Azim Premji University

CSE Working Paper

#53

Telephone surveys for data collection - some reflections

Rosa Abraham and Mridhula Mohan

September 2023

Centre for Sustainable Employment

cse@azimpremjiuniversity.edu.in

Telephone surveys for data collection - some reflections[1]

Rosa Abraham[2] and Mridhula Mohan[3]

1. Introduction

The last few years have seen an upheaval in practices of data collection and survey methods. Even before the pandemic, several data collection endeavors had begun the transition to digital, computer-assisted, and tablet-based surveys. India's labor force surveys themselves had moved away from traditional paper-based surveys to computer-assisted PI techniques. The Covid-19 pandemic imposed a massive shock to these practices. Across several countries, ongoing surveys had to be prematurely terminated or put on hold in the interest of the safety of enumerators and interviewers.

However, as the pandemic progressed, data on the impacts of the pandemic at the economic, social, and health levels became crucial for researchers, policymakers, and the general public. The months during the lockdown, and subsequently saw a mushrooming of phone and web-based surveys conducted by independent agencies and national research organizations to address the urgent need for recent data on the impact of the economic lockdown across various dimensions. In India, multiple phone surveys emerged collecting crucial information on various aspects including the impact of the pandemic on employment, health, awareness about Covid-19, access to government schemes, the welfare of migrant workers, etc. (NCAER 2021, ActionAid 2020, (RCRC 2021; Dalberg 2020)

The transition to a phone survey raises several issues including access to respondents, ensuring proper administering of consent, monitoring of enumerators and data quality, ensuring the privacy of interviews and respondent safety when administering sensitive questions. These concerns may be broadly classified into those related to limitations in the reach and coverage of phone surveys and those related to the quality of data emerging from phone-based surveys. The former, we refer to as issues related to *sampling,* and the latter issues related to *reporting.* The most immediate *sampling* concern in the context of a phone survey is the coverage and reach of such surveys. The available sample is automatically restricted to those individuals who have access to a phone. This in itself introduces an implicit bias into the sample. However,

---

[1] The India Working Survey 2021 on which this paper is based is a collaborative project between researchers at IIM Bangalore, Azim Premji University and University of Western Australia. The project was supported by the Initiative for What Works to Advance Women and Girls in the Economy (IWWAGE), the Indian Institute of Management Bangalore (IIMB), and Azim Premji University.

[2] Assistant Professor, Azim Premji University

[3] Post-graduate scholar, University of Gottingen

even with access to a phone, there may be further constraints in participation. For many, an active connection may not be available despite owning a phone. Participating in a phone call with a near stranger for an extended period may not be an option for some owing to several restrictions including social norms that frown on conversations with outsiders (particularly for women), the burden of household responsibilities, and other time constraints. This is particularly the case when the survey has a particular intended respondent (say, female respondent) rather than any individual who answers the phone. The nature of phone surveys can also raise certain *reporting* issues since enumerators are not able to make actual observations (as in field surveys) or engage in a more detailed dialogue to verify a certain response. Phone surveys are also typically shorter and do not often allow for a free flow of conversation which may further restrict enumerators from asking more probing questions, including questions to aid recall or recovery-type questions. Finally, privacy, something which is often difficult even in field-based surveys particularly in developing countries, may not be possible and may not even be verifiable when doing a phone survey. The conditions of the pandemic are likely to exacerbate all if not some of these issues.

Both sampling and reporting errors are likely to disproportionately affect certain populations. Women, older people, and those from marginalized communities are more likely to not have access to a functional phone connection. The digital gender divide has been well-documented in India and elsewhere (GSMA 2019). Alvi *et al.* (2020) find that women were less likely to own a phone and even if they did, were less likely than men to have a working recharged connection. Although the gender gap in women's access to mobiles has narrowed globally in recent years (GSMA 2019), in India, there is still substantial variation in the ownership of mobiles between men and women, with only 59 percent of women owning mobile phones compared to 80 percent of men. Besides the issue of comparatively limited access to women over the phone, there are also issues of privacy (Alvi *et al.* 2020) particularly when asking sensitive questions (Saini and Gupta 2020), being able to provide their own responses to questions (rather than being prompted by someone) as well as the availability of 'free time' for women to participate in such surveys (Mathur 2020).

A detailed protocol can go a long way in reducing sampling and reporting errors. A protocol that envisages the possible scenarios that can arise when households are contacted via phone and accounts for these in enumerator responses minimizes confusion and streamlines processes, leaving little to the discretion of enumerators. It can also increase response rates consequently. Another tool for monitoring and assessing survey quality is to incorporate post-survey interviewer observations. These can provide some insight into the nature of the interview and whether there are any systematic differences across certain groups.

This paper investigates these issues in the context of a large phone-based survey, the India Working Survey (IWS), spanning a sample of individuals in two states in India. The survey was conducted to understand the impact of the pandemic and the subsequent lockdowns and mobility restrictions on the employment and earnings of individuals and households. We use the para data emerging from this data - namely call record data (call attempts, time of call, rescheduling requests, etc.), and post-survey interview questions on the nature of the survey to assess the potential extent of sampling and reporting errors.

The next section provides a brief overview of the India Working Survey. We then describe in detail the protocol that was adapted to reach our respondents. Section 3 describes the response rates, and in the context of the response rates, how the protocol worked in ensuring higher responses with a focus on differences by gender.

2. The IWS Protocol

The India Working Survey was conducted over the months of February and March 2020 in the states of Karnataka and Rajasthan. The objective of the survey was to capture the impact of social identity on labor market outcomes, including the measurement of work. The survey interviewed two randomly chosen male and female respondents from each household in two separate interviews. Although the survey was intended to be state representative, it had to be prematurely terminated owing to the pandemic. In the months of August and September, we re-contacted the respondents from the field survey over the phone. The intention of the phone survey was to understand the impact of the pandemic on employment, household's access to relief measures, and coping strategies adopted.

For the phone survey, we designed a detailed protocol to consider various possible scenarios and streamline the process for the enumerators. The design of the IWS protocol aims to minimize sampling and reporting issues. Therefore, the protocol includes as many scenarios as possible that could arise during a telephone survey. The workflow (shown in Figure 1) begins with the supervisor assigning respondents to enumerators for the day on the previous night. Following this, an automatic SMS is sent to these respondents informing them about the survey. The enumerator then calls the respondent and fills out the questionnaire (hereafter referred to as 'forms') depending on the response and in the end, marks the form as 'closed'.

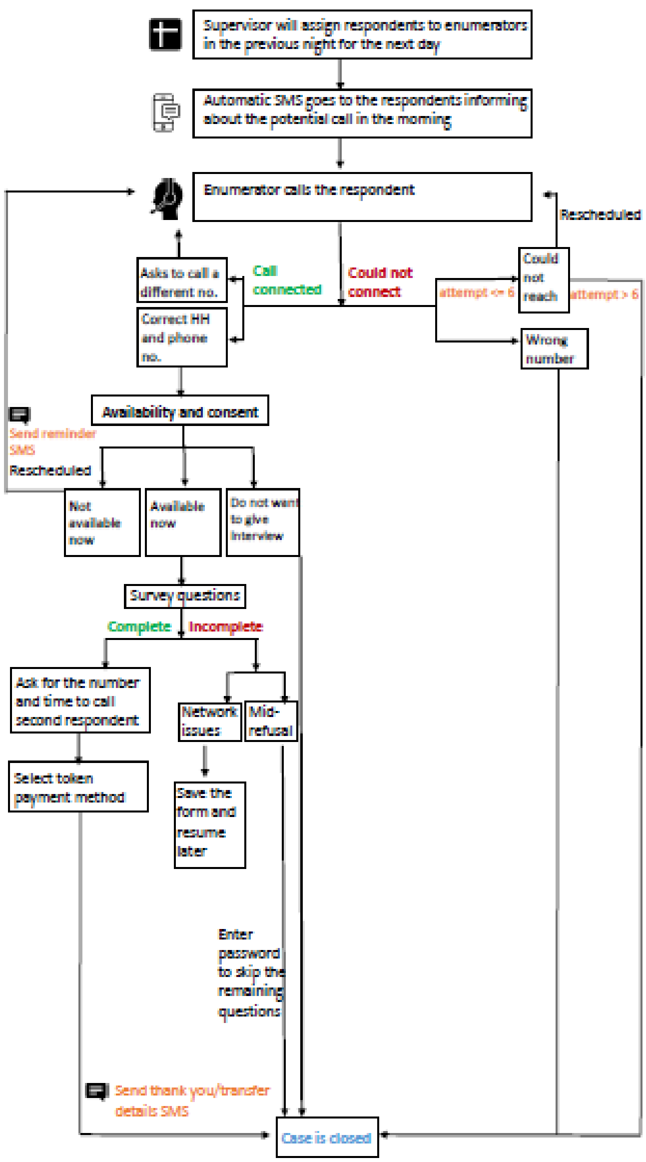


Figure 1: The IWS protocol

The leverage of the IWS protocol is that it maximizes the response rates and quality of data while keeping the process simple and easy for the enumerators. Therefore, we streamline processes and incorporate 'guides' for the enumerators, wherever possible. To begin, we pay particular attention to the allocation of respondents to be surveyed. Rather than giving the enumerators a complete list of respondents they need to call at once, we assign the supervisors to allocate a fixed number of respondents they should call each day. During the allocation, we also ensure that only female respondents are assigned to female enumerators and male respondents to male enumerators.

In contrast to face-to-face surveys, telephone surveys have an additional hurdle of getting the phone call connected. Sometimes respondents are missed even though they are temporarily unreachable because of network issues. To minimize such attrition, we also include such connection related scenarios to the protocol. For each of the different scenarios, we include instructions that appear on the screen to help the enumerators navigate the survey better. Among them, a scenario where it is possible to maximize reach is through reschedules. To ensure that we do not lose a potential respondent because their phone was unreachable, the survey is automatically rescheduled to four hours after the first call. If this automatically calculated time slot does not work for the enumerator, they are given the option to reschedule the survey to a date and time of their convenience. We also allow the enumerators or respondents themselves to reschedule a survey up to 6 times. To ensure that the enumerators remember to call the respondent, we send an automatic SMS reminder to the enumerators one hour before the rescheduled time. This makes it easier for the enumerator to keep track of the rescheduled calls easily. A separate SMS reminder is also sent to the respondent to ensure their availability.

Another scenario particular to telephone surveys is when enumerators are asked to call a different number to reach the respondent. For such cases, we allow the enumerators to add the new phone number directly to the form instead of making a note of it separately. Therefore, when the enumerator calls the respondent again, the call is made to this new number without having to copy it again.

The protocol also accounts for unexpected call drops mid-survey. In these cases, it is possible for the enumerators to save the form and resume it from the point they stopped in the previous attempt without having to repeat the completed questions again. A closely related scenario is if the respondent refuses to continue in the middle of the survey. Enumerators are given a password that they can use to skip the remaining questions and close the survey.

To help with survey monitoring, we also include a provision to add a supervisor to the call during a survey, time stamps, and questions for the enumerators at the end of the survey. In addition, specific information from each call that helps monitoring is published into a spreadsheet automatically and in real-time and shared with both the enumerators and supervisors.

3. Findings

3.1 Phone surveys and Sampling coverage

Sharing a phone number and consent to call back

The IWS field survey interviewed approximately 5,951 individuals and 3,646 households over the course of two months.[4] Of these completed interviews, 5171 agreed to share their landline or mobile phone number. The remaining either did not have a phone or did not want to share their number. There was a clear difference in the responses between men and women when asked to share their numbers. Among men, 89 percent were willing to provide mobile or landline numbers, about 5 percent refused to share and 6 percent did not have a number to share. For women, only 85 percent agreed to share, about 7 percent refused and 8 percent reported not having a mobile/landline. Several studies have pointed towards the lack of access to and ownership of mobile phones among women (GSMA 2019, Alvi *et al* 2020). We see this within our sample too, with a clear gendered difference in both ownership as well as the agreement to share a number.

To the 5,171 individuals who agreed to share their number, we also asked their permission to call back on the number provided in case of any follow up questions or clarifications. Almost all respondents consented to a callback, giving us a final sample of 5,117 individuals and we did not see any gender-specific differences here.

This final sample closely resembled the initial sample in terms of social composition - 56 percent were women, and 36 percent were from marginalized (SC or ST) communities. About 27 percent and 51 percent of men and women were illiterate. Only on this aspect was there some variation in the distribution between the IWS field sample and the phone survey consented sample. In the larger sample, about 29 and 54 percent of men and women were illiterate. Therefore, the phone survey sample is marginally more educated, which is not surprising given that access to the phone and reaching over the phone are likely to be correlated with higher incomes and education levels. Owing to the premature termination of the field survey due to Covid-19, the sample is predominantly rural (82 percent). Some of the numbers provided were incomplete or non-functional leaving us with a final sample of 4515 individuals.

Response rate

Of the 4,515 individual phone numbers we had, we were able to reach/contact nearly 75 percent of the numbers. For the remaining 25 percent, we either got a wrong number message or the number was not reachable. This could happen if the phone was out of coverage, if it was switched off. or not recharged. We did not see any difference between men and women in this aspect.

[4] We do not consider those households that were approached and refused to participate, as well as those that were not available (household locked, adult households members not present or available etc). While household refusal and non-availability does have consequences on the nature of the sample and introduce a selection bias, we do not consider this in this analysis.

In terms of response rate (where response rate is defined as the total successfully completed interviews as a share of all numbers available), within the sample of all numbers available, we were able to complete interviews with about 66 percent of the respondents. There is a marginal difference between men and women, with response rates higher among men, and a higher share of women refusing, (9.7 percent compared to 8.2 percent).

Table 1: Final interview status

		Men (%)	Women (%)	Overall (%)
1	Complete	66.4	65.1	65.7
2	Refused	8.2	9.7	9.0
3	Main respondent passed away	0.3	0.2	0.2
4	Respondent rescheduled	0.6	0.1	0.3
5	Wrong number	6.8	6.8	6.8
6	Could not reach. Clos	17.7	18.1	17.9
7	Total	100	100	100
8	N	*2005*	*2509*	*4514*
	Able to contact (1+2+3+4)	75.5	75.1	75.2

Source: Authors' calculations using IWS phone survey data

Getting access to women respondents

As we saw earlier, women were less likely to share a phone number compared to men. However, even with having a phone number to reach a woman at, there are still difficulties in having direct access to a woman via the phone. Indeed, this is apparent when we look at how many times the respondent themselves answered the phone when our enumerators called.

For male respondents, 71 percent of the times we called, our intended respondent, the man, answered the phone. However, when we contacted women on the phone number they had provided, there was only a 40 percent chance that it was the woman themselves who answered the phone. Clearly, there is a very stark difference in the likelihood of reaching the intended respondent when they were a woman, even when we had phone numbers provided by the women themselves.

The phone survey protocol anticipated this situation of the intended respondent not answering the phone and provided the enumerators with the appropriate responses in this scenario as described in the earlier section. In the event that the main respondent themselves did not answer the phone, we requested to speak with the main respondent, or if they were unavailable, rescheduled the interview. We also anticipated refusals at this stage, and this was also incorporated into the protocol. It could also arise that the respondent's number may have changed, and this was also added to the protocol. Table 2 lays out the scenarios that occurred when the main respondent did not answer the phone.

Table 2: Distribution of possible outcomes when the phone was answered

	Men (%)	Women (%)	Overall (%)
Main Respondent available	3.31	13.77	11.17
Main Respondent unavailable	73.62	71.43	71.98
Refuses to hand over the phone	6.87	6.72	6.76
Asked to call a different number	15.21	7.82	9.67
Total	100	100	100

Source: Authors' calculations using IWS phone survey data

Interestingly, even though women did not answer the phone call themselves, when requested to come to the phone, a relatively high share of women were available to speak. So, for about 14 percent of the calls placed to women, we were able to reach the main respondent. Therefore, although women were less likely to answer the phone, ultimately, there were no major differences in eventually *accessing* women over the phone compared to men.

The likelihood of the phone not being handed over to the intended respondent, as well as the chances of the main respondent not being available at the time the call was placed was similar for both men and women. And for men, there was a high share for whom their numbers had changed, and we were provided alternate numbers with which to contact them. For the majority of calls (72 percent), the respondent was unavailable at that time, and we had to request a reschedule. This was the case for both men and women. The IWS protocol allowed for six such rescheduling requests. And as we see later, this went a long way in maximizing our responses.

Once we reached our main respondent over the phone, we re-introduced ourselves wherever necessary, as well as spoke to them about the purpose of the interview and the approximate time it would take. We then went on to ask them if they were willing to participate in the survey. Notably, women, in 84 percent of instances[5], were willing to participate at the time of the call. For men, the corresponding number is lower and only 68 percent of cases agreed to participate at that time. A large share of men (28 percent) requested rescheduling of the call.

Rescheduling of calls and response rates

As detailed in the protocol, up to six attempts would be made to a given phone number. Any call made by the enumerator to the assigned number is counted as a call attempt, whether or not the call is connected. The only reason a call would be closed with less than six attempts being made is if the respondent or the person who answered the phone refused to participate or if the intended respondent was no longer available (had passed away).

Table 3: Final attempt number of completed calls

Call attempt at which interview completed	Men (%)	Women (%)	Overall (%)
1	44.1	29.5	36.05
2	26.9	28.89	27.99
3	13.3	17.38	15.55
4	7.74	11.08	9.58
5	4.58	6.79	5.8
6	3.38	6.36	5.02
	100	100	100

Source: Authors' calculations using IWS phone survey data

A successful call is one where the respondent agrees to participate and the interview is completed.
Table 3 shows the distribution of completed and successful calls by the call attempt at which they were completed. Among men, nearly half of calls (44 percent) were likely to be completed in the first attempt. For women, on the other hand, the first attempt yields comparatively lower response rates, with only 30 percent of calls being completed. Half of the women's completed calls occur only by the second attempt. Notably, for women, at every successive call attempt, there is a relatively higher share of completed calls compared to men. By the fifth attempt, up to 93 percent of the completed calls had occurred. A similar

[5] Here, we refer to every call attempt as a unique instance, rather than every respondent/phone number. So, if a respondent was called twice (the first call was rescheduled and the second one was completed), both these calls are considered here.

number for men occurred earlier at the fourth attempt. So, for women, the initial low response coming at the first attempt is made up for by subsequent call attempts to the same number. In fact, a large share of completed interviews occur at the sixth and final attempt. This highlights the need for a strict protocol and multiple re-attempts being made to the respective number to secure a response. This can ensure higher response rates, particularly among women.

In fact, we can trace what the typical responses are for every call attempt made. Table 4 shows the distribution of responses for each call attempt, for men and women separately. Two things are notable here. First, for men, completed interviews are highest in the first attempt, and subsequently taper off. This suggests that by connecting with men in the very first interaction itself, enumerators are able to secure a complete interview. For women, in contrast, the share of complete calls in the first attempt is low at 17 percent. However, this share picks up in the second attempt. And, in all subsequent attempts, the share of completed calls for women is higher than for men, (something seen in the earlier table as well). Therefore, while the initial response from women is not positive, it takes a few attempts to ensure a positive response and completed interviews from women.

Table 4: Distribution of interview outcomes by call attempt number

	Men					
	Attempt 1	Attempt 2	Attempt 3	Attempt 4	Attempt 5	Attempt 6
Completed	26.4	24.1	16.9	12.6	8.9	7.2
Refused	2.4	2.6	3.2	2.2	1.9	1.4
Respondent reschedule	13.3	8.9	6.8	4.8	3.8	2.1
Respondent passed away	0.1	0.0	0.0	0.1	0.4	0.2
Wrong number/alternate	6.8	3.4	2.7	1.6	1.9	1.0
Automatic reschedule	50.9	61.1	70.4	78.8	82.7	5.3
Could not reach, closed	0.0	0.0	0.0	0.0	0.0	82.9
Total	*100*	*100*	*100*	*100*	*100*	*100*
	Women					
Completed	17.1	22.1	18.3	15.1	11.3	10.9
Refused	2.2	3.1	2.8	1.9	2.1	3.1
Respondent reschedule	4.5	2.8	2.7	2.1	1.5	0.4
Respondent passed away	0.1	0.1	0.1	0.0	0.1	0.0
Wrong number/alternate	7.1	3.9	2.7	1.6	0.9	1.2

Automatic reschedule	69.1	68.0	73.5	79.4	83.3	12.8
Could not reach, closed	0.0	0.0	0.0	0.0	0.0	71.7
Total	*100*	*100*	*100*	*100*	*100*	*100*

Source: Authors' calculations using IWS phone survey data

Second, the share of automatic rescheduling is much higher for women compared to men. Automatic rescheduling happens when the call does not connect, and the enumerator has made less than six attempts. When this happens, the particular number remains in the enumerator's assigned cases, and they are required to recontact them at a later point. So, for women, the likelihood of the phone number not connecting at all was relatively high. In the first attempt, for women about 70 percent of calls do not connect and are automatically rescheduled. For men, the share was only 50 percent. In fact, for men, they were more likely to be contacted and then request a reschedule at a later date. As the attempts increase, and more and more viable phone numbers are removed from the enumerator databases, the share of this category steadily increases. By the sixth attempt, these cases are automatically closed.

Table 5: Final outcomes of rescheduled calls

	Final outcomes of every manually rescheduled call	Final outcomes of every automatically rescheduled call
Completed	60.5	49.09
Refused	10.06	7.76
Respondent rescheduled	2.04	0.27
Respondent passed away	0.15	0.18
Wrong number/alternate number provided	0.58	3.03
Automatic reschedule	3.35	4.31
Could not reach, closed.	23.32	35.36

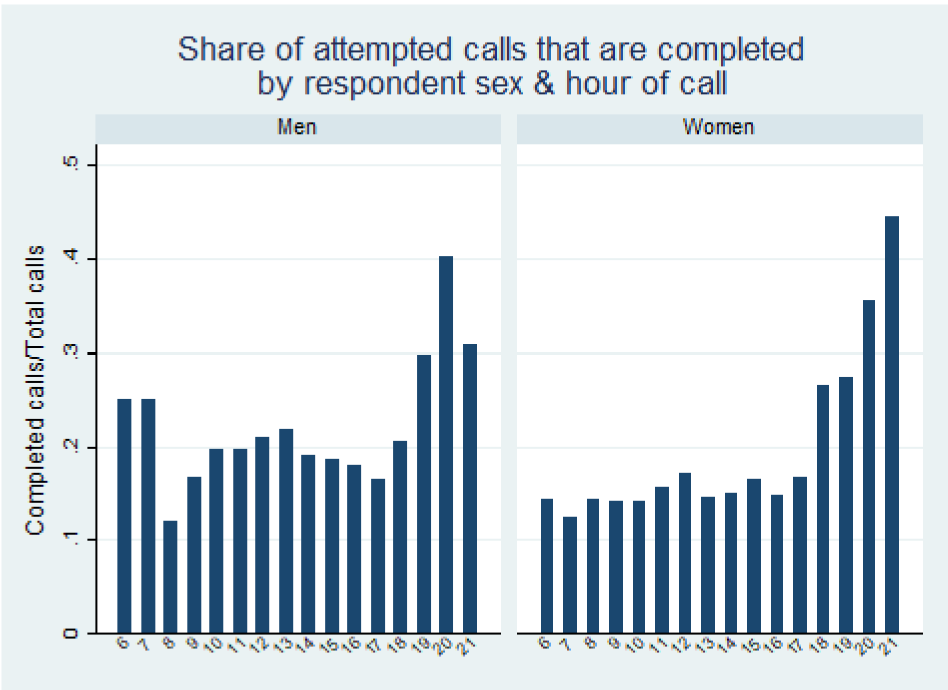
Source: Authors' calculations using IWS phone survey data

Notably, about 60 percent of respondent-requested rescheduled calls culminated in a completed interview. There was no significant difference between men and women. And, about 50 percent of automatically rescheduled interviews were eventually completed.

Response rates by the time of the call

Finally, we examine the response rate by time of call. While this was not used in our protocol, the information from this para data can be useful for planning future surveys and protocols. We are able to identify those times at which calls were more likely to be completed as well as what times respondents specified when they requested a rescheduling. This can give a better idea of respondent availability and how response rates in future surveys can be increased by timing calls effectively.

Figure 2: Response rate by time of call



Source: Authors' calculations using IWS phone survey data

As Figure 1 shows, for men and women response rates are highest in the latter half of the day. For men, there is also a time in the early morning when response rates are high, as well as some time mid-day when there is a small increase in response rates. For women, response rates are highest after 6 PM, and there is no other intermediate time during which their response rates are high. This likely points towards the lesser availability of women owing to work within and outside the home at all times of the day.

3.2 Phone surveys and Reporting of responses

At the end of a completed survey, after the call had ended, we asked our enumerators to make a series of observations on the nature of the participation of the respondent. Specifically, we asked them the following five questions:

Were the responses given by someone else on behalf of the respondent? (*proxy*)
Did the respondent put the call on speakerphone while the interview was going on? (*speaker*)
Did you hear any other individuals helping the respondent in understanding the questions? (*helped*)
Did you hear any other individuals prompting the respondent in their responses? (*prompting*)
Did he/she seem in a hurry to finish the interview? (*hurry*)
Did they seem interested and willing to provide answers? (*interested*)

Table 6 provides the share of instances where enumerators confirmed the occurrence of each of these outcomes, by gender of the respondent.

Table 6: Enumerator observations on survey quality by gender of the respondent

	Proxy	Speaker	Helped	Prompting	Hurry	Interested
Men	11	18	7	8	13	92
Women	13	23	20	20	17	89

Source: Authors' calculations using IWS phone survey data

We find that in the case of female respondents, responses were more likely to be given by proxy, 13 percent compared to 11 percent of men. However, in general, there are only a few instances where proxies are reported on behalf of the main respondent. Nearly a quarter of female interviews had the phone placed on speaker, much smaller than the 18 percent observed among male interviews. In the case of help in understanding questions as well as enumerators observing a second person prompting on responses, there is a substantial difference between male and female interviews. Women respondents were about three times more likely to receive help or be prompted in responses. Women respondents seemed more hurried in responding to questions perhaps as a result of the multiple demands on their time at home/work. There is not much difference between men and women in their perceived interest in participating in the survey.

The IWS survey did not ask any obviously sensitive questions to the female respondents (domestic violence, autonomy at home, etc.). However, even then, we see that female respondents are likely to not have privacy when participating in the phone survey. Therefore, phone surveys that seek to ask more sensitive questions that may put women at risk must heed extreme caution.

4. Conclusion/Discussion

Phone surveys were extremely valuable in understanding how the lockdown and pandemic were affecting individuals and data from these surveys have been crucial in informing public opinion and policy. However, there are serious issues of sampling, proper consent administration, privacy, and data quality when conducting such surveys. As we show, a detailed and well-thought-out protocol can go a long way in improving the reach and coverage of such surveys. Multiple rescheduling attempts and a streamlined protocol that does not leave much to enumerator discretion may be crucial in making sure that otherwise excluded individuals are brought into the extent possible. Para data from surveys can also provide valuable information on the quality of surveys. As we found, the phone was more likely to be put on speaker, the respondent was more likely to receive help in understanding and responding to questions, and women were more likely to be in a hurry and less interested in participating. Therefore, caution must be taken when relying on phone surveys both in the process of administration, the use of data, and the nature of questions being asked. As the ILO (2020) recommends, such, rapid (phone) surveys should not be used as a replacement for labor force surveys to produce estimates of key labor market indicators. In general phone surveys cannot substitute for field surveys. However, as we find, a good protocol and the use of para data can go a critical way in effectively reducing some of the typical problems associated with phone surveys.

References

Action Aid. 2020. 'Workers in the Time of Covid-19: Round II of the National Study on Informal Workers'. Action Aid.

Alvi, M., Gupta, S., Meinzen-Dick, R. Ringler, C. 2020. Phone Surveys to Understand Gendered impacts of Covid-19: A cautionary note.

Dalberg. 2020. 'Efficacy of Government Entitlements for Low-Income Families during Covid-19'. https://impactsofcovid.in/assets/front_assets/download/20.07.27_Efficacy%20 of%20government%20entitlements%20for%20low-income%20families%20during%20 Covid-19.pdf.

GSMA, 2019. Connected Women: The Mobile Gender Gap Report 2019. GSMA. https://www.gsma.com/mobilefordevelopment/wp-content/uploads/2019/02/GSMA-The-Mobile-Gender-Gap-Report-2019.pdf

Mathur, M. 2020. Reflections on reaching women over the phone in rural India. https://medium.com/idinsight-blog/reflections-on-reaching-women-over-the-phone-in-rural-india-29fffdbe296a

NCAER, 2021. https://www.ncaer.org/image/userfiles/file/DCVTS4/DCVTS4_Presentation.pdf

RCRC. 2021. 'RCRC Household Survey Round3'. https://www.rcrc.in/wp-content/ uploads/2021/03/RCRC_householdsurvey_round3.pdf. RCRC. 2020. "RCRC Household Survey Round 1." RCRC Centre for Monitoring Rural India. rcrc.in.

Saini, S., Gupta, S. Just a call away? Learnings from Phone Surveys in India and Nepal. IFPRI. Published on SurveyCTO blog: https://www.surveycto.com/case-studies/ifpri-covid19-phone-surveys/