Northwestern
GLOBAL POVERTY
RESEARCH LAB

Remote Surveying in a Pandemic: Research Synthesis

Executive Summary

As part of IPA's response to COVID-19, many existing and new data collections have shifted to remote data collection modes including computer-assisted telephone interviews (CATI), interactive voice response (IVR) and SMS surveys.

To support this effort, IPA examined existing evidence from remote surveys in low and middle-income countries, which is synthesized here. The source data, a categorized list of research papers, is available online. We invite clarifications and updates by email, including suggestions of relevant literature we may have missed.

Though there is extensive evidence to inform study and protocol design, it comes with several caveats, stemming from inconsistent reporting, lack of explicit testing, and limited evidence in some modes and topics. We identified four key lessons from the evidence that can be used to inform decision-making:

- **Sampling frames** have dramatic effects on research protocols, representativeness and response rates, as demonstrated by meaningful differences in the proportion of connected phone numbers by sampling frame. Existing samples connected with 63 percent of unique numbers attempted, whereas random digit dialing samples connected with 19 percent of unique numbers attempted.
- **Estimated costs** vary substantively by mode, with IVR being the cheapest mode and CATI the most costly. Using data from 27 studies and 18 countries, cost per completed survey ranged from $11.97 using CATI to $4.86 using IVR and $7.75 using SMS (all in 2020 USD).
- **Response rates**, which we report as the number of complete interviews divided by unique numbers attempted, are over 250 percent higher in CATI than IVR or SMS based on data from 41 studies and 20 countries. CATI surveys using existing samples averaged response rates of 56 percent compared to an average of 33 percent using RDD samples.

This document is part of a series reviewing existing evidence on implementing surveys using computer-assisted telephone interviewing (CATI) and other remote survey modes. It was made possible with the generous support from and collaboration with Northwestern University's Global Poverty Research Lab (GPRL). It was prepared by Savanna Henderson and Michael Rosenbaum with helpful input from Janina Roemer and additional review by Steven Glazerman, Doug Parkerson, and Shana Warren.

  - Promised monetary incentives meaningfully increase response rates: incentives of less than $1 increase response rates by five percentage points. There are steep diminishing marginal returns to incentive size.
  - Response rates over panel waves are highly context-dependent, though it appears that limiting the time window between surveys and collecting contact information for additional sim cards may help reduce attrition.
- Respondents to remote surveys are not **representative** when compared to a nationally representative survey, based on evidence from 15 studies and 11 countries. This is based on commonly-measured socio-demographic variables including gender, education, age, and urbanicity available from the evidence. Differences between response mode is not clear, nor are ways to make surveys more representative.

Inclusion Criteria for Literature

Papers are included in this review if they meet the following criteria:

- SMS, IVR or CATI was used to collect data for social science research.
- Research takes place in or is relevant for a low- or middle-income country (LMIC).
- Research was conducted after 2013.[1]
- Manuscripts reported information on costs, response rates, or representativeness.

As of June 2020, we prioritized evidence that can inform study designs viable during COVID-19. Some studies were excluded from response rate estimates based on their design. These studies provided respondents with a cell phone and/or solar charger during a face-to-face baseline survey, which is not possible in the COVID-19 context. This includes the World Bank's "Listening to" series of evaluations—which primarily provide mobile phones and SIM cards, as well as pre-notification of forthcoming calls, during baseline face-to-face surveys.

A note on study samples

Sampling frames for phone surveys can be either *new* ("cold calls") or from *pre-existing* lists. There are three primary approaches for developing a new sampling frame for phone surveys:

Unscreened Random Digit Dialing (RDD): Numbers are randomly generated in alignment with target countries' mobile number formations and subsequently contacted. While RDD tends to be representative of those with working phones, it inevitably generates a large proportion of ineligible numbers such as non-connected and non-residential numbers. This requires more call attempts for the same size sample. Figure 1 compares the average proportion of connected numbers per unique number attempted by sample type.

Screened RDD: Mobile Network Operator (MNO): A list of active numbers is procured through a Mobile Network Operator, generally through the services of a third party such as GeoPoll or Sample Solutions. These numbers are checked for activity by the telecom and are randomly sampled from the universe of telecom customers. For MNO to be representative, the sample must reflect the MNO's subscriber share, which depends on the

[1] Given recent changes to mobile penetration rates in LMICs, evidence prior to 2015 may no longer be representative. 2014 was included to ensure that the Ebola response was in our literature review. We made exceptions for Heath et al., 2017; Cantor, O'Hare & O'Conner, 2007; Singer & Ye 2013; James & Bolstein 1992; Edwards et al., 2002; Edwards, Dillman & Smyth, 2014; Dillon 2012; Ballivian, Azevedo & Durbin, 2015; Mahfoud et al., 2014; Francisco et al., 2011; Ferreira et al., 2011; Medway, 2012; Stecklov, Weinreb & Carletto, 2017, which began research earlier but provided information on protocols or results of A/B testing.t.

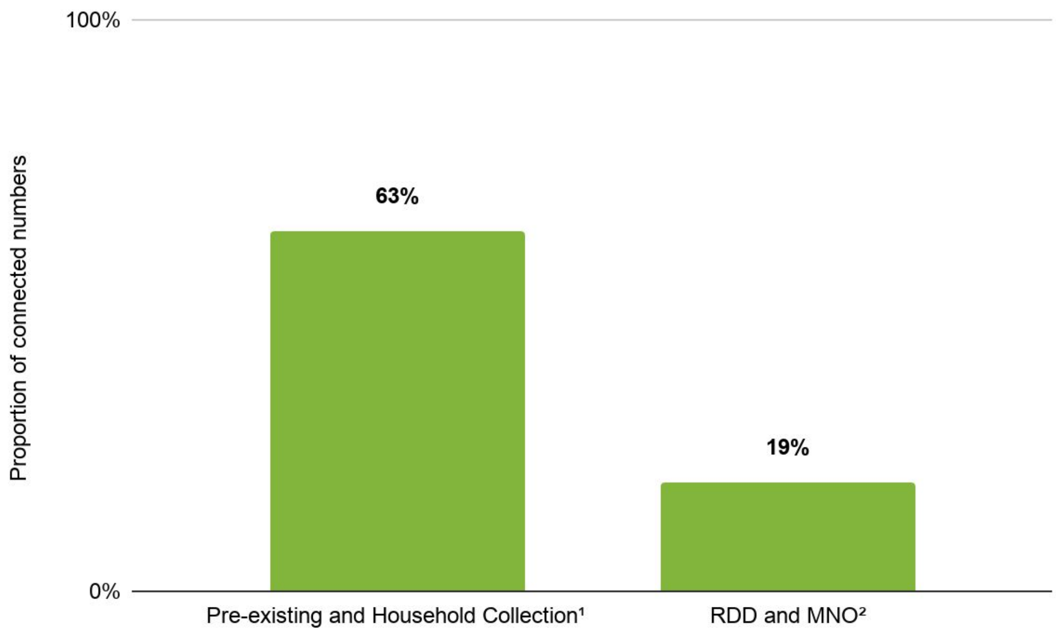
third party/vendor and their relationship with all national operators. Although number lists are supposed to be up-to-date with active numbers, slightly outdated lists may include disconnected numbers and more seriously, be missing new numbers, which could result in a sample that systematically under-represents some groups, notably migrants, younger individuals, and people who have recently experienced a positive income shock.

Household Collection: Numbers are collected during an in-person baseline survey. This first interview often collects demographic and baseline information and then invites the eligible respondent to participate in the panel or subsequent survey.

- During a global pandemic like COVID-19, face to face household data collection is not a viable option for developing a new sample but many existing panels that collected phone numbers in-person can be used as a sampling frame if the purpose to which the numbers are used is consistent with the consent originally provided by the respondent. To understand how this sampling frame will impact response rates, representativeness and cost, we differentiate sampling from *pre-existing* lists in the evidence review when possible.
- **Pre-existing**: A list of numbers are available from a previous survey or program. Representative surveys can be used to ensure that the sample is as representative as possible.[2] The representativeness and quality of a pre-existing list depends on the nature of the previous survey or program.

[2] See Kastelic et al., 2015, which used numbers collected during the Liberia Institute of Statistics & Geo-Information Services' Household Income and Expenditure Survey (2015).

Figure 1. Average proportion of connected numbers per unique number attempted by sample frame



Note: We limit this figure to studies that reported AAPOR disposition codes 4.30, 4.31 and/or 4.32 (Non-working/disconnected number, non-working number, disconnected number) with the exception of Australia (9), which was excluded as phone usage patterns for high-income countries may not apply to LMICs.

[1] *Lebanon (43); Liberia (28, 46).*

[2] *Bangladesh (19, 48); Ghana (31); Liberia (42); Mozambique (53); Philippines (4); Tanzania (48); Uganda (19); Zambia (5).*

Evidence on sample differences

The source of the sampling frame will influence a survey's cost, response rates, and representativeness. Evidence on the cost of sampling approaches and explicit testing of representativeness across samples in LMICs are somewhat sparse. Differences in sample characteristics make the choice between these sample types important and reported data may not be relevant to all contexts. In this document we explore the evidence on remote surveying modes and their various costs. Evidence on response rates by sample types is presented in a subsequent section. Evidence on representativeness is also presented in a subsequent section, but limited literature prevented any analysis of how sampling frames affect representativeness .

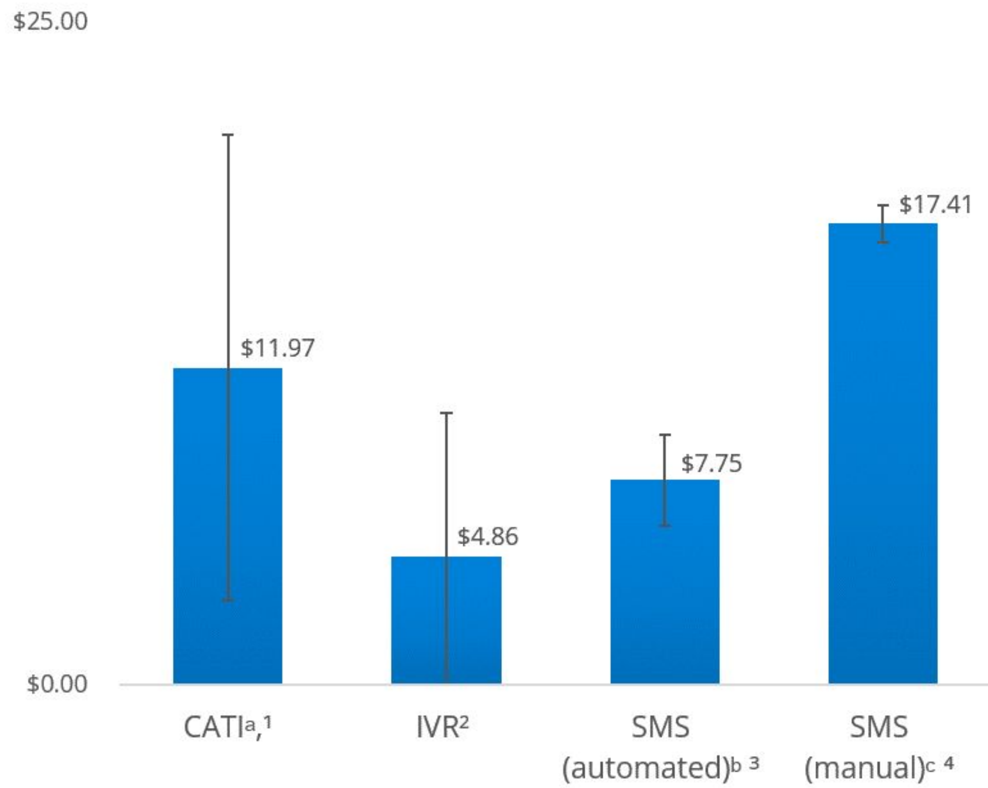
How much will a remote survey cost?

Self-administered modes—IVR and SMS—are cheaper to implement than CATI because they do not require the same personnel costs, (human interviewers and supervisors). Higher CATI costs may be partially offset by the higher response rate and potentially higher quality data. Differences in response rate require larger sample sizes and offset some of this difference. Response rates are discussed below. Lau et al. compares the costs of cross-sectional mobile phone survey modes, accounting for fixed and variable costs, between samples of varying sizes and find that for a survey with 3,000 completed responses, SMS costs 24 percent the cost of CATI, and IVR costs 43 percent the cost of CATI (Lau et al., 2019a).

Exact costs per completed survey depend on many characteristics such as sample size, survey duration, airtime costs and monetary incentives. Another cost could come from the provision of mobile phones and/or chargers to respondent households, which is recommended as good practice among populations with low phone ownership and/or access to ensure that a sample can be contacted and is representative (Dabalen et al., 2016).

Although many papers did not report costs or accounted for costs differently, we describe reported costs by survey mode in Figure 2, and present more details in the source data containing the source studies. We do not standardize these rates beyond converting to 2020 US dollars. The wide range of costs reflects the variation in approaches to implementing the same type of survey. SMS costs were separated in the figure below by approach (automated vs. manual), which significantly impacts cost. The estimated cost for a study that used interviewers to manually send messages instead of automating the messaging is based on interview time, including time to send texts and wait for the respondent to reply until all questions are answered, as well as the cost of text messaging itself (West, Ghimire & Axinn, 2015). All sample frames were included in calculating costs per completed survey. CATI costs primarily consisted of pre-existing and household collection samples, while IVR costs primarily consisted of RDD and MNO samples.

Figure 2: Average Cost per Completed Survey (2020 USD)



Note: Average survey cost by mode is displayed in this table with standard deviation displayed. Studies account costs differently and have different lengths and protocols. Studies reporting costs for multiple treatments are only cited once below.

[a] A portion of these estimates do not include fixed costs and underestimate total survey costs.

[b] Automated SMS utilizes technology that can schedule messages rather than rely on human interviewers to send messages to participants

[c] Manual SMS utilizes human interviewers to send messages to participants

[1] Kenya (32); Lebanon (43); Madagascar (7); Malawi (7, 49); Mozambique (53); Nepal (55); Peru (3); Senegal (7); Sierra Leone (42); South Africa (18); Tanzania (7); Togo (7).

[2] Afghanistan (40); Bangladesh (19); Ethiopia (40); Ghana (24, 31, 54); Mozambique (40); Peru (3); Sierra Leone (42); Uganda (19); Zimbabwe (40).

[3] Liberia (17); Peru (3).

[4] Nepal (55)

Cost considerations: Incentives

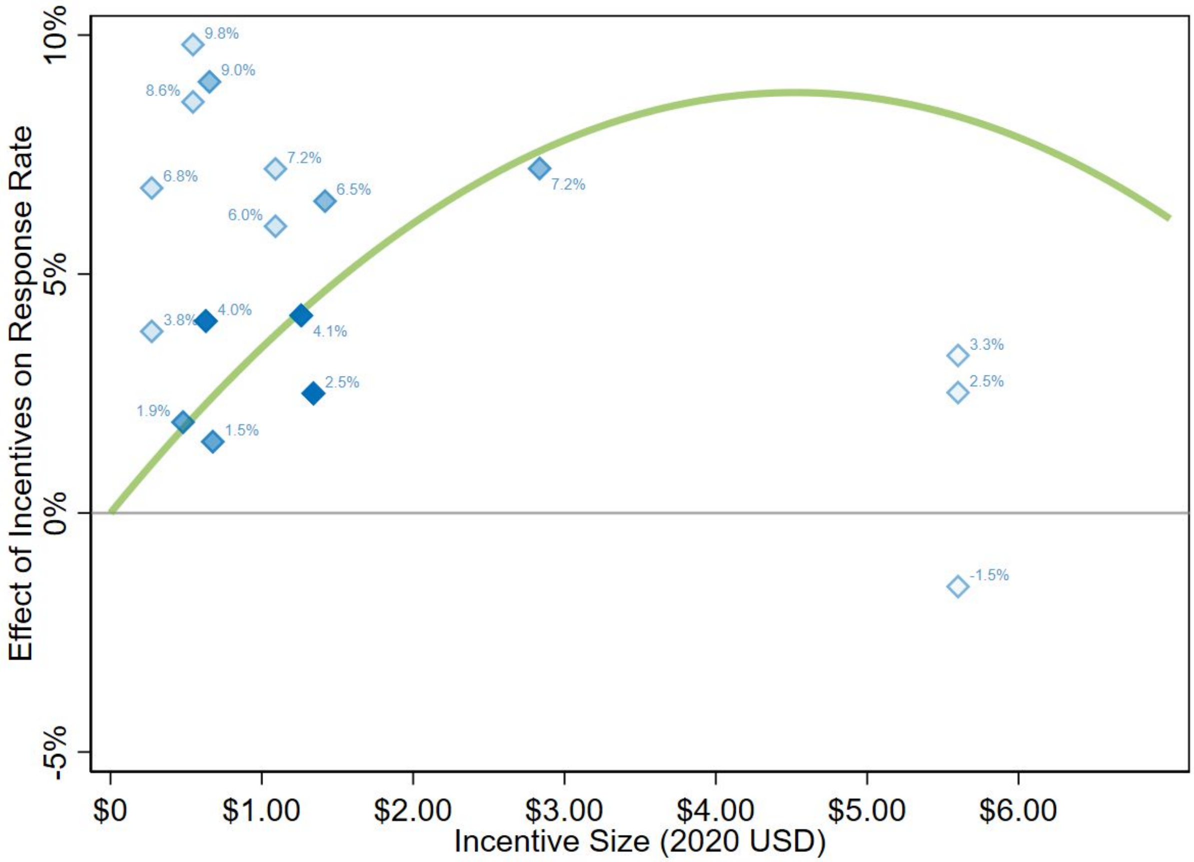
There is a rich literature on survey incentives and response rates in the developed world. These studies differentiate incentives to complete a survey based on the following motivations summarized by Singer & Ye, 2013:

Table 1: Behavioral Mechanisms to Promote Survey Completion

Motivation	**Behavioral Mechanisms**
Altruism	Responsibility; trust
Egoism	Reciprocity; payment
Survey characteristics	Inquisitiveness

There is some consensus that monetary incentives, the most widely tested, increase response rates by reducing refusal rate, but do so with diminishing returns as the size of incentives increases (Singer & Ye, 2013). These findings apply to the developing world, based on a number of experiments. Survey incentives, ranging in size from \$0.10 to \$3.00, increase response rate by around 5 percentage points, although this effect ranges from 2 to 10 percentage points depending on context and incentive size (Gibson et al., 2019; Lau et al., 2018b; Leo & Morello, 2016; Velthausz et al., 2016; Ballivan, Azevedo & Durbin, 2015). A few studies have explicitly tested the impact of financial incentives on various forms of response rates (partials, refusals, break-offs). We display the effect of this in 2020 USD in Figure 3 on completions per attempts:

Figure 3: Incentive Effects on Response Rate



Note: This displays a quadratic estimated curve in green of the response rate effect of incentives based on existing evidence around incentives in LMICs. Data from (Gibson et. al., 2019; Morello & Leo, 2016; Ballivan, Azevedo & Durbin, 2015; Leo et. al., 2015). Each point corresponds to a data collection (unique mode, country, and incentive), where the transparency corresponds to the sample size as a percentage of the maximum sample size of eligible phone numbers. No covariates are included in these regressions as not all studies had open data. Therefore, there may be discrepancies between effect sizes displayed in the text. Code to reproduce this analysis and figure is located [on Github](#).

There is some evidence, but a very limited amount, on the role that incentives play on the quality of responses. Incentives may not affect sample composition but may affect response quality by motivating strategic responses as well as decreasing satisfying behavior (Lau et al., 2018b; Stecklov, Weinreb & Carletto, 2017; Medway, 2012).

What response rates can be expected?

We report response rates as the number of complete interviews divided by unique numbers attempted. This provides more actionable information on how the study's measurement approach may affect response rate and costs.[3] To construct response rates for the reviewed literature, we collected American Association for Public Opinion Research (AAPOR) disposition code data when available.[4] These are linked in our source data. The AAPOR response rates authors reported are also included in the database.

By mode and sampling frame

Figure 4 displays response rates from the literature by mode and sampling frame. Raw response dispositions are reported in our source data. Surveys using CATI, the only interviewer-administered mode, report the highest response rates conditional on sample type. Despite the added costs of interviewers, the return in data quality as well as response rate may be cost effective. Self-administered modes such as IVR and SMS may have lower response rates compared to CATI due to respondents' difficulties with the technologies, such as SMS and manipulating the keypad, issues with providing responses in the requested format, or illiteracy.[5] Additionally, respondents who struggle to understand a question or how to respond may be systematically excluded. Strategies for clarification in the IVR context may exacerbate partial response rates; unanswered questions over IVR would be repeated three times before the system would hang up (Pariyo et al., 2019).

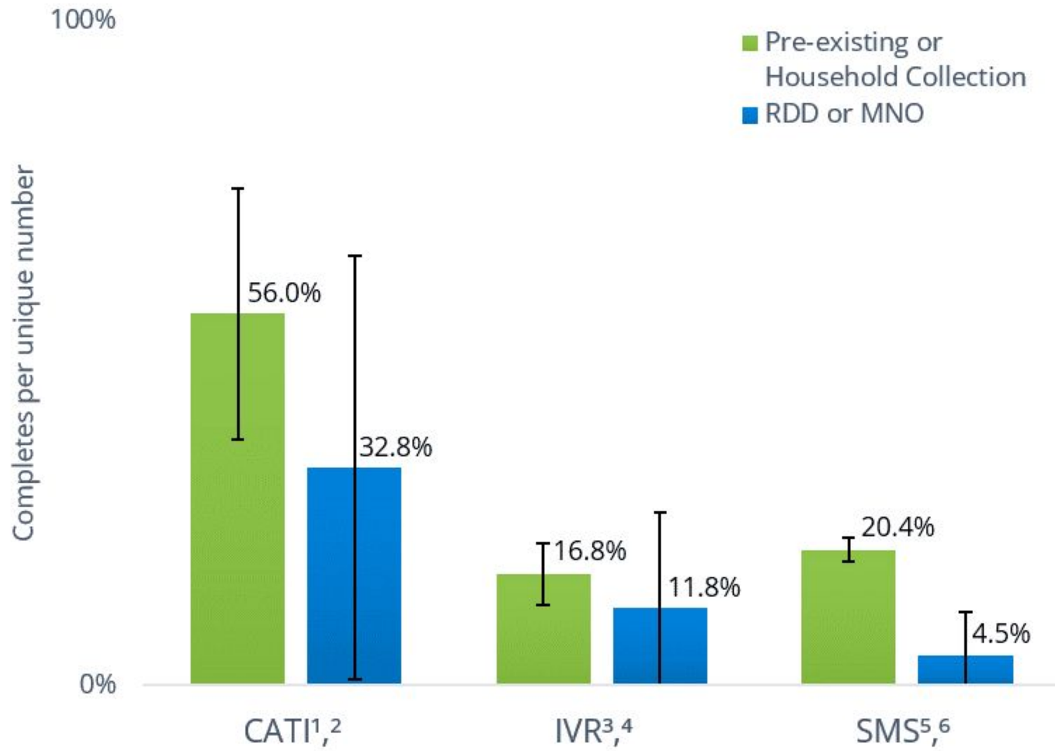
This evidence does not account for the COVID-19 context. It is not yet clear how COVID-19 will affect response rates or sample composition. For example, even if availability to respond increases due to local lockdown/stay-at-home orders, populations that only have electricity access in public spaces will likely have reduced phone access and usage.

[3] AAPOR outlines standardized disposition codes and outcome rates, including six response rate formulas, for surveys conducted by phone. Though there are variations to each of the six AAPOR response rates, they are broadly defined as the number of complete interviews divided by the number of eligible units in the sample. Notably, this does not include ineligible units, such as non-working or disconnected numbers, in the sample in the denominator, which may result in artificially high response rates when used for cost. This is relevant across sampling frames but especially pronounced in RDD, which generates a large proportion of non-working numbers.

[4] There are four categories of disposition codes: 1. Interview; 2. Eligible, Non-interview; 3. Unknown eligibility, Non-interview; 4. Not eligible.

[5] In Peru and Honduras (Gallup, 2012) SMS survey participants were required to enter the numeric character "1" for "yes" responses and "2" for"no" responses. Qualitative evidence finds that this process is a learned skill in the IVR context and may result in non-response or error (Daftary et. al., 2017).

Figure 4: Response Rate by Mode and Sample Type



Note: We report an unweighted average between study protocols (ex. a study that experimentally varied pre-notification contact would count as two observations) with standard deviation in parentheses. We do not control for various differences in study protocols that likely affect response rates. We exclude studies that do not provide exact numbers to calculate response rates. Studies reporting response rates for multiple treatments are only cited once below.

[1] *Ghana (23); Kenya (32); Lebanon (43); Liberia (28, 46); Peru (3); Sierra Leone (29); Turkey (47)*

[2] *Australia (9); Bangladesh (48); Liberia (42); Mozambique (53); Nigeria (37); Tanzania (48)*

[3] *India (30); Ghana (54); Malawi (1); Nigeria (1); Peru (3)*

[4] *Afghanistan (40); Bangladesh (19, 48); Ethiopia (40); Ghana (1, 24, 31); Liberia (42); Mozambique (40); Nigeria (37); Tanzania (48); Uganda (19); Zambia (5); Zimbabwe (40)*

[5] *Kenya (27); Peru (3)*

[6] *Ghana (34); Kenya (34); Nigeria (34, 37); Philippines (4); Uganda (34)*

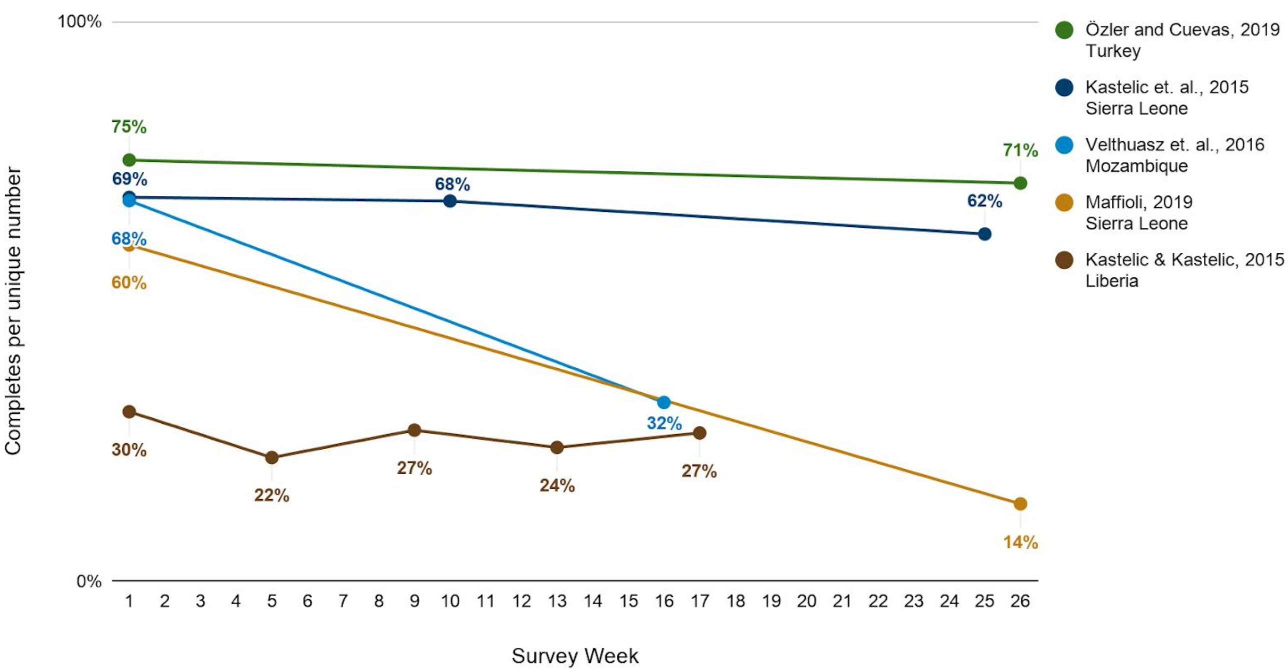
Across Panel Waves

Remote survey methods have the advantage of being cost-effective to conduct multiple waves, which may offset the disadvantage of having to use shorter questionnaires. This means that researchers need evidence on attrition from one wave to the next. High attrition reduces sample size and hence statistical power, and may introduce bias if there is differential (non-random) attrition with respect to an experiment's treatment arms. Mode, sampling frame, contact and survey protocols (which are explored in a separate section) likely have an impact on attrition, but evidence explicitly testing this is very limited.

Response rates across waves for CATI surveys are shown in Figure 5 below.[6] Although there is limited evidence, response rates changes, as well as qualitative evidence highlight the importance of context. Even among CATI surveys with relatively similar protocols—not providing phones, providing substantial incentives, using mobile numbers collected during in-person panels—baseline response rates and follow-up attrition vary substantively.

This context could be affected by COVID-19. The Sierra Leone and Liberia surveys that occurred during the Ebola crisis, with both countries in the midst of quarantine and lockdown measures, had very low attrition rates. Notably, Sierra Leone, in 2013, had response rates similar to Turkey in 2019, despite differences in mobile phone penetration and electrical infrastructure.[7] The other two studies lost about half or more of their sample to follow up. This is likely due to phone usage patterns in LMICs along with the large time window between baseline and follow-up, which was 14 weeks in Mozambique and 24 weeks in Sierra Leone (Maffioli, 2019). Multi-sim card ownership is common in LMICs (Li, 2019) with users switching between sim cards based on cost and connectivity advantages. In Liberia, 43 percent of numbers were permanently switched off and 28 percent were not ringing at the time of follow-up. In Mozambique, 39 percent of respondents noted they swap their sim cards every four months and during follow-up, 42 percent of numbers were unreachable. Limiting the time window between surveys and collecting contact information for additional sim cards may help reduce attrition.

Figure 5: Response Rates Across Panel Waves



[6] We did not include IVR and SMS response rates across waves due to the lack of studies available.
[7] High-income countries have increased mobile phone access and usage patterns than LMI countries that make comparisons on attrition suspect.

Note: Studies that provided phones to respondents were not included (Ballivan, Azevedo & Durbin, 2015 and *Heath et al., 2017). Sierra Leone (Kastelic et al., 2015) and Liberia (Kastelic et al., 2015) surveys occurred during the Ebola crisis; Phone use in Turkey (Özler and Cuevas, 2019) may not be representative of other LMI countries due to levels of cell phone infrastructure development and government. Velhausz et. al. and Maffioli both point to high rates of inactive numbers, changes in SIMS, and high costs of tracking as reasons for high attrition rates, although various protocol adjustments including incentives and SMS contact may be able to ameliorate some of this attrition (2013; 2019).*

Can response rates be improved with contact and survey protocols?

Mode and sample type are major determinants of response rates, but some studies have experimentally tested contact protocol choices such as pre-interview notification texts. Pre-survey SMS contact alone increased response rates by 3-8 percentage points (Kasy & Sautmann, 2019; Leo, Brian & Morello, 2016). When combined with a monetary incentive, pre-survey SMS contact increased response rates by 4-13 percentage points (Velthausz et al., 2016). Other studies have reported anecdotal and observational measures of various protocol changes (Özler & Cuevas). However, those approaches are not easily comparable between studies that survey different populations over time and do not detail salient aspects of protocols.

Questionnaire length and wording of questions likely impact response rates, but there is limited evidence explicitly testing survey protocols. Based on common practice in household surveys using mobile phones, Dabalen et al. suggest survey length should be restricted to 15-20 minutes and roughly 20 questions. Remote survey modes may affect how respondents comprehend individual items. Pariyo et al. describe this as a trade-off between survey time, cognitive burden of processing long questions, and clarity in wording. Interviewers may find it harder to deliver the survey instrument due to the remote mode's limitations in establishing rapport and providing visual feedback on comprehension.

Anecdotally, various study teams have suggested some study protocols that have increased response rate:

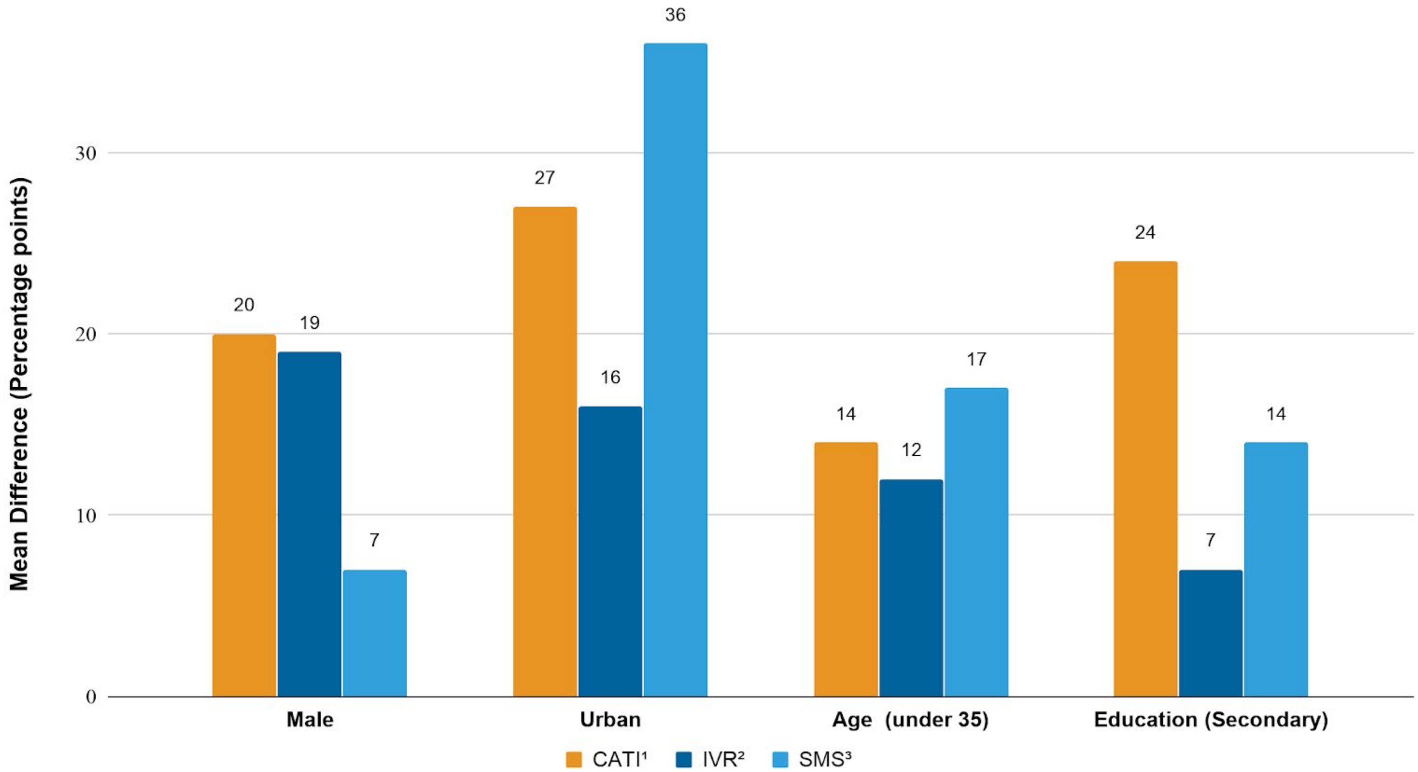
- Calling respondents who gave a "soft refusal" back with a supervisor to address questions about the legitimacy of the survey and convert refusals to "completes" (Suri, 2020)
- Offering an opportunity for respondents to reschedule the phone surveys
- Increasing number of call attempts, eligible times of day, and days of week for call attempts (Özler & Cuevas, 2019)
- Reducing survey length (Lau, Sanders & Lombaard, 2019)

Can remotely-collected survey data be representative?

At most, remote survey modes are representative of the population that owns or has access to a working, connected mobile phone. While LMICs on average have rising mobile penetration rates (ITU, 2018), there is still considerable variation in mobile phone access and use. The average phone owner in sub-Saharan Africa is more likely to be male, educated, and urban according to findings from Gallup surveys (Gallup, 2011). Multi-SIM card ownership is common in LMICs, due to variable network quality and price sensitivity, and family members or friends may share a SIM card (Li, 2019). The gender gap is also pronounced in LMICs: women are on average eight percent less likely to own a mobile phone than men, and use a smaller range of mobile phone services, such as SMS and internet-based uses, than men (GSMA, 2020).

Evidence in the developing world indicates representativeness is a significant obstacle for CATI surveys, and worse for IVR and SMS (Ballivan, Azevedo & Durbin, 2015; Gibson et al., 2019; Lau et al., 2019a; Leo et al., 2015). Determining the magnitude and cause of these differences is an empirical challenge. Some studies compare sample characteristics to "ground truth" from administrative data or a face-to-face survey, while others compare sample characteristics to those of DHS surveys. We report this information below by mode for commonly-measured socio-demographic variables: gender, urbanicity, age and education attainment. These, along with other variables less consistently measured, are all reported in our source data. As seen in Figure 6, all phone modes over-represent respondents who are urban, educated, young (below age 35), and male, compared to the benchmark.

Figure 6: Representativeness of Remote Surveys Compared to Nationally Representative Surveys



Note: This reflects the difference between remote surveys and nationally representative household surveys. Differences are reported here across four domains common in reported surveys. We limit this graph to studies that report any of these data with the exception of Australia (9), which was excluded as representativeness data for high-income countries may not apply to LMICs .

[1] *Liberia (28); Nigeria (37)*

[2] *Afghanistan (40); Ethiopia (40); Ghana (31); Mozambique (40); Nigeria (37); Zambia (5); Zimbabwe (40)*

[3] *Ghana (34); Kenya (34); Nigeria (34, 37); Philippines (4); Uganda (34)*

References

1. Amaya, Ashley, Charles Lau, Yaa Owusu-Amoah, and Jocelyn Light. "Evaluation of gaining cooperation methods for IVR surveys in low-and middle-income countries." *Survey Methods: Insights from the Field* (2018): 1-8.
2. Arthi, Vellore, Kathleen Beegle, Joachim De Weerdt, and Amparo Palacios-López. *Not your average job: measuring farm labor in Tanzania*. The World Bank, (2016).
3. Ballivian, A, João Pedro Azevedo and Will Durbin. "Using Mobile Phones for High- Frequency Data Collection." In *Mobile Research Methods: Opportunities and Challenges of Mobile Research Methodologies*, 21–39. London: Ubiquity Press, 2015.
4. Bloomberg Philanthropies Data for Health Initiative. "Executive Summary: Philippines NCD Mobile Phone Survey 2018." (2020).
5. Bloomberg Philanthropies Data for Health Initiative. "Executive Summary: Zambia NCD Mobile Phone Survey 2017." (2020). https://0df04920-08a1-4b70-a12c-25629cb5ffc4.filesusr.com/ugd/7a26dc_8d7efc17f3d649fd8f3db4c8fa7c0bcc.pdf
6. Cantor, David, Barbara C. O'Hare, and Kathleen S. O'Connor. "The use of monetary incentives to reduce nonresponse in random digit dial telephone surveys." *Advances in telephone survey methodology* (2008): 471-498.
7. Dabalen, Andrew, Alvin Etang, Johannes Hoogeveen, Elvis Mushi, Youdi Schipper, and Johannes von Engelhardt. *Mobile phone panel surveys in developing countries: a practical guide for microdata collection*. The World Bank, (2016).
8. Daftary, Amrita, Yael Hirsch-Moverman, Getnet M. Kassie, Zenebe Melaku, Tsigereda Gadisa, Suzue Saito, and Andrea A. Howard. "A qualitative evaluation of the acceptability of an interactive voice response system to enhance adherence to isoniazid preventive therapy among people living with HIV in Ethiopia." *AIDS and Behavior* 21, no. 11 (2017): 3057-3067.
9. Dal Grande, Eleonora, Catherine Ruth Chittleborough, Stefano Campostrini, Maureen Dollard, and Anne Winifred Taylor. "Pre-survey text messages (SMS) improve participation rate in an Australian mobile telephone survey: An experimental study." *PloS one* 11, no. 2 (2016).
10. Dillon, Brian. "Using mobile phones to collect panel data in developing countries." *Journal of international development* 24, no. 4 (2012): 518-527.
11. Edwards, Michelle L., Don A. Dillman, and Jolene D. Smyth. "An experimental test of the effects of survey sponsorship on internet and mail survey response." *Public Opinion Quarterly* 78, no. 3 (2014): 734-750.
12. Edwards, Phil, Ian Roberts, Mike Clarke, Carolyn DiGuiseppi, Sarah Pratap, Reinhard Wentz, and Irene Kwan. "Increasing response rates to postal questionnaires: systematic review." *Bmj* 324, no. 7347 (2002): 1183.
13. Feng, Shuo, Karen A. Grépin, and Rumi Chunara. "Tracking health seeking behavior during an Ebola outbreak via mobile phones and SMS." *NPJ digital medicine* 1, no. 1 (2018): 1-8.
14. Ferreira, Aline Dayrell, Cibele Comini César, Deborah Carvalho Malta, Amanda Cristina de Souza Andrade, Cynthia Graciane Carvalho Ramos, Fernando Augusto Proietti, Regina Tomie Ivata Bernal, and Waleska Teixeira Caiaffa. "Validity of data collected by telephone survey: a comparison of VIGITEL 2008 and'Saude em Beaga'survey." *Revista Brasileira de Epidemiologia* 14 (2011): 16-30.
15. Francisco, Priscila Maria Stolses Bergamo, Marilisa Berti de Azevedo Barros, Neuber José Segri, Maria Cecília Goi Porto Alves, Chester Luiz Galvão Cesar, and Deborah Carvalho Malta. "Comparison of estimates for the self-reported chronic conditions among household survey and telephone survey-Campinas (SP), Brazil." *Revista Brasileira de Epidemiologia* 14 (2011): 5-15.

16. Gallup. 2011. Mobile Phone Access Varies Widely in Sub-Saharan Africa. https://news.gallup.com/poll/149519/mobile-phone-access-varies-widely-sub-saharan-africa.aspx
17. Gallup. "The World Bank Listening to LAC (L2L) Pilot Final Report." (2012).
18. Garlick, Robert, Kate Orkin, and Simon Quinn. "Call me maybe: Experimental evidence on frequency and medium effects in microenterprise surveys." *The World Bank Economic Review* 34, no. 2 (2020): 418-443.
19. Gibson, Dustin G., Adaeze C. Wosu, George William Pariyo, Saifuddin Ahmed, Joseph Ali, Alain B. Labrique, Iqbal Ansary Khan, Elizeus Rutebemberwa, Meerjady Sabrina Flora, and Adnan A. Hyder. "Effect of airtime incentives on response and cooperation rates in non-communicable disease interactive voice response surveys: randomised controlled trials in Bangladesh and Uganda." *BMJ global health* 4, no. 5 (2019): e001604.
20. Gibson, Dustin G., Amanda Pereira, Brooke A. Farrenkopf, Alain B. Labrique, George W. Pariyo, and Adnan A. Hyder. "Mobile phone surveys for collecting population-level estimates in low-and middle-income countries: a literature review." *Journal of medical Internet research* 19, no. 5 (2017): e139.
21. Greenleaf, Abigail R., Dustin G. Gibson, Christelle Khattar, Alain B. Labrique, and George W. Pariyo. "Building the evidence base for remote data collection in low-and middle-income countries: comparing reliability and accuracy across survey modalities." *Journal of medical Internet research* 19, no. 5 (2017): e140.
22. GSMA. 2020. Connected Women: The Mobile Gender Gap Report 2020.
23. Heath, Rachel, Ghazala Mansuri, D. Sharma, B. Rijkers, and W. Seitz. "Measuring Employment: Experimental Evidence from Ghana." (2017).
24. Heward-Mills, Nii Lante and Yaa Amankwaa Owusu-Amoah. "USAID Communicate for Health: Mobile Cohort Study - Follow-up Survey 2018." USAID Communicate for Health. (2018).
25. James, Jeannine M., and Richard Bolstein. "Large monetary incentives and their effect on mail survey response rates." *Public Opinion Quarterly* 56, no. 4 (1992): 442-453.
26. Janevic, Mary R., Amparo C. Aruquipa Yujra, Nicolle Marinec, Juvenal Aguilar, James E. Aikens, Rosa Tarrazona, and John D. Piette. "Feasibility of an interactive voice response system for monitoring depressive symptoms in a lower-middle income Latin American country." *International journal of mental health systems* 10, no. 1 (2016): 59.
27. Johnson, Douglas. "Collecting data from mHealth users via SMS surveys: A case study in Kenya." *Survey Practice* 9, no. 1 (2016): 2824.
28. Kastelic, Kristen Himelein, and Jonathan G. Kastelic. "The socio-economic impacts of Ebola in Liberia: Results from a high frequency cell phone survey round five." *Washington, DC: The World Bank Group* (2015).
29. Kastelic, Kristen Himelein, Mauro Testaverde, Abubakarr Turay, and Samuel Turay. "The socio-economic impacts of Ebola in Sierra Leone: Results from a high frequency cell phone survey round three." (2015).
30. Kasy, Maximilian, and Anja Sautmann. "Adaptive treatment assignment in experiments for policy choice." (2019).
31. L'Engle, Kelly Eunice Sefa, Edward Akolgo Adimazoya, Emmanuel Yartey, Rachel Lenzi, Cindy Tarpo, Nii Lante Heward-Mills, Katherine Lew, and Yvonne Ampeh. "Survey research with a random digit dial national mobile phone sample in Ghana: methods and sample quality." *PloS one* 13, no. 1 (2018).
32. Lamanna, Christine, Kusum Hachhethu, Sabrina Chesterman, Gaurav Singhal, Beatrice Mwongela, Mary Ng'endo, Silvia Passeri et al. "Strengths and limitations of computer assisted telephone interviews (CATI) for nutrition data collection in rural Kenya." *PloS one* 14, no. 1 (2019).
33. Lau, Charles Q., Eric Johnson, Ashley Amaya, Patricia LeBaron, and Herschel Sanders. "High stakes, low resources: what mode (s) should youth employment training programs use to track

alumni? Evidence from South Africa." *Journal of International Development* 30, no. 7 (2018a): 1166-1185.

34. Lau, Charles Q., Ansie Lombaard, Melissa Baker, Joe Eyerman, and Lisa Thalji. "How representative are SMS surveys in Africa? Experimental evidence from four countries." *International Journal of Public Opinion Research* 31, no. 2 (2018b): 309-330.
35. Lau, Charles, and Nicolas di Tada. "Identifying Non-Working phone numbers for response rate calculations in Africa." *Surv Pract* 11 (2018c).
36. Lau, Charles, Ehud Gachugu, Eric Johnson, and Leenisha Marks. "Using SMS Technology to Survey Low-Income Youth: Lessons from a Vocational Education Tracking Study in Kenya." *Journal of International Development* 30, no. 6 (2018d): 1060-1063.
37. Lau, Charles Q., Alexandra Cronberg, Leenisha Marks, and Ashley Amaya. "In Search of the Optimal Mode for Mobile Phone Surveys in Developing Countries. A Comparison of IVR, SMS, and CATI in Nigeria." In *Survey Research Methods*, vol. 13, no. 3, pp. 305-318. (2019a).
38. Lau, Charles Q., Herschel Sanders, and Ansie Lombaard. "Questionnaire Design in Short Message Service (SMS) Surveys." *Field Methods* 31, no. 3 (2019b): 214-229.
39. Leo, Ben, and Robert Morello. "Practical Considerations with Using Mobile Phone Survey Incentives: Experiences in Ghana and Tanzania." *Center for Global Development Working Paper* 431 (2016).
40. Leo, Benjamin, Robert Morello, Jonathan Mellon, Tiago Peixoto, and Stephen T. Davenport. "Do mobile phone surveys work in poor countries?." *Center for Global Development Working Paper* 398 (2015).
41. Li, Tracey. UN Mobile Data Training Workshop: Mobile data analysis in low-income country settings. https://unstats.un.org/unsd/bigdata/conferences/2019/workshops/mobile-phone/day2/04%20MPD%20-%20Resource-scarce%20settings%20-%20Flowminder.pdf
42. Maffioli, Elisa M. "Relying Solely on Mobile Phone Technology: Sampling and Gathering Survey Data in Challenging Settings." (2019).
43. Mahfoud, Ziyad, Lilian Ghandour, Blanche Ghandour, Ali H. Mokdad, and Abla M. Sibai. "Cell phone and face-to-face interview responses in population-based surveys: how do they compare?" *Field methods* 27, no. 1 (2015): 39-54.
44. Medway, Rebecca. "Beyond response rates: The effect of prepaid incentives on measurement error." PhD diss., (2012).
45. Meuleman, Bart, Arnim Langer, and Annelies G. Blom. "Can incentive effects in web surveys be generalized to non-western countries? Conditional and unconditional cash incentives in a web survey of Ghanaian university students." *Social Science Computer Review* 36, no. 2 (2018): 231-250.
46. Morse, Ben, Karen A. Grépin, Robert A. Blair, and Lily Tsai. "Patterns of demand for non-Ebola health services during and after the Ebola outbreak: panel survey evidence from Monrovia, Liberia." *BMJ global health* 1, no. 1 (2016): e000007.
47. Özler, Berk, and P. Facundo Cuevas. "Reducing Attrition in Phone Surveys." World Bank Blogs, November 21, 2019. https://blogs.worldbank.org/impactevaluations/reducing-attrition-phone-surveys.
48. Pariyo, George W., Abigail R. Greenleaf, Dustin G. Gibson, Joseph Ali, Hannah Selig, Alain B. Labrique, Gulam Muhammed Al Kibria et al. "Does mobile phone survey method matter? Reliability of computer-assisted telephone interviews and interactive voice response non-communicable diseases risk factor surveys in low and middle income countries." *PloS one* 14, no. 4 (2019).
49. Pattnaik, Anooj, Diwakar Mohan, Sam Chipokosa, Sautso Wachepa, Hans Katengeza, Amos Misomali, and Melissa A. Marx. "Testing the validity and feasibility of using a mobile phone-based method to assess the strength of implementation of family planning programs in Malawi." *BMC health services research* 20, no. 1 (2020): 1-9.

50. Singer, Eleanor, and Cong Ye. "The use and effects of incentives in surveys." *The ANNALS of the American Academy of Political and Social Science* 645, no. 1 (2013): 112-141.
51. Stecklov, Guy, Alexander Weinreb, and Calogero Carletto. "Can incentives improve survey data quality in developing countries?: results from a field experiment in India." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 181, no. 4 (2018): 1033-1056.
52. Suri, Tavneet. 2020. Phone Surveys: Use and Adaptation to COVID-19. Adaptations for phone surveys: A Webinar with Tavneet Suri. J-PAL. https://www.povertyactionlab.org/blog/3-20-20/best-practices-conducting-phone-surveys
53. Velthausz, Daan, Rotafina Donco, Hannah Skelly and Marga Eichleay. “Mozambique Mobile Access and Usage Study: Computer-Assisted Telephone Interview (CATI) Survey Results.” United States Agency for International Development. (2016).
54. Vogel, Lara D., Levi Goertz, Suhuyini S. Shani, Mark Boots, Louis Dorval, and N. Ewen Wang. "A mobile-based healthcare utilization assessment in rural Ghana." *Procedia engineering* 159 (2016): 366-368.
55. West, Brady T., Dirgha Ghimire, and William G. Axinn. "Evaluating a modular design approach to collecting survey data using text messages." In *Survey research methods*, vol. 9, no. 2, p. 111. NIH Public Access, (2015).
56. World Bank. “Listening to Tajikistan - Household Survey: Background, Implementation, and Methods (English).” Washington, D.C. : World Bank Group. (2017). http://documents.worldbank.org/curated/en/624621538136672723/Listening-to-Tajikistan-Household-Survey-Background-Implementation-and-Methods