

Continuity of household surveys after the coronavirus disease (COVID-19) pandemic

December 2020

Summary

In the context of the global coronavirus disease (COVID-19) pandemic, the Economic Commission for Latin America and the Caribbean (ECLAC) is preparing a series of short publications with key policy recommendations for this period. This note, the third on household surveys, offers some considerations regarding the continuity of this type of statistical instrument in a context of initial lifting of the restrictions on movement in the region that were put in place to contain the spread of the pandemic.

Introduction¹

In the COVID-19 pandemic, most countries of Latin America and the Caribbean took measures to restrict citizens' movement. These restrictions prevented household surveys from being taken face-to-face. As a result, some surveys were suspended, and others were conducted by telephone.

The change in the data collection mode from face-to-face to telephone—which was necessary to continue producing employment and income statistics, that are particularly important in the pandemic—has posed new challenges for national statistical offices and other public agencies that take surveys. These challenges were shared and analysed in various virtual seminars as part of the work of the Knowledge Transmission Network of the Statistical Conference of the Americas of ECLAC, as well as during the nineteenth meeting of the Executive Committee of the Conference (see ECLAC, 2020a and 2020b).

The ECLAC Statistics Division also made some recommendations to guarantee the quality of statistics with the changes in collection methodology and has provided technical assistance to several countries for their implementation (see ECLAC, 2020c and 2020d). The recommendations focused on the inherent properties of sampling strategies, namely the survey sampling designs and estimation procedures used to draw inferences from collected data. In short, it was recommended to:

- Base publication of official statistics from household surveys on probabilistic selection of samples and not on predictive models.
- Define probabilistic monitoring panels, based on the most recent months of face-to-face collection, to conduct surveys by telephone, maintaining the population contained in the national statistical offices' master sampling frame as the reference population.
- If response rates were low, publish statistics at the national level, without the usual disaggregation, explicitly mentioning the period over which the information was gathered.
- Use models to correct selection and coverage bias, which would allow for adjustment of weights based on the auxiliary information available in the monitoring panel.

¹ The cut-off date for the information used to prepare this report is 31 October 2020, unless otherwise indicated.

Summary

Introduction

A. Suggestions for mixed-mode household surveys in the return to a new normal

B. Comparability and the impact of COVID-19

C. Conclusions

Bibliography

Since the start of the pandemic, most of the countries of the region have satisfactorily addressed the challenges posed by the exceptional situation, continuing to publish official labour market statistics obtained through telephone or mixed-mode surveys. The progressive lifting of restrictions on movement is creating new circumstances, in which face-to-face household surveys are gradually being resumed. This document sets out some key considerations for this period. The procedures that were recommended and implemented under the restrictions on movement should not be abandoned, but adopted as regular processes for household surveys. It is also proposed that flexibility be maintained given the possibility of new lockdowns, and that consideration be given to a post-pandemic period in which there cannot be reliance on face-to-face surveys alone.

A. Suggestions for mixed-mode household surveys in the return to a new normal

Telephone surveys have proved to be a valuable alternative when movement restrictions are in place, and can be reconciled with face-to-face operations. This experience can be drawn on in the future, so that survey logistics offer several ways of gathering data from selected individuals: in person, by phone or online.

Mixed-mode surveys offer greater flexibility in the current situation. In view of the risk of a fresh rise in COVID-19 cases and a need to re-establish restrictions on people's movement, a mixed-mode approach allows for continued gathering of information. In addition, the telephone survey mode offers new ways of increasing the efficiency of surveys and reducing their cost.

In the medium term, as new data collection methods become accepted by respondents, household surveys are expected to be conducted using a combination of different collection modes, even after the pandemic is considered to be over.

However, certain details do not lend themselves to collection by telephone, and some population groups will only be accessible in person, such as those living in slums and rural areas. In such cases, face-to-face interviews seem the most appropriate way to gather information.²

The following section provides some reflections on conducting household surveys with continued mixed-mode information collection.

1. Sample selection and disposition codes

On the whole, mixed-mode surveys will continue to be subject to the usual selection of primary sampling units with corresponding subsampling down to households and individuals. This selection is based on the probability measure induced from the (most often stratified and multi-stage) complex sampling design from the cartographical area frame in the country. This selection is exhaustive in that any area of the country's mapping grid can be selected, and therefore any household or person.

The map updates that record and list dwellings with private households must be strengthened through surveys that gather exhaustive contact information of pre-selected households: mobile phone numbers, landlines and emails. In addition, in these surveys it is recommendable to record the number of occupants of the home, along with their basic sociodemographic data and, if possible, information that enables the residents of the home to be linked to external administrative records.

The household surveys regularly performed in the region can be divided into two large groups: those with independent cross-sectional collection and those with some form of rotation in their sampling. For surveys with rotation structures, in which dwellings are part of samples for a certain number of periods, mixed-mode surveying could take the form of initial face-to-face interviews, followed by remote surveying in subsequent periods.

² Naturally, all necessary measures should be taken to minimize the risk of transmission of the coronavirus disease in face-to-face interviews. These include continuous and systematic monitoring of the health of the team working in the field; basic equipment and biosafety protocols for the protection of the respondent and the interviewer; and the required sanitization and disinfection material.

Therefore, although it may be feasible to carry out face-to-face surveys, telephone contact could be used to conduct interviews, and even when interviews cannot be conducted, it will always be possible to adjust sampling weights to eliminate selection bias, owing to non-response and lack of coverage, since more detailed information collected in previous periods will be available.

Surveys that do not have built-in rotation schemes may also benefit from mixed-mode surveying, provided that contact information and some basic household characteristics are collected before the survey starts, to include them in the identification and elimination of coverage and selection bias. Consequently, map updates will be vital for adjusting sampling weights. However, there are clear benefits of rotating schemes in this type of situation, because actively searching for new respondents is easier than in cross-sectional surveying.

It is important to take into account the definition of disposition codes (eligible respondent individuals, eligible non-respondent individuals, individuals with unknown eligibility, and ineligible individuals) for different collection modes and their comparability. While there are standard classifications (see AAPOR, 2016) for the disposition codes used in most surveys in the region (face-to-face and telephone), it is strongly recommended to revisit each possible eligibility event, its meaning according to the collection mode used in the survey, and its corresponding classification in the survey. It would be incorrect to apply traditional telephone codes to mixed-mode surveys.

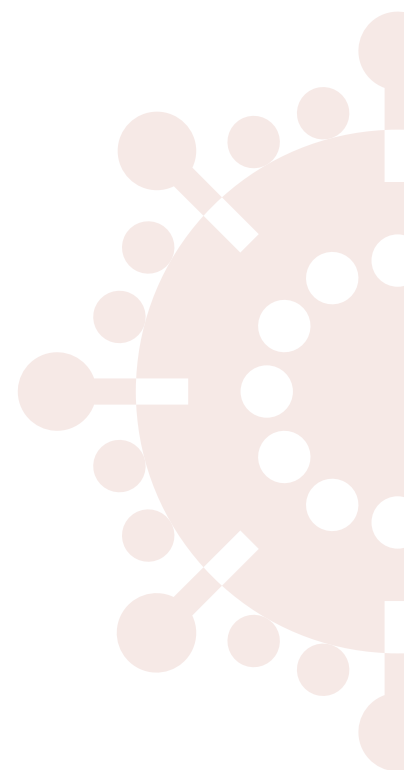
Telephone disposition codes are defined based on the assumption that random digit dialling (RDD) is used. Therefore, a selected number could correspond to a household (which is part of the target population) or to another type of entity that is not part of the target population. For example, telephone survey disposition codes would classify an engaged telephone number as “unknown eligibility” (impossible to determine whether it is part of the target population). In household surveys conducted by national statistical offices, where telephone numbers are linked directly to sampling units in an area frame, this event cannot be categorized as “unknown eligibility” but should instead be classified as “eligible non-respondent”. Therefore, it is advisable to carefully review the events associated with both face-to-face and telephone survey disposition codes and to formulate an appropriate classification for mixed-mode surveys.

2. Suitability of the mixed-mode format for all surveys

In most cases, telephone surveys do not capture all the constructs studied in face-to-face household surveys. Less time is available for telephone interviews than for face-to-face interviews, so questionnaires need to be shortened. Otherwise, surveys may not be completed properly owing to respondent fatigue or effective response rates may decline. In recent months, several countries in the region have used shortened questionnaires in their labour force telephone surveys, generally prioritizing information on the labour market and income, and excluding other commonly collected variables.

Therefore, in a mixed-mode survey, it is advisable to have two versions of the questionnaire, which allow for joint calculation of the main indicators generated through the survey. Thus, when an area is in lockdown and face-to-face surveys are not allowed, the adapted questionnaire can be used for telephone surveys. Reconciliation does not merely entail shortening the telephone questionnaire, but also checking whether there are questions that could be dispensed with in the face-to-face questionnaire.

In addition, certain household surveys do not seem feasible in a telephone format. One example is household budget surveys (income and expenditure), which require extensive contact over several days between the interviewer and the respondent. Obtaining this information by telephone puts a considerable burden on the respondent. The pandemic itself can also accentuate complex issues that are part of the constructs studied by victimization surveys or time-use surveys. In the case of victim studies, strict protocols to ensure respondents' safety cannot be applied in a telephone survey. In the case of time use, lockdowns during the pandemic may disrupt the usual structure of the population's regular activities.



3. Eliminating bias

One of the greatest challenges in changing the way surveys are collected comes from coverage and selection bias. It is assumed, as a fundamental principle, that households that respond to a telephone survey do not have similar characteristics to non-respondent or uncontacted households. Moreover, a failure to respond may be linked to the phenomenon being measured. For example, households that do not provide telephone contact information may have more people out of work, or non-response rates may be higher among households that have fallen into poverty because of the impact of the pandemic on their income. The document “Recommendations for eliminating selection bias in household surveys during the coronavirus disease (COVID-19) pandemic” (ECLAC, 2020d) provides technical references on procedures for detecting and correcting possible bias. It is recommended that these processes be adopted by national statistical offices for all household surveys, even after the end of lockdowns.

National statistical offices that have access to different administrative records may use them to determine whether selected individuals (respondents and non-respondents) are included in their databases. Thus, the statistical models to control for non-response bias will be complemented more robustly. For example, information on whether an individual has received State assistance, or is affiliated to a social protection scheme, or has not been identified in the pension contribution base in recent months can be a powerful tool for detecting and eliminating potential selection bias in this type of mixed-mode survey.

Soon, it will be possible to study the differences between the characteristics of face-to-face and telephone respondents. For instance, before making corrections for non-response and calibrating sampling weights, countries with rotating surveys will be able to compare the expanded population structures of their respondents and identify possible biases. In some cases, it will be possible to include information from administrative records as auxiliary variables in the adjustment of weights. The fact that persons or households are included in administrative records is enough to differentiate them from other respondents, and sociodemographic patterns can be established to account for the differences between respondents and non-respondents. Many non-respondents could be people who did respond in face-to-face surveys. Correct identification of such patterns will help to preserve the valuable properties of the household survey estimators, the most important of which is a lack of bias.

B. Comparability and the impact of COVID-19

The pandemic forced most countries to change aspects of their information gathering and analysis methodologies, as summarized below:

- It changed the survey mode from face-to-face to telephone (or mixed-mode, in some cases), as well as the definitions of the eligibility structure of the selected dwellings and their corresponding disposition codes.
- The supervision scheme for interviewers was changed and, in some cases, map updates on the number of private households in the selected primary sampling units were eliminated.
- A new weighting scheme was introduced, seeking to eliminate coverage bias (not all households in previous surveys had telephone contact numbers) and non-response bias (some households contacted by telephone did not respond to the questionnaire).
- Calibration schemes were revisited and to increase the flexibility of estimation methodologies the number of calibration restrictions was limited.

These differences pose a challenge for national statistical offices, in terms of the comparability of the series generated during the exceptional circumstances of the pandemic. Considering that two events (telephone surveys and the pandemic) occurred at the same time, one because of the other, it is not simple to separate the change of mode (from face-to-face to telephone) from the effects and impacts on official figures reported by national statistical offices on the basis of household surveys.

However, there must not be a fall into the false dilemma of choosing between comparability or a lack of bias. The aforementioned differences make it clear that planning, surveying, execution and analysis in pre- and post-pandemic surveys are not comparable. The question is not whether methods are comparable, but whether the changes have had a significant impact on estimation of official statistics. The answer to this question will contribute to proper adjustment of the time series produced, using the results of each survey that national statistical offices carry out. Thus, it may be found —as a result of all the changes and adjustments made to collection and analysis of household surveys to eliminate bias— that there is no significant impact on the figures produced by an entity with respect to previous surveys.

Evidently, because of the pandemic, the magnitude of estimated social indicators changed. There are clear examples in the economically active population, the unemployed, people with reduced income, and people who have fallen into poverty. The comparability of the figures is not affected by these changes in magnitude, but by the possible over- or underestimation that collection mode changes and subsequent adjustments may have caused.

One sure way to separate the change in collection mode from the effect of the pandemic is to maintain telephone and mixed-mode surveys in the post-pandemic period for a reasonable time. This can be done in several ways. National statistical offices can perform separate parallel surveys with appropriate sample sizes. It is also possible to have a nationally representative telephone subsample, with larger sizes for face-to-face surveys as they require greater disaggregation. Another option could be to gradually include face-to-face surveying and simultaneously reduce telephone surveying.

In surveys that have rotation groups, different modes of collection could be considered for these groups. Assuming that each of the rotation groups is a representative sample of the country's situation and is distributed throughout the country, and that, with large samples, such as those in national household surveys, the sampling estimators have a normal distribution, it is possible to contrast hypotheses about the most important indicators in the rotation groups. This could be used to estimate the impact of the collection mode change on figures and determine subsequent actions to bring the figures produced during the pandemic into line with an indicator's time series. These actions will make it possible to determine, through statistical modelling, whether the change in collection mode affected the estimation of social indicators, and if it did, in what regard.

Ultimately, it is likely that the time series will not suffer from breaks, because of the adjustments to eliminate bias. If, because of parallel surveying in the return to normal, the change of mode is found to have had a significant effect on estimation of official figures, it would be advisable to adjust the time series retrospectively to make the measurements comparable. A further argument in favour of parallel surveying is that this splicing could not be done immediately after the pandemic; instead there should be a prudent wait, in order to correctly identify the components that had the greatest effect on estimation.

C. Conclusions

After several months of conducting household surveys by telephone to address movement restrictions put in place because of the pandemic, a new situation is emerging, in which face-to-face surveys are gradually being resumed. This document highlights some key considerations for this period, focusing on maintaining some of the procedures adopted, either temporarily or as part of the normal processes for household surveys.

In particular, the following recommendations are made:

- Consider a transition period during which telephone surveys continue, thus maintaining flexibility for the possibility of new lockdowns in some areas, and gradually reduce their use over time.
- Assess the possibility of continuing to use mixed-mode surveys in the pandemic and thoroughly study the feasibility of using these survey modes in the post-pandemic period, by redesigning surveys.



- Reconcile face-to-face and telephone questionnaires to maintain comparability when measuring constructs in both survey modes.
- Include processes for collecting contact information and basic household data during map updates, prior to completion of the questionnaire.
- Adapt the disposition codes of the selected dwellings and households, maintaining comparability between face-to-face and telephone surveys, reclassifying according to the reality of the selection from an area frame.
- Adopt the technical procedures for eliminating coverage and selection bias that were adopted while movement restrictions were in place as regular processes for household surveys.
- Evaluate, through statistical procedures, the effect of the change of mode on the taking of the surveys through parallel operations and, if necessary, readjust and smooth the time series.

Bibliography

AAPOR (American Association for Public Opinion Research) (2016), *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys* [online] https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf.

ECLAC (Economic Commission for Latin America and the Caribbean) (2020a), “Contingency plans and recommendations for statistical production in the face of the health emergency generated by COVID-19”, Knowledge Transfer Network (KTN) [online] <https://rtc-cea.cepal.org/en/conectados-rtc/covid-19-contingency-plans-of-the-national-statistical-offices>.

_____(2020b), “Nineteenth meeting of the Executive Committee of the Statistical Conference of the Americas of ECLAC”, August [online] <https://www.cepal.org/en/events/nineteenth-meeting-executive-committee-statistical-conference-americas-eclac>.

_____(2020c), “Recommendations for the publication of official statistics from household surveys in the context of the coronavirus disease (COVID-19) pandemic”, *COVID-19 Reports*, Santiago, April.

_____(2020d), “Recommendations for eliminating selection bias in household surveys during the coronavirus disease (COVID-19) pandemic”, *COVID-19 Reports*, Santiago, May.

This document is part of a series of reports prepared by the Economic Commission for Latin America and the Caribbean (ECLAC) on the evolution and effects of the COVID-19 pandemic in Latin America and the Caribbean. It was prepared by the Statistics Division, directed by Rolando Ocampo, under the general coordination of Alicia Bárcena, Executive Secretary of ECLAC.

Copyright © United Nations, 2020



Economic Commission for Latin America and the Caribbean (ECLAC)
Comisión Económica para América Latina y el Caribe (CEPAL)
www.eclac.org