

The pitfalls and potential of high-frequency phone surveys during COVID-19

Jeffery C Tanner

Phone surveys can be particularly useful in times – such as during the current pandemic – when it is difficult to conduct face-to-face surveys, but can present challenges.

The COVID-19 pandemic has torn through lives and livelihoods across the globe. Forcibly displaced people are among the most vulnerable but there is little robust data to provide insights into how their needs are evolving through the pandemic.¹ Health considerations limit options for gathering data during this challenging period, with traditional face-to-face data collection efforts halted out of concern for enumerator and public health, and because of travel restrictions, lockdowns and social distancing. One alternative is the use of mobile phone surveys. This approach has proven useful in gathering data in risky environments, in remote areas, in areas where enumerator safety is a concern, where responsiveness to new and changing data needs is important, and in contexts where monitoring unfolding situations is needed. They have also been found to yield high-quality data and to be cost-effective.²

In the face of the COVID-19 pandemic, the World Bank rolled out robust high-frequency (that is, multi-round) phone surveys (HFPS) to collect socioeconomic data in some 100 countries.³ In collaboration with the World Bank, UNHCR and national statistical offices, the World Bank–UNHCR Joint Data Center on Forced Displacement⁴ is supporting the integration of booster samples of displaced people in at least two rounds of ongoing and planned COVID-19 HFPS in 12 or so of those countries. This will enable a better understanding of changes in welfare, vulnerabilities and prospects for displaced people as a result of the COVID-19 pandemic, and enable better responses to be designed.⁵

Although HFPS can generate crucial data during the pandemic, there are challenges in applying this method, particularly in the context of displacement. The main

challenges can be classified as threats to the robustness of the sample and limitations to the information collected. Fortunately, with foresight and careful planning there are also ways to overcome those challenges.

Sampling

A statistically sound survey should be representative of a defined population and have a large enough sample to have the statistical power to detect meaningful differences between groups or across time. Because phone coverage is uneven – due to ownership (affected by demographic and economic factors), network coverage or electric power availability, for example – HFPS samples may not be representative of the target displaced population, thus skewing the analytical results. Similarly, bias can occur when some groups are more or less likely to be selected to be surveyed ('undercoverage' or 'differential coverage') or when some respondents refuse to participate ('non-response') or refuse to continue in subsequent rounds ('attrition').

Identifying specific sources of potential bias is critical in data collection exercises to maximise the ability to a) prevent bias through design, b) correct for it through cleaning, reweighting and analysis and c) identify the direction of bias in order to facilitate use of the data in policymaking. Integral to the value of survey data is the correct identification and implementation of a robust sampling frame, sampling strategy, and analytical weights – the primary tools for overcoming sample bias. If bias persists in HFPS data, it is most likely to be biased upwards, meaning that the data may indicate that the population is better off than it is. Even if this is so, HFPS data are still useful for policymaking if outcomes

observed in the data are low enough to indicate that intervention is merited.

In general, three approaches can be used to build a sampling frame: using a representative survey, using a list, and using random digit dialling.⁶ The first two rely on the existence of comprehensive information at a time not long prior to starting the phone survey. Some countries are fortunate to have had a national data collection effort that included collecting data and phone numbers on displaced populations, from which a representative sample can be drawn. In other instances, UNHCR registration data can be used, so long as the registry data – particularly phone numbers – are valid and current; however, unregistered populations would be missing. Typically, personal information may be used to help elicit a balanced sample but (to help mitigate data protection concerns) only the phone numbers need to be passed to the enumeration team. The third approach – random digit dialling (using randomly generated phone numbers) – is seldom cost-effective or feasible when surveying displaced persons but may be appropriate in certain cases.

All three of these methods are used in current JDC-supported survey activities. In Chad, a nationally representative survey from 2018 that included refugees for the first time is being used as a sampling frame. In Djibouti and Ethiopia, UNHCR registration data are being used. And in Ecuador the survey team will partner (as they have done previously) with a local telecommunications firm to determine the geographic areas and collections of phone numbers most likely to belong to Venezuelans in refugee-like situations in the country.

If appropriate complementary data are available, 're-weighting' can be used to help correct sample data to recover underrepresented parts of the true population. This process is complex and challenging but if properly applied can overcome issues of attrition, non-response and partial coverage to make the sample a useful approximation of the actual population.

Preventing a biased sample is of course better than correcting for one – and

incentives can help. The World Bank survey in Ecuador on the impact of COVID-19 on host and refugee populations provided airtime credit that met or exceeded the time spent on the survey. Outside an epidemic, there are other incentive options, such as direct provision of mobile phones or small solar charging stations.⁷

Instrument design and implementation

There are various practical considerations to bear in mind relating to potential limitations on what and how much can be collected.

Surveys need to be kept short to limit respondent fatigue, particularly when repeated interactions are planned, and this inevitably limits the potential depth and breadth of data gathering. Moreover, question complexity tends to reduce data quality, particularly in phone surveys. There are three options to mitigate these constraints to at least some degree. Different respondents can be randomly assigned a different module of questions while answers from a common set of modules are used to impute data for the randomly missing modules. Alternatively, modules that are asked of the entire sample can be rotated in and out from one round of the survey to the next, as was done in Ethiopia by including a module on locust swarms in the second round of that HFPS. This approach works best when more dynamic outcomes (such as food security and employment) are included in all rounds, and more static outcomes (for example, demographics or histories) are rotated. Finally, in some cases it is possible to use a smaller set of questions from a long module to impute an overall outcome score (as with consumption poverty).

Some topics are difficult to cover by phone survey. In general, topics that are challenging to survey face to face are even more difficult over the phone, although the impersonal nature of a phone survey may make it easier for some respondents to discuss certain topics (such as those involving social stigma like mental health, domestic abuse, xenophobia or sexual behaviour).

The evidence base on the reliability of phone surveys compared with face-to-face

surveys for displaced populations is still quite nascent. However, several studies have found no difference between mobile phone and landline surveys or mobile phone and face-to-face surveys.⁸ It is hoped that planned JDC-supported work in Jordan might be able to offer insights into this question, as the respondent sample is split between face-to-face and phone surveys.

Despite these challenges, phone surveys are generally better suited to collecting socioeconomic microdata than other remote data-collection options. SMS options may have slightly better coverage and initial compliance but questionnaires generally need to be limited to fewer than five questions to avoid respondent fatigue. Paper-based surveys pose significant logistical challenges in a setting without postal addresses or a functioning postal system. Internet-based surveys generally suffer worse sample bias than phone surveys as they require even more technology and developing a robust sampling frame is far more challenging. And all these options assume a fairly high functioning level of literacy, which is not an issue for HFPS.

There are several practical guides on implementing a phone survey – guides that cover topics including the paramount importance of data protection and collaboration with the national statistical office (NSO).⁹ Such collaboration may appear time-consuming during a crisis but is important in order to help build capacity,

encourage the inclusion of displaced populations in regular national data collection efforts, and build buy-in for the statistics generated to be used in designing a national COVID-19 response and recovery plan that includes displaced people. Phone surveys are not appropriate in all instances, but with a sound sampling strategy and instrument they can produce reliable data.

Jeffery C Tanner jtanner@worldbank.org
Senior Economist, World Bank–UNHCR Joint Data Center on Forced Displacement

1. World Bank (2020) 'Highly Vulnerable Yet Largely Invisible: Forcibly Displaced Populations in the COVID-19 Induced Recession' bit.ly/JDC-Paper-series-No1
2. Dabalen A, Etang A, Hoogeveen J, Mushi E, Schipper Y, von Engelhardt J (2016) *Mobile Phone Panel Surveys in Developing Countries: A Practical Guide for Microdata Collection*, World Bank <https://openknowledge.worldbank.org/handle/10986/24595>
3. See World Bank COVID-19 High Frequency Monitoring Dashboard bit.ly/WB-C19-HFDashboard
4. A World Bank–UNHCR joint initiative established 2019 www.jointdatacenter.org
5. Most surveys include modules on one or more of the following topics: health; education; financial access; food security; consumption (use of goods and services); labour market outcomes (e.g. employment status and sector, changes in labour income); WASH/health habits; COVID knowledge and behaviours, including compensating mechanisms; and aid/assistance. To that, the JDC has frequently added a module on displacement histories and, where possible, attitudes towards host/displaced populations.
6. Himelein K et al (2020) *High Frequency Mobile Phone Surveys of Households to Assess the Impacts of COVID-19: Guidelines on Sampling Design*, World Bank Group bit.ly/WorldBank-Himelein-et-al-2020
7. Etang A and Himelein K (2020) 'Monitoring the Ebola Crisis Using Mobile Phone Surveys', *Data Collection in Fragile States; Innovations from Africa and Beyond* <https://link.springer.com/book/10.1007%2F978-3-030-25120-8>
8. See endnote 2.
9. See for example UN ESCAP (2020) 'Engaging with development partners to stand-up COVID-19 rapid assessment surveys: Five tips for National Statistical Offices', *Stats Brief Issue 25* bit.ly/UNESCAP-stats-brief-25; UNICEF (2020) 'MICS Plus; A Step-By-Step Guide to Implementation' bit.ly/UNICEF-MICS-guide-2020; WFP (2017) 'Conducting Mobile Surveys Responsibly' bit.ly/WFP-mobile-surveys-2017



Datasets to be publicly available

JDC-supported datasets (and accompanying documentation) on displaced populations from roughly 12 countries will be de-identified and published alongside the corresponding data on national (host) populations. These will be publicly available in the World Bank microdata library with links in the UNHCR microdata library at <https://microdata.worldbank.org>.

Briefs on the trends found in these data will be used in dialogue with host governments to illustrate the challenges that these vulnerable populations face during the pandemic.

JDC Literature Review updates and database

The World Bank–UNHCR Joint Data Center on Forced Displacement (JDC) publishes monthly literature review updates. These are added to a literature review database and also to a compilation product, searchable under 13 themes. www.jointdatacenter.org/resources