should be tested to avoid data loss or other complications.
- If PAPI was being used, choices will need to be made about how to submit the data. A data entry software may already have been developed and this could be made available to the interviewers, along with training on how to use it. Alternatively, it may be feasible to generate a data entry application (or even a CAPI/CATI solution) using software available such as CSPro. This should build on any existing capacities within the NSO, such as preferred data entry software and expertise. However, the necessary technology (laptops, PCs, tablets, internet connections, secure methods of data transmission) will need to be made available to the interviewers.
- If interviewers continue to use PAPI along with telephone interviews, the logistics will need to be put in place to enable the forms to be returned and entered. The NSO offices may or may not be accessible so any processes and systems may need to be remotely implemented. NSOs will need to assess the feasibility of this, e.g. relocating servers. Alternatively, the forms could be collected but stored for later entry, meaning a delay in publication but at least no loss of data/coverage. In this case, consideration is still required of appropriate storage to ensure confidentiality and security of the data collected.
- It may also be the case that capacity exists for other surveys undertaken by the organization that could be leveraged. For example, if telephone interviews were being implemented for an establishment survey then systems and/or interviewers (subject to appropriate training for interviewers and adaptation of systems) could be used for the LFS.

Assuming data collection can go ahead the attention will switch to ex-post assessment of the achieved sample to assess bias. The overall objective of all the steps and processes must be to achieve an unbiased representative sample to the extent possible.

2.3 Assessing suitability of PAPI/CAPI questionnaire for telephone interviewing

PAPI or CAPI interviews are typically longer than would be recommended for a telephone interview causing several countries to consider a reduction in the survey content in order to maintain response levels. A separate note [“Monitoring labour markets amid lockdowns to contain the COVID-19 virus: Essential labour force survey content and treatment of special groups”](#) has been published regarding essential LFS content, highlighting the need to retain, as a minimum, those parts of the questionnaire required to generate key labour market indicators and main disaggregations, as well as guidance on treatment of special cases that may become more prevalent during this period, such as extended absences from employment of unknown duration. Other adjustments to content, such as the introduction to the survey provided at the start, may also be required, especially if there has been no prior introductory contact.

Countries may also wish to consider the addition of survey content, in particular to more directly assess the impact of COVID-19 on respondents' working situation. This will be more likely for those using multi-mode data collection and is discussed further in [section 3.3](#).

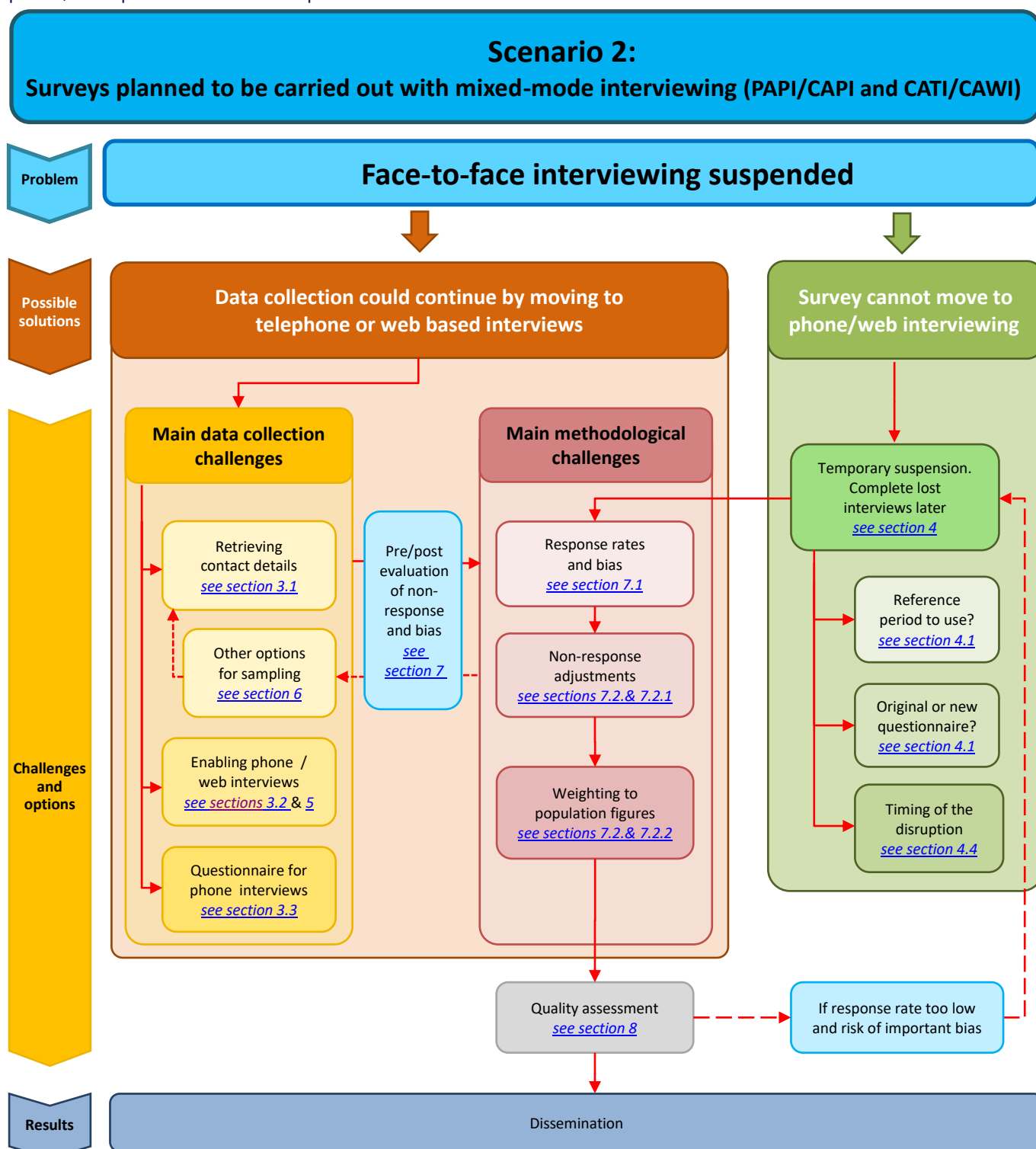
When considering any alterations to content, NSOs must carefully assess the impact this will have on the data generated, for example loss of indicators. Any changes should be subject to some form of testing, even if on a small scale to avoid operational or response difficulties when implemented. All changes must be documented and communicated to users when the data is being disseminated.

2.4 Other options and challenges

There will be a variety of other options and challenges facing countries under this scenario, in common with other scenarios. One challenge of note will be handling any temporary break in data collection ([see section 4](#)). Other options can include changing the sample and this is discussed in [section 6](#). Other challenges will relate to issues such as handling non-response and weighting, as discussed in [section 7](#). Another general challenge or issue, not discussed in detail in this report, is the possibility of a mode effect, whereby respondents may respond differently by telephone, even keeping all other elements constant. There may be some possibility to assess this through testing but most countries are unlikely to have this opportunity in the current environment. As a minimum the possibility of this effect needs to be recognized.

3 Scenario 2: Multi-mode data collection with partial face-to-face interviewing

This scenario covers many countries in Europe and selected countries in other parts of the world. In almost all cases there will be a panel design. The most typical scenario will be a face-to-face interview for the first interview, with CATI (and possibly CAWI) for subsequent interviews. Some face-to-face interviewing will typically continue to take place for the subsequent interviews, but limited to those for whom CATI or CAWI are not possible, for example households with no phone, or respondents who state a preference for a face-to-face interview.



As with those countries under scenario 1, the immediate response for the majority of these countries is to move any remaining face-to-face interviewing to one of the other modes, mostly CATI.

Essentially the range of challenges and options will broadly be similar to those outlined for scenario 1, but some specific issues still warrant attention.

3.1 Creating a database of contact details

Assuming a panel design is already in use the challenge will only relate to those respondents scheduled for face-to-face interviewing. Depending on the panel design, telephone coverage in the country, age profile of the population etc., this could still be a large proportion of all interviews. As such, while not as substantial as those without a panel design the challenge of obtaining contact details will remain significant.

The various options and considerations set out under scenario 1 remain relevant ([see section 2.1](#)), including:

- Assessing possible sources for contact details
- Linking data to the sample frame or selected sample (based on the identified sources)
- Manually adding contact details through searches
- Using introductory letters to introduce the survey and obtain contact details/arrange interviews
- Re-using previous panels in place of introducing new panels ([see section 2.1.1](#))

In this situation, like with scenario 1, one of the most important issues is assessing whether the sample of respondents actually interviewed is still representative of the full theoretical sample, and unbiased estimates can be produced. The issue of bias is discussed further in [section 7.1](#).

3.2 Implementing the new interviewing approach

For the majority of these cases the technology for computer assisted collection and transmission of data will almost certainly exist. Nonetheless, it is possible that arrangements need to be made to enable interviewing to take place from alternative locations. For some of these countries telephone interviewing may already be taking place from interviewer homes (prior to the disruption), while for others call centres or central facilities may be used.

- A general assumption is that the countries retain their existing approach to sampling of individuals within FSUs/households (other alternatives are discussed under [section 6](#)). In general this involves interviewing all members of the household above an assigned age that designates the working age population. To the extent possible, as with face-to-face interviewing, direct interviewing should be attempted for as many household members as possible and proxy used where necessary. Under the current circumstances an increase in proxy response rates may occur, and countries need to assess the possible impacts of this, particularly if the increase is very substantial. Any failure to achieve full response (either direct or proxy) among household members should be treated in line with existing practices.
- For those countries that are already implementing telephone interviews from the interviewer's home, the change should be relatively straightforward to implement but there can be an increase in the volume of interviews the interviewers must complete. This must be factored into planning.
- For countries already using CATI through central facilities/call centres it may be feasible to add the extra interviews to their schedule and use the existing facilities. However, in some cases those call centres are being closed due to COVID-19 restrictions forcing those operations to be relocated, typically to the homes of the call centre interviewers. This will create a range of logistical and technological challenges, such as ensuring necessary hardware and software can be provided, necessary data security and transfer to the centralized server can be maintained, a suitable environment for interviewing is available and data connectivity can be achieved.
 - An additional related case will occur where the closure of a call centre leads to a situation where the interviews need to be re-assigned to other interviewers because the CATI systems cannot be decentralized.

One country facing this situation is moving all interviews (both those scheduled for face-to-face and CATI) to be completed by existing field interviewers who usually collect information by CAPI. Those interviewers will now complete all interviews using existing CAPI software but making contact by telephone. This represents a large increase in volume of work for the interviewers and will likely require additional resources to be allocated (extra interviewers, technology etc.).

- As with scenario 1, capacity may exist within the organization to support the LFS operation, for example other surveys being completed by telephone. Several countries have identified priority data collection operations, with the LFS generally being one of the top priority activities. The significance of this is that resources will be re-allocated from other activities to sustain the LFS operation if needed.

3.3 Reviewing questionnaire content

Again, the other options identified under scenario 1 are generally relevant ([see section 2.3](#)) but may be slightly different for those with multi-mode operations. For example, some of these countries may indeed consider the reduction of survey content to reduce respondent burden. However, in many of the countries under scenario 2 the questionnaire has already been adapted to a telephone mode, and it is more common that they are considering addition of a limited number of questions specifically on the impact of COVID-19, or possibly even a parallel survey to more directly address COVID-19 related labour market impacts ([see one example from Australia](#)). Like reductions in content, additions to the LFS questionnaire, while potentially very valuable, need to be considered carefully so as not to introduce unintended impacts on existing measures or have a detrimental impact on response rates due to increased burden. Among the topics countries are considering for additional questions are; whether absences from work are related to COVID-19, whether differences in hours worked (versus usual hours) are related to COVID-19, and teleworking arrangements. The ILO has already published a note on [essential LFS content](#) and plans to develop additional guidance on possible alternative/additional content.

3.4 Other options and challenges

As with scenario 1 there will be a variety of other options and challenges facing countries under the scenario as discussed in subsequent sections of this note.

4 Temporarily suspending interviewing

Temporary suspensions of interviewing are widespread due to COVID-19 restrictions. Their duration will be variable and quite often unknown. They can refer to the entire sample or only part of the sample. In this section we address some of the options available to deal with these suspensions.

Some examples, of situations where this can be faced include:

- A country moving to telephone interviewing but needing to suspend interviewing temporarily to create systems and processes to enable this change.
- A country not changing mode but facing a disruption due to office closures or restrictions on movement. This can exist for example for countries operating CATI interviewing at a central facility or call centre but the facility has been closed.
- A country running a periodic survey (e.g. annual survey covering only part of the year) deciding to suspend interviewing until the lockdown is ended. This may or may not include a change in mode.
- A country managed to moved to telephone interviewing but the achieved sample was either too small to produce estimates with the desired precision, and/or the achieved sample was not representative of all groups of population (hence estimates would be highly biased if the rest of the sample is not interviewed). The country could decide to temporarily suspend interviewing until conditions allow data collection on a wider scale.

Whichever the specific case a disruption of data collection of some duration will be the outcome. The guidance under

previous sections provides information on some of the possible challenges to be faced to recommence interviewing, for example by a new mode. However, there will remain gaps in the data that need to be filled if possible. Some of the issues and options for this are discussed below. Any of these options will likely still have negative impacts on data quality of varying degrees, for example the possibility to produce usual disaggregations, precision of indicators etc. These impacts should be assessed to the extent possible, and reported transparently (for example appropriate flags to indicate discontinuities in series, publishing the data as separate series etc.).

In some cases quality assessment may lead to a conclusion that data is not of sufficient quality to publish. Quality assessment and reporting are discussed further in [section 8](#).

4.1 Choosing the appropriate reference period to use when the survey recommences

One important consideration when a suspension of data collection occurs, is the reference period to be used when the survey recommences, linked to the duration of the disruption. This also links to the current strategy for assignment of reference periods. Some countries use a fixed reference week, while others use a rolling week (always one week before the interview). The implications of suspensions in interviewing are somewhat different.

For those using fixed reference periods there is a trade-off to consider between maintaining original planned reference periods, or moving to more recent reference periods and/or changing the types of questions asked. If the nature of the periods covered is different (e.g. different points in seasonal/agricultural cycles etc.) then the change in reference period would be likely to have an impact on results and cause breaks in the series involved, in addition to any shifts in the series which would inevitably occur due to the economic impacts of COVID-19. However, retaining original reference periods entails recall bias.

For those with rolling reference periods the change in the period covered will be the main concern, as the interview period will be different than originally intended if the collection period moves. This raises the question as to whether to stick with rolling reference periods, or attempt to implement fixed periods, covering the originally intended survey period.

4.1.1 Retaining current content and reference periods

Allowing a number of weeks after the reference week to cover all households is in any case a normal LFS practice for those using fixed reference periods, but typically the number of weeks allowed is low (two or three weeks perhaps). An important difference compared with the current situation is that now the delay is applying to all (or a large part of) the sample, rather than the low number of households who could not be covered in the initial weeks. As a result possible recall biases, for example for variables like working time, or even employment in the case of casual working activities, become more important. However, in the case of the COVID-19 lockdown the unusual nature of the circumstances would at least make it likely that the timing of main changes in situation would be recalled, allowing some information on the working situation for the desired reference period to be captured. This option will be particularly attractive if the duration of the lockdown/restrictions is short (a few weeks). However, it may have implications for the level of detail that can be collected, meaning some reduction of content may need to be considered. It may also impact the comparability of the data collected due to recall effects and this would need to be assessed.

For those with fixed reference weeks the manner of implementation of this will simply be to retain the original fixed reference weeks. For those with rolling weeks this would in fact involve shifting to fixed weeks covering the originally planned period, alternatively if rolling weeks are retained it needs to be understood that this will have an impact on results if the periods covered are different in nature.

Summarising the pros and cons of this approach we can say:

- **Pros:** Does not require any redesign of the survey or other methodologies
- **Cons:** Introduces potential recall issues for a large proportion (or all) of the sample. Creates major pressure on timelines to complete all interviews quickly once operations recommence.

4.1.2 Changing reference period and/or using retrospective questions

For any longer duration of suspension it would remain advisable to attempt to collect information from the missed households covering the lockdown period, but in this case it could also be considered to update the questionnaire content to ask retrospective questions, phrased differently to the normal LFS questions (either alongside normal LFS short reference period type questions or in replacement of them). These could, for instance, ask about any changes in employment situation and when/why they occurred (for example whether they were COVID-19 related). This will most likely not be sufficient to fully apply criteria to assign LFS status in line with international standards, so there will be limitations on the use or interpretation of the data generated. Nonetheless, it may provide useful information to avoid a complete loss of data for the periods involved. The pros and cons of this approach include:

- **Pros:** A redesigned questionnaire should reduce recall bias versus the maintenance of original fixed reference weeks although recall biases can undoubtedly remain if the duration of the suspension is relatively long.
- **Cons:** Requires the questionnaire to be updated. This should be tested, even if briefly on a small scale to avoid operational problems when implemented. Additional care will be needed to integrate the data collected in this manner with existing data if the disruption only covered part of a survey instance. Even if successful the comparability to previous periods may not be high.

For both longer and shorter disruptions there will be resource implications assuming there is a regular survey process that needs to continue. Specifically, the exercise to cover the missed sample is likely to overlap with the ongoing survey (except for those with irregular periodic surveys). For example, if the last two weeks of the first quarter of 2020 were suspended then they would need to be covered at the same time as the early interviews for the second quarter. This may require a temporary increase in interviewing resources to avoid knock on effects, which could evidently become greater the longer the disruption and/or the larger the number of households which have not been covered. This may ultimately make this approach unrealistic.

Other possible approaches could be evaluated such as a hybrid approach that re-interviews the missed households asking them both about a recent reference week and the period of the lockdown for which they were originally scheduled to be interviewed, i.e. they are used both for the current period and to get some retrospective information for the missed period. Along with all other options involving changes in content or design, this would require careful development, implementation and assessment of the results generated.

4.2 Covering missed interviews at a later date in case of low response rates

Even where operations are not completely suspended, countries may still consider one of the above approaches necessary if response rates were low during the normally planned data collection period. Furthermore, the period of the disruption will almost certainly be substantially different (in labour market terms) from the period that was more comprehensively covered, meaning using the data as originally collected would not provide representative estimates for the full period. Attempts to increase coverage could take the form of a minor extension in the period allowed to collect the data in order to raise the response rate to acceptable levels, or it could involve applying the retrospective questioning approach mentioned above if a longer period is required to cover the missed sample. In all likelihood the low response will be biased, for example poor coverage of rural areas, so any attempt to increase coverage retrospectively also needs to attempt to reduce bias. In addition, the change in approach or additional recall bias involved can make integration of the data complicated. As such, while potentially worth considering, this is an approach with many complications, both to implement and to subsequently use the data.

4.3 Periodic survey

Another specific case which does arise is a periodic survey - for example an annual survey with data collection in one or two periods in the year, or a quarterly survey with one week collection per month. In this case timelines may not be as constrained as other cases because the next round of survey collection may be planned for a later date, for example if the current survey round was planned for May (now to be suspended), with the next one in November. The most likely approach will be to suspend the collection until it is feasible to recommence through the appropriate mode. Some

countries in this position are suspending interviews by the period of time required to prepare for telephone interviewing in the anticipation that any easing of restrictions will not enable face-to-face interviewing to recommence.

If the survey is infrequent (e.g. once every 2 or 5 years) this may be a less substantial problem than a more regular survey (e.g. quarterly or annual periodic) but nonetheless consideration should be given to the likely impact on any time-series or the interpretation of the data generated. For instance, if the time period now covered has different seasonal characteristics than the originally planned period (e.g. covering a low agricultural period as opposed to a peak period) this would have an impact on the results generated, all other things being equal.

Like all options, the identification of the most suitable approach requires assessment and planning to avoid unintended consequences for the data generated, followed by transparent dissemination of information about the survey.

4.4 Timing of the disruption (during a cycle or covering a full cycle)

An overarching consideration that is touched on in the above discussion is that a disruption may occur at different points in a survey cycle. One additional challenge worth noting in the case of a disruption mid-cycle will be the challenge of integrating data from two separate approaches, for example where half of the survey has been completed as normal before the disruption, and the other half completed at a later stage, possibly with a variety of different methodologies. Alternatively, it may be the case that the other half cannot be completed, thus creating the need to decide how to use the collected data. This creates a variety of challenges such as the need to treat and assess the data. Some relevant issues related to non-response/weighting are discussed in [section 7](#), while issues of assessment of data quality are discussed in [section 8](#).

5 Other mode-related scenarios

The two main scenarios outlined in [sections 2](#) and [section 3](#) cover many of the challenges and options NSOs are facing to maintain LFS data collection at this time. However, the scenarios are not exhaustive given the many variations of approaches in use for the LFS across countries. Some other selected scenarios or options are worth a brief mention, as they are being faced or considered by some countries, although they will not be addressed in detail here.

- **Introducing CAWI:** Some countries ([see note on the impact of COVID 19 on the collection of labour market statistics](#)) made reference to plans to introduce CAWI. While CAWI is being increasingly used in the LFS, it is not used (or likely to be used) in isolation as the only mode of data collection. Experience to date shows that CAWI can be useful to increase response in certain groups (e.g. those for whom it may be difficult to arrange interviews), but the main benefit is to reduce the cost of data collection. However, if used in isolation it will lead to partial and biased response without other modes. Therefore, it is not recommended to plan to introduce CAWI as the only mode of collection for the LFS. Further, the introduction of CAWI is a significant development. The questionnaire will require adaptation for self-completion, online guides will need to be developed, survey management processes and systems will need to be updated etc. This will clearly be more feasible if the NSO has experience of CAWI surveys, but would nonetheless require a good development and testing process before implementation.
- **Disruptions for countries fully using CATI/CAWI:** In general, countries already entirely using CATI/CAWI are not anticipating major disruptions to their data collection operations. Indeed some countries in this situation reported some positive impacts on response rates due to increased availability of respondents or increased positive perception of Government services during the COVID-19 crisis. Nonetheless, some of these countries are facing disruptions to their CATI operations if being required to move from centralized telephone facilities to decentralized interviewing. The guidance presented in the other sections of this note can be of relevance to these countries.

6 Other options or changes - sampling

The discussion under the scenarios above assumed that methodologies are unchanged other than those practically necessary to enable a change in approach. In general this will be recommended, if feasible, in order to avoid additional sources of inconsistency or incomparability. This will be particularly likely if a panel design is in place as coverage problems may be limited to only a subset of the sample, and changing the design for part of the sample will create complexities to integrate the samples during processing and analysis.

However, this may not be feasible in all cases, or at a minimum countries may need to assess alternatives versus maintenance of the status quo, for example in case early assessments lead to a conclusion that achieving high coverage of the existing sample will be impossible.

Some of the specific options countries are considering are discussed in this section, specifically in relation to sampling. In considering these options countries should evaluate them against other options available, including temporary suspension of data collection with the possibility to attempt to recover missed sample at a later date as discussed in [section 4](#).

One option not addressed in any detail in this note is the possibility to over-sample or increase the size of the sample. This could improve precision but would not necessarily address biases. It is generally assumed that resource constraints will not make this a feasible approach, nonetheless some countries may wish to consider it, for example by selecting extra FSUs in each PSU.

6.1 Using the same frame but selecting a new (reduced) sample

If for various reasons (lack of contact details, insufficient resources) it becomes clear that it is impossible to achieve good coverage of the existing sample, then consideration could be given to reducing the sample size but maintaining the sample frame and main elements of the sample design. However, if this is done it should be pursued in a planned way to avoid possible biases. Sample reduction does not mean that the NSOs interview only FSUs with available telephone numbers or only select those units when picking the new sample. Either of these approaches would introduce bias and reduce the probability of selection of some valid units to zero.

The most straightforward and easiest to recommend option would be to maintain existing primary sampling units (PSUs), if used, and select a smaller number of FSUs (generally households) within each PSU. The scale of reduction should be even across PSUs.

It could also be considered to change more important elements of the design such as strata, number of PSUs etc., for example by reducing the number of PSUs within each strata. However, this could generate additional inconsistencies versus existing approaches and should be planned with great care.

Inevitably, a reduction in sample will lead to a reduction of precision but this could be preferable to a larger but biased sample if insufficient resources were available to contact all households. Selecting the appropriate size should be done in the same way as the original sample was selected, typically based on a chosen target level of precision for key indicators such as the employment to population ratio and unemployment rate.

Perhaps, another way of framing this option is to say that NSOs should consider if sufficient resources and contact details are available to achieve complete coverage of the existing sample. If not, and additional resources cannot be allocated, any reduction in sample size should be done in a planned way, rather than allowing foreseeable biases to arise.

6.1.1 Other sub-sampling approaches

Some countries are considering a move from household sampling to sub-sampling individuals within households. This needs to be distinguished from countries who, by design, directly sample individuals for interviewing from a sample frame of individuals, which is a perfectly valid approach when designed and implemented well (as with household sampling).

The immediate impact of sub-sampling individuals within households (for example 1 individual per household) will evidently be a substantial reduction in sample size versus the planned sample. However, this potentially creates additional

challenges that require consideration, including:

- How to handle the process of selection of individuals within the household. If just one individual per household is selected then this could be random selection, done by some means by the interviewer following a basic household listing/roster exercise at the beginning of an interview. This may be complex and subject to error. Also in this case individuals in different households will have different probabilities of selection, i.e. lower probabilities of selection in larger households if the number of individuals to be selected in each household is fixed.
- Various sources of bias could arise if the respondent selection process cannot be well designed and implemented. It may be difficult if not impossible to correct for these biases, or even evaluate them after collection has been completed.
- Weighting approaches would need to be updated as the sample is no longer household based.

While these challenges may not be impossible to resolve, they create significant complexity, require new processes and methodologies to be created, will lead to potentially major (possibly biased) reductions in sample, and the reduction in cost to the NSO will not be proportional to the reduction in achieved sample as contact still needs to be made with the same number of households. With all these considerations in mind, it would not generally be recommended to switch to such an approach, rather consider other ways to reduce sample size in a planned way as discussed in the previous sub-section.

6.2 Using a new frame

It could be the case in some countries that a (relatively) comprehensive source of contact details exist but that it is impossible to link it to the existing sampling frame or selected sample. This could then become a viable option to use as a sampling frame but caution is certainly needed in such a change of approach. Firstly, as with any sampling frame an assessment of coverage will be needed. The issues to be considered will differ depending on the source, some of which are listed below:

- **A register of mobile phone numbers:** In this case, the registers covers mobile phones assigned to individuals and may or may not have unique addresses (meaning it may not be possible to link/collapse the information to households). Other potential issues will include:
 - Incomplete coverage (potentially biased), for example in rural areas with poor network coverage, among poorer households, cross-border or migrant workers, individuals choosing not to be included on directories etc.
 - Irregular updating meaning some listed numbers will no longer be active or more recent activations not included
 - Accurate information may or may not be recorded for all mobile phone numbers (e.g. prepaid phone numbers with limited or inaccurate contact information)
 - Individuals may have multiple phone numbers.

All these issues should be considered in an initial assessment of the source with the guiding principles generally adopted for sample frames - for example, whether there are known probabilities of selection, non-zero probabilities etc. If this is the only source used it becomes unlikely that a sample of households can continue to be used, instead becoming a sample of mobile phone numbers (not individuals, unless there is a high degree of confidence that individuals can be uniquely identified and associated with one or multiple phone numbers). In addition, legal issues may arise if some of the data required is not publicly available (e.g. confidential customer contact information or the details of customers who have chosen to be ex-directory).

- **A register of landline numbers:** The advantage of this approach is that it generally identifies a unique residence (although in limited cases a single residence could have multiple landline numbers) but coverage will not be complete given increasing numbers of people and households who only use mobile phones. As with mobile phone registers an assessment will be needed to establish the scale of any bias, and any approach to sampling will need to reflect the

additional information available, such as address etc.

Alternatively, multiple sources may be linked together, for example different administrative registers from tax databases etc. The guiding principles will remain the same as those outlined under scenario 1, namely:

- Assess the information available on the frame
- Based on that information assess coverage of the frame against any other comprehensive source with full population coverage (e.g. the original sampling frame of the survey). In particular it will need to be assessed if the coverage of the new frame is biased in any important way. If it can be established that the coverage is incomplete but unbiased (this may be difficult to verify) then using the new frame would be less problematic than if it were biased.
- The above assessments may be supplemented by testing, for example take a selection of the units on the new frame and attempt contact to verify the information.
- Assess possible sampling approaches based on the information available, for example, what units can be uniquely identified for sampling (people, mobile phone numbers, households)?; what, if any stratification information is available?; will there be known non-zero probabilities of selection for all units to allow design weights to be calculated?.

Using this information an informed decision can be taken on the viability of the selection of a sample from an alternative source. This should include a comparison to the continuation with the existing frame and design to identify which option is best (or least bad). Given the complexities and potential biases involved it will remain the preference for most NSOs to attempt to maintain the existing sample frame and design, but if alternatives are to be considered it is important that they are adequately assessed to enable an informed decision to be taken on their use. If the sample frame is changed, this will inevitably have implications for other stages of the process, in particular weighting as discussed in [section 7](#).

7 Ex-post (after data collection) challenges or issues to consider

The above sections deal with ex-ante challenges of design and the challenge of actually completing the survey. At this point we switch our attention to ex-post challenges once the data has been collected by whatever mode or methodology has been implemented.

7.1 Response rates and bias

Maintaining the existing sampling design will be relatively most straightforward if the potential remains to achieve good coverage of that sample. As described before, good coverage could be maintained either by continuing interviews with a different mode during the lockdown, or completing the data collection operations at a later stage, or potentially a combination of the two.

For many reasons it may not be possible to achieve a representative sample with as high response levels as desired and this could lead to bias in coverage in various ways. While the total response rate will remain an important indicator to monitor, the assessment of bias will require us to look beyond it. A high response rate could still be biased if the achieved sample is unrepresentative, while a lower level of response may in fact be less biased if non-response is randomly distributed.

Given that many countries are attempting to introduce or increase telephone interviewing at this time some particular potential sources of bias are worthy of note, such as:

- Some groups do not have landline or mobile phone or do not want their numbers to be published on phone directories. The challenge of obtaining contact details has already been discussed above.
- Other groups are difficult to contact although contact details are available - generally because they have lifestyles that keep them out of their houses for most of time during the day and week, hence being difficult to find them home. In the current crisis availability may improve for many people, but for some (e.g. front line healthcare workers) their

availability may in fact be much more limited.

Bias arising from non-response is not a new issue, but in the current context of widespread lockdowns and possible increased non-response, the problem assumes higher relevance. Indeed, one of the primary mechanisms used by countries to avoid biases as a result of differential non-response is to use mixed modes, for instance, continuing to use face-to-face interviewing to cover cases which are difficult to contact, something likely to be impossible in the current circumstances. Consequently, increases in bias become likely while fewer options are available to deal with it during data collection.

These or various other issues (for example insufficient resources to fully cover the sample) could lead to a bias in the coverage. This becomes true if the FSUs who are actually interviewed have different characteristics as regards their labour market attachment/characteristics from those who are not - i.e. those who could not be contacted were either more or less likely to be employed/unemployed than those who could be contacted. These biases could appear either in estimates at the total level of different sub-groups.

To the extent possible NSOs should evaluate in advance the proportion of the original sample that could be effectively covered with new approaches/modes, for instance by reviewing coverage level of contact details for different groups. This may allow us to anticipate and attempt to minimize possible bias, but it may not reflect the reality once data collection has been completed. Furthermore, predicting bias in advance is not always easy as it is often linked to characteristics that are not known/present on the sampling frame and may only become visible once the data has been collected.

Handling non-response ex-post starts with an assessment of possible bias. This involves a comparison of the characteristics of the actual achieved sample with the sampling frame, population figures or past samples. If the assessment suggests that possible biases may exist - for example few households covered in some rural areas but higher coverage in urban areas - NSOs will have (at least) two options:

- **Attempt to achieve additional interviews** for those groups that have been less well covered. This can be feasible where the undercoverage is clear and based on variables included on the sampling frame (for example lower coverage in some cells in the selected sample than others) as the additional interviewing can be targeted to certain FSUs or PSUs. However, if the undercoverage is shown on other criteria (e.g. very low coverage of youth) and that information does not exist on the sampling frame then additional interviews may not be an efficient means to reduce bias.

One approach which can be considered if additional interviews are to be attempted is to publish initial provisional estimates, following original timelines, using the original sample and provide updated final estimates once additional data has been collected. Like all options this has pros and cons including a general dislike among users for revisions and the obvious additional burden to the NSO to follow such an approach, traded off against timeliness and accuracy of the estimates.

- **Try to reduce the bias with non-response adjustments and weighting** (next sub-section).

7.2 Non-response adjustment and weighting systems

Non-response adjustment and weighting approaches are closely linked parts of a strategy that can reduce bias in estimates when achieved samples are biased. In some countries they are in fact dealt with in the same weighting step, whereby the achieved sample is weighted to some external population constraints. Calibration approaches are also used for this purpose, generally allowing a greater set of constraints to be imposed than more basic post-stratification weighting procedures. Non-response, weighting and calibration are discussed in many available guides and manuals, such as two manuals from United Nations Statistics Division (see: [Designing Household Survey Samples: Practical Guidelines](#) and [Household Sample Surveys in Developing and Transition Countries](#)). The guidance of these and other manuals will not be repeated here, instead key points of note for LFS in the context of the current disruptions are highlighted.

As a general point it is not recommended that countries choose to alter their weighting and non-response adjustment processes at short notice unless necessary, as it will introduce an additional source of incomparability versus existing

methodologies. In addition, it will place an extra burden on the NSO at a time when many challenges are being faced and these methodologies can be complex to review or introduce.

Nonetheless, countries may feel this is required if there are major non-response issues, or wish to assess alternatives. Additionally it may be necessary if some important element of methodology has changed, such as a move from household based sampling to sampling of individuals.

7.2.1 Non-response adjustments

In cases where separate non-response adjustments are being used, they introduce adjustments to impose the characteristics of the originally selected sample. This is done before any post-stratification weighting or calibration. A key challenge is that any adjustment requires some information about both responding and non-responding households. Typically, in practice this can be limited to the information used to stratify the sample and for this reason non-response adjustment sometimes only involves adding an increased weight to each responding household based on the response rate within the strata/sample cell. Some countries are able to use auxiliary information, for example from administrative sources or additional information from the frame to provide some other basis for adjustment. The possible sources and range of information will vary substantially by country.

Some countries have started to simulate possible updated approaches to non-response adjustment and weighting using old samples with assumed/modelled levels of non-response, for example selectively removing some records and testing alternative approaches to non-response adjustment. The estimates generated are assessed to see if different approaches generate more or less volatile results, or have impacts on the ability to create disaggregations etc. Where a change from face-to-face to telephone interviewing is involved, the simulation process could remove some records to simulate non-availability of contact information, as described [in chapter 7 of this paper from The Netherlands](#). These types of studies are extremely useful in the current context and NSOs should consider how they can simulate the results of different approaches before their introduction.

Also, countries that use panel designs could build non-response adjustment models making use of longitudinal data, i.e. comparing the structure of the longitudinal sub-samples over time and correcting for panel attrition and non-response. The same applies for countries that decide to re-use old samples, as described in [section 2.1.1](#). In other words, the information from previous rates of completion and attrition, including by mode (if available) can be used to simulate the impact of changes in approach based on some assumptions or the results of any tests done of the new approach.

7.2.2 Weighting/calibration

Again, unless it is considered necessary, it is not recommended to change weighting approach as the changes may introduce incomparability over time. For this reason, as a general recommendation when any change of weighting or calibration approach is planned, consideration must also be given to recalculation of previous series using the updated weighting approach to remove this potential source of incomparability – this will also be a useful exercise to assess the impact of the changes.

Weighting methods are typically used to obtain estimates that are generally more accurate than those obtained using the design weight (i.e. the reciprocal of the probability of inclusion) only. Weighting methods allow us to equate the joint and/or marginal sample distributions with the population distributions. They also help to eliminate fluctuations from population totals due to the sampling strategy, and help to reduce the negative effects of total non-response and the under/over-coverage.

Several methods are available to weight sample data taking into account auxiliary information, including:

- **Post-stratification** - also known as cell weighting adjustment. This approach has been widely used for many years because of its ease of implementation.
- **Raking, generalized regression and calibration** are more complex adjustment methods that generally incorporate additional auxiliary information as compared with post-stratification. These approaches are increasingly used in recent years as the availability of software to enable them has increased.

While generally more flexible/advanced, calibration is also more complex. Therefore, while it is potentially attractive to move to calibration, by and large countries are suggested to continue with their current weighting approach, or at least allow sufficient time to develop any alternative.

The assessment or implementation of possible weighting alternatives is not a straightforward task. For example, let us consider countries that include sex in their current weighting framework but do not weight by age group. It would be possible that weighted sample estimates for certain age groups would be regularly under- or over-estimated (as they are not benchmarked to population figures). However, they may still be comparable if the under- or over-representation of age-groups within the sample is stable over time (e.g. age-group 15-29 is under-represented by 30 per cent in all survey rounds, 30-64 are always over-represented by 10 per cent etc.).

Under the current exceptional circumstances, the response rate can drop and the proportions represented by the different age groups might change as well, in other words either new bias is introduced, or the nature of existing bias could be changed. Countries could be tempted to introduce a benchmark by age groups to reduce the bias. However, while this can be a valid strategy, it will affect comparability with earlier estimates, meaning consideration must also be given to applying the new approach to previous periods and recalculating estimates.

It cannot be assumed that non-response adjustment or weighting can correct for all biases as it will only apply corrections on certain selected criteria/variables, meaning other variables not adjusted for could still lead to biases. For example, persons aged 18-29 years old living alone or as couple could have very different labour market characteristics from persons of similar age and sex living with their parents (e.g. the former being more likely to be employed with permanent contracts, higher pay etc.). If non-response adjustment or weighting corrects for age but not household type then any differences in response rates between those types of households could still generate bias. As such, identifying the correct set of variables to use in non-response adjustment and weighting processes will be essential to reduce biases.

There is no single recommended approach to weighting or calibration that would necessarily generate improved results in response to the disruptions face as a results of COVID-19. If the primary problem expected is that there may be differential non-response for some subgroups of the population, then the existing weighting scheme needs to be reviewed considering how well it may or may not correct for any biases that are identified, and on that basis whether superior alternatives may be available.

If alternative weighting approaches are being considered it is recommended to test them using available data from previous periods to an assessment of the results generated, for instance to assess if results are more volatile under some approaches than others, the impacts on disaggregated results etc.

The options mentioned [in section 6](#) (changes to sampling) in particular may require an update to weighting or calibration approaches, for instance if a change was made to a household based from to an individual based frame. The specific case of reusing an old sample ([see section 2.1.1](#)) also may warrant a specific calibration of the achieved sample in the new round to its shape or composition from previous rounds.

7.3 Summary on bias, non-response and weighting

It is likely that many countries will face difficulties to achieve usual levels of response due to the many challenges they face in continuing data collection for the LFS. A thorough assessment of biases created by non-response will be needed to evaluate the impact this will have on estimates. This can be done to some degree ex-ante and inform the strategy adopted for the survey operation. However, an ex-post assessment will also be required.

Non-response adjustment and weighting approaches can correct for biases. Where robust weighting and non-response adjustments are already being used it is generally recommended to continue to use them, unless forced to change, as changes will reduce the comparability within time series.

If changes in approach are being made, like all changes discussed in this note, they should be planned and implemented carefully, considering various options, and taking into account national resources and data availability.

8 Assessing data quality and reporting the data

The above covers some of the actions, which can be envisaged to face the challenge of maintaining a large scale survey of the nature of the LFS. The correct approach to take in any given country can only be determined through an evaluation of the circumstances in the country and the different options which may be available. This evaluation needs to be holistic considering the many implications of possible approaches for data quality. In other words, the choices made should be guided ex-ante by expectations of impacts on data quality, alongside resource implications or any other considerations relevant at the national level.

However, regardless of ex-ante evaluations, impacts on data quality will be almost inevitable, and will only become more fully understood once the data has been collected and processed. This section provides brief guidance on the task of the assessment of data quality.

There is a high likelihood for many countries that the disruptions being caused by COVID-19 will lead to a break in series for methodological reasons, independent of the shifts in series which will inevitably arise as a result of the economic impacts of COVID-19. These breaks in series could arise from any of the challenges or choices referenced above, for example reduced response rates, mode effects, increased non-response errors or many other sources.

These possibilities create a very significant emphasis on the assessment of data quality, a common task for all NSOs, but of even higher priority in the current circumstances.

Quality assessment and assurance is covered by a variety of guidance, and subject to many differences of approach across countries. The [UN National Quality Assurance Framework Manual](#) provides an overview of the elements required for quality assurance at the national level, including for individual statistical products, as well as the national statistical system. While the manual identifies several key dimensions to assess the quality of a statistic, it is not highly prescriptive. A more prescriptive/detailed approach is presented in the [ESS Handbook for Quality Reports](#), which, while set in a European context, can be a useful reference for NSOs seeking to implement a comprehensive quality reporting approach, or elements of one.

While many dimensions can be used to describe the quality of a statistic, those of most immediate relevance to COVID-19 related disruptions will be *accuracy* and *comparability*. Other identified dimensions such as *timeliness* and *punctuality*, *coherence* etc., will undoubtedly also be important, although less immediately and directly impacted by disruptions or changes to LFS data collection. Table 1 below shows some recommended headings and indicators to cover in quality reporting for *accuracy* and *comparability* with further detail provided in the ESS handbook.

Table 1. Main headings for quality reporting under the dimensions of accuracy and comparability

Dimension/indicator	Possible impact
Accuracy	Lower accuracy due to the combined impacts
Sampling errors	Likely increase in sampling errors if achieved samples fall
Coverage errors	May increase, for example if a change in sampling approach/frame is introduced with incomplete coverage
Measurement errors	May increase if a change in mode is implemented without sufficient preparation/testing
Processing errors	May increase if changes to processing systems and practices need to be changed
Non-response errors	Likely to increase due to difficulties in maintaining response rates
Comparability	Possible breaks in series (one or multiple)
Comparability over time	Any change in methodology, practices or achieved response may create breaks in series. There is the potential that this may occur multiple times if changes are made in multiple periods. Simultaneously there will be actual shifts in time series due to the real world impacts of COVID-19 on the labour market.

Source: [Handbook on data quality assessment methods and tools](#), Eurostat

NSOs in general should include quality assessment in their regular processes, even if not applying something as comprehensive as the European model. For example, sampling errors may already be calculated and can continue to be reviewed. They are however likely to be insufficient to fully describe the impact on data quality of disrupted LFS data collection.

All NSOs are recommended to undertake additional assessment of quality of any statistics generated, using the above headings and others of relevance. This will be important both to consider possible adjustments to methodology (e.g. imputation, non-response adjustment, weighting and estimation), as well as to allow for proper assessment of data quality before and alongside dissemination of the data.

However, no evaluation framework or approach can completely replace or remove the role of expert judgement. Compilers can use any assessments completed to enable them to make decisions on when and how to report the statistics, but that decision also needs to be built on an understanding of the country context.

It is already being reported by some countries that, despite all efforts, low and biased response has been leading to implausible results, for example an increase in employment that can be directly linked to an inability to achieve representative coverage of the population. In this situation some countries are understandably taking the decision to delay publication to consider options (such as those outlined in earlier sections) to achieve higher coverage or address bias. While this is not a decision to be taken lightly it may well be preferable to publishing results according to normal timetables, despite clear data quality concerns, something that could potentially lead to serious reputational damage. The circumstances under which it would be advisable to delay publication are in general very limited for very good reasons, but those circumstances are more likely to exist where data collection has been significantly impacted, in the way currently occurring in many countries. Data compilers will need to be thorough in their assessment and apply judgement in deciding when and how to communicate the statistics. Any delay should be explained clearly and transparently to users, along with an indication of remedial actions being taken to enable publication of the statistics in future.

When the data can be published, as a minimum if it is concluded that breaks in published series have occurred (or are likely to have occurred) due to changes in mode, methodology, response rates etc., this should be indicated to users, even if the scale of the break cannot be estimated. This should extend to the inclusion of appropriate flags on databases, or potentially creation of new series if the differences are particularly substantial leading to a conclusion of a major break/incomparability. Even in a case where a new series is being presented, in publications it would be expected that both old and new series (if relevant) are presented and explained in as clear and transparent a manner as possible, potentially through additional ad hoc information notes or FAQs to ensure their interpretation is clear.

If there have been changes in methodology or practice but it is unclear that they have led to breaks, explanatory or methodological notes should still clarify the changes which have occurred, along with the reporting of any available quality indicators (such as sampling errors, response rates etc.). The need for some caution in interpretation of changes among users of the data should be noted.

In some cases, it may be ascertained that despite all efforts, usable data could not be generated, even where some data could be collected. Hopefully this can be avoided, but in this instance the possibility to holistically and transparently assess data quality is equally critical to support the process of explaining this to users if data cannot be published for some periods or series.

Regarding the type or content of publications to be disseminated, some flexibility of approach is also highly recommended. Countries may wish to consider additional or new publications, alterations to current publication content etc., and any important methodological changes should also be clearly described.

9 Summary and conclusions

While this note touches on many scenarios, challenges and options, it still does not do full justice to the scale of the challenge NSOs or other data producers face at this time. All NSO operations are being impacted in some ways, meaning the challenge is far from over if data collection has been completed reasonably successfully. Further the impacts are evolving over time, making planning far more difficult. A good plan in April may no longer be appropriate in May if conditions change substantially etc.

Nonetheless, options do exist for most challenges that can, at the very least, form a basis of the development of a strategy to re-establish the flow of labour data, even if it has been temporarily disrupted.

The ILO will continue to generate further guidance and provide direct support to countries in the development and implementation plans to navigate through the unprecedented circumstances.

Like many other domains, the COVID-19 pandemic prompts those involved in the compilation of labour statistics to consider what we can learn about our practices and methodologies, and in particular how we can make them more flexible and resilient in the future. We can say with certainty that CATI and CAWI based collection is less impacted by a crisis which causes restrictions in movement and contact than face-to-face interviewing. Computer based approaches in general will have advantages over paper and pencil. We also can learn about the need to ensure we have good contact information for households and individuals to enable contact to be achieved when needed. Other elements of processes can also be improved or made more flexible/resilient, such as approaches to questionnaire design (modular approaches) or weighting etc. Even if unable to make updates to approach in the short-term, all countries should consider how to update their practices in the longer term to create greater resilience.