Hello, can you hear me? Impact of speaker phones on responses in phone surveys during COVID-19

by Muzna Fatima Alvi, Shweta Gupta, Prapti Barooah, Claudia Ringler and Ruth Meinzen-Dick

Hello, can you hear me?

Impact of speaker phones on responses in phone surveys during COVID-19[1]

Muzna Fatima Alvi[2] , Shweta Gupta[2], Prapti Barooah[2], Claudia Ringler[3], Ruth Meinzen-Dick[3]

Abstract

Using phone survey data from India and Nepal, and propensity score matching techniques, we test the impact of speaker phone use on women's response to sensitive and non-sensitive survey questions. Our results show that 65% women in India and 47% women in Nepal had their phone on speaker during the survey and this impacted responses to questions on intra-household decision-making, causing women to under-report their agency over their own income and income of their husband/primary male earning member. Contrary to expected social norms, it did not impact the time reported on household care activities by women or their spouses. On objective questions of crisis coping strategies, we find no impact of speaker phone. The behavior of putting the phone on speaker can be partially explained by trust with the enumerator, as we find speakerphone use to be lower when women are matched with the same enumerators in the second round. There is significant heterogeneity in speaker use and its impact on responses. Our findings have important implication for design and analysis of phone survey data.

Keywords: respondent privacy, response bias, women, phone surveys, speaker phone, South Asia

[1] This research is supported by The German Federal Ministry for Economic Cooperation and Development in India, and United States Agency for International Development (USAID) in Nepal. We are thankful for the help and support provided by SEWA sisters, including our respondents. We received survey support from the team at All India Disaster Mitigation Institute-Ahmedabad in India, and Institute for Integrated Development Studies in Nepal. Comments on earlier drafts and concept notes by Ruth Meinzen-Dick and Agnes Quisumbing are gratefully acknowledged. All remaining errors are our own.

[2] International Food Policy Research Institute, New Delhi, India
[3] International Food Policy Research Institute, Washington D. C., U.S.A.

Introduction

As COVID-19 continues to spread globally, the reliance on phone surveys is expected to grow as well. It is imperative to know potential ways by which responses and data quality can be compromised in phone surveys. How survey responses are impacted due to the presence of third parties, has been investigated extensively, but evidence on the direction of impact is mixed, based both on type of third party, and type of information. However, little to no evidence exists on the various channels by which respondent privacy can be violated in surveys. This question becomes much more important in phone surveys, where unlike face to face surveys, surveyors can't confirm respondent privacy a priori. Conventional ways of confirming respondent's privacy over the phone – asking directly or through background noise- may not be sufficient, and may even jeopardize the safety of respondents, especially vulnerable groups like women and children. It is also important to consider how privacy concerns for women might be reinforced by persistent gender gap in mobile phone ownership, which is especially acute in South Asia.

The most ostensible drawback of phone surveys compared to in-person surveys comes from lack of physical interaction between the surveyor and respondent. Facial expressions, ease in explaining and conducting surveys, and comfort due to familiarity between the surveyor and respondent (Blaydes & Gillum, 2013) play an important role in improving the quality of information collected, which is absent in phone interviews. These factors could also play a role in determining respondent privacy, as trust and rapport building may reduce intrusion by third parties.

Data reliability concerns are heightened when surveys cover potentially sensitive questions. Such questions suffer from problems such as social desirability bias, which could distort survey estimates (Krumpal, 2011), with the extent of sensitivity recognized as a major factor influencing this bias (Stocké, 2004; Kryson, 1998). Even the most innocuous incident or

activity could potentially act as a source of response bias. For instance, studies find that the mere presence of a white foreigner could impact behavior in experiments (Cilliers at al., 2015) among predominantly black respondents in Africa. More recently, studies have highlighted the role of cultural differences and geographical variations in respondent privacy during the interview (Diop et al, 2015; Mneimneh et al, 2015).

In light of these challenges, our paper demonstrates a possible channel by which respondent privacy can be breached in phone surveys, and how it impacts response to questions on intra-household decision making. We conducted phone surveys of self-employed women in Gujarat, India and women farmers in Dang, Nepal to study the socio-economic impacts of COVID-19 on livelihoods. We find that a significant proportion of women were answering the survey while on speakerphone in both countries. Using linear regression and propensity score matching, we show that speaker phone biases answers to questions that can be perceived as sensitive, such as intra-household decision making over own and spousal income, but doesn't impact objective questions such as household coping strategies to deal with income loss. We also find strong geographic differences in both use and impacts of speaker use. Our paper contributes to the growing body of literature on best practices for data collection, especially in the context of remote data collection and speakerphone use.

Background

There are many factors that govern whether presence of third party will impact survey responses or not and how. The first, is the relationship of the bystanders with the respondent, and if questions asked are about, or related to, the bystander. For instance, parental presence makes youth & young adults less likely to report substance abuse (Aquilino, 1997; Aquilino et al, 2000; Gfroerer, 1985; Hoyt & Chaloupka, 1994; Moskowitz, 2004). Evidence on spousal presence is mixed. Casterline & Chidambaram (1984) and, Pollner & Adams (1994) confirm

impact of spousal presence on reported substance abuse, but Aquilino (1997) finds no such impacts. Anderson & Silver (1987) confirm no impact of partner presence on subjective questions like standard of living among Soviet couples. One must also consider if the respondent may face repercussions from third party from disclosure of key information (Edwards et al (1998); Aquilino et al. 2000). Familiar bystanders could lead to larger impact on responses as opposed to strangers (Mavletova & Couper, 2013).

Moreover, these impacts are larger for sensitive questions than those that collect neutral information (Krumpal, 2011; Mneimneh et al, 2015; Krysan, 1998). Much of it is governed by the social desirability bias which is the tendency to give socially desirable responses instead of those that are reflective of true feelings (Grim, 2010). Respondents' willingness to report their answers accurately and honestly is influenced by the perceived privacy of the survey setting, the perceived legitimacy of the survey and rapport between the interviewer and respondent (Holbrook et al.2003; Stocke 2007).

In situations when alternative modes are not feasible, it is necessary, but also more challenging, to find ways to establish respondent privacy. For instance, in face-to-face interviews, one can directly observe if respondent is surrounded by someone and not ask sensitive questions. However, third party presence in phone surveys can impact survey responses to the extent that questions are overheard by bystanders. Various techniques have been deployed to overcome this, such as item count technique (Tsuchiya et al, 2007), however, their applicability might be limited in settings where respondent is not literate or takes time to understand the methods.

An important role is played by the circumstances in which phone survey is conducted. Lynn and Kaminska (2011) find poor line quality, multitasking during a survey, and distractions due to active or passive involvement of others, can affect answers, and this is more likely in mobile phone surveys. There is little to no evidence on potential ways by which respondent's

privacy can be breached in mobile phone surveys. Some evidence comes from Brick et al, (2007) where 5% of respondents report being in someone else's home and from Lynn & Kaminska (2011) where 2% respondents felt that their responses were affected by bystanders. A few discuss use of caller ID and answering machines to screen calls on survey participation (Kempf & Remington, 2007; Roth et al 2001), however the use of speakerphone is not studied. The use of speaker phone while talking over phone is fairly common, and one might use it due to network issues, hearing disability and even while multi-tasking. All these issues related to privacy and bias are thus likely to be present in mobile phone survey data as well.

Data and methods

Our data comes from multi-round phone surveys conducted in India and Nepal. In India, we conducted phone surveys of women who were members of SEWA[4] (Self Employed Women's Association), a trade union of women that is organized as a collection of smaller self-help groups (SHGs). The survey was conducted in Gujarat, a state on the western coast of India. SEWA members are predominantly poor, self-employed women workers engaged in various forms of employment in the unorganized sector.

We selected respondents from among SEWA membership lists, and SEWA staff contacted them in advance, telling them to expect our call. Out of a list of 930 respondents obtained from SEWA, 627 respondents were surveyed from both rural and urban areas in nine districts in round 1 (Figure 1).

In Nepal, phone surveys of women who were maize farmers were conducted in 4 municipalities of Dang district- Dangisharan, Lamahi, Rapti and Shantinagar. Dang is located in Lumbini province in the Nepal Terai region (Figure 2). These farmers had been previously

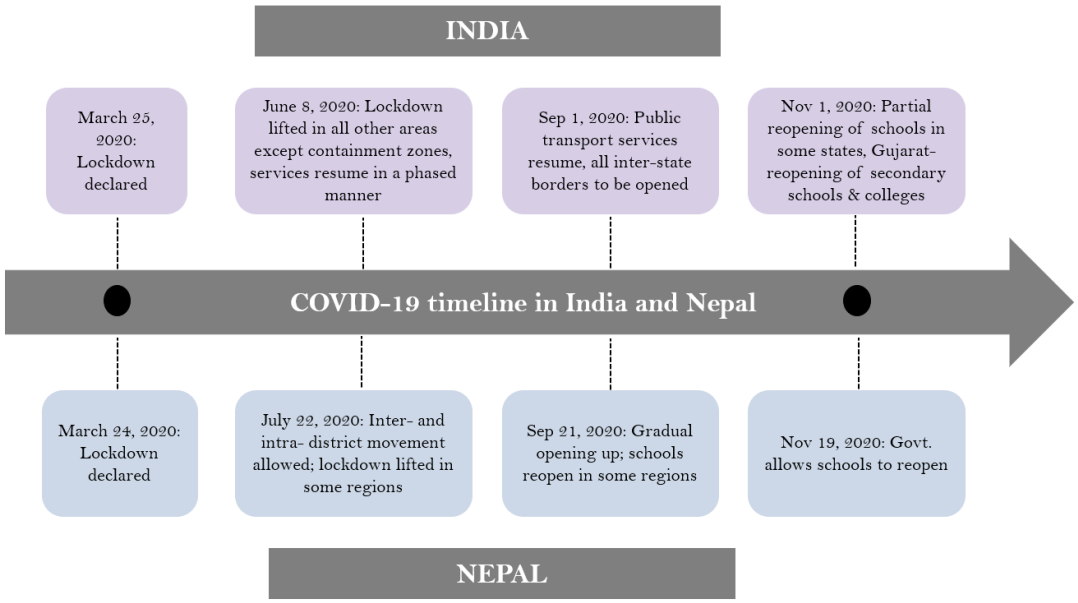
[4] https://www.sewa.org/

contacted during a household listing exercise conducted before the Covid-19 pandemic began, for a pest management experiment. Out of the record of 695 phone numbers, 534 women were surveyed. In order to track the short to medium term impacts of the pandemic and to trace the evolution of household resilience strategies, 5 rounds of surveys with the same set of respondents had been planned between May 2020 to July 2021. In round 2, the number of participants in Gujarat declined to 567 and those in Dang declined to 490 due to attrition. This paper uses data from the 1st and 2nd round of surveys from both regions.

Figure 1: Sampled districts in Gujarat

Figure 2: Sampled district in Nepal

Figure 3: Timeline for surveys conducted in Gujarat and Dang



The respondents were asked questions about their demographic profile, primary occupation, household asset position, questions on intra-household decision making on earnings, sources of extension used by farmers, impact of the pandemic on household's income, access to food and water, mobility, migration and questions on conflict and disagreement between the respondent and the primary male decision maker of the household (See Appendix for questionnaires used in both rounds).

The survey team was made up almost entirely of female enumerators. In Gujarat, many enumerators were previously affiliated with SEWA while in Dang, the respondents were already familiar with enumerators due to prior household listing exercise, which. helped build rapport with the respondents in both countries. Additionally, as enumerators in Gujarat were familiar with SEWA livelihoods and local procedures, they were also able to provide valuable insights that improved the articulation of survey questions.

Verbal consent was obtained from all participants before conducting their surveys. In Gujarat, all participants were provided with a food kit of 300 INR value and in Dang, phone credit of 100 NPR was provided to all study participants for each round. A team of 30 enumerators in

Gujarat and 5 enumerators in Dang was trained virtually in early May 2020 and June 2020 respectively and the surveys were conducted using the Computer Assisted Telephone Interview (CATI) platform of SurveyCTO.

Speaker phone to detect respondent's privacy

One way to confirm privacy in phone surveys, is by directly asking the respondent if she is around any person who can hear the responses. She can be then requested to switch to a private place. Asking the respondent if she is alone is in itself a sensitive question. It raises unnecessary suspicion by the respondent, and anyone listening to the conversation. The respondent may then be asked, or forced, to remain in a place where their answers can be monitored to avoid any risk or conflict with other members of the household. Even when respondents say they are in a private space, that may not be accurate. This raises questions about the credibility and reliability of information that is sensitive in nature or which the respondents may be uncomfortable to disclose unless they are alone.

In Gujarat, we had been forewarned by SEWA that in recruiting respondents for the survey, they noticed that several women were answering the phone while on speaker, sometimes on their own volition, but often as a result of a request by spouses or in-laws. This was important information to know beforehand, to ensure that sensitive questions on intra-household conflict were not asked to those who put their phone on speaker during the interview. We thus included questions on speaker use in all subsequent phone surveys with women. To avoid conflict, we asked the respondent indirectly if her phone was on speaker while the survey was being conducted. The question was worded as:

> *"I'm having a bit of trouble with the phone connection. Is the phone on speaker on your side? It's not a problem if it is, but I just need to adjust a setting if it is."*

The question was asked in the beginning of the survey and again towards the end before the module on intra-household conflict. If the respondent said that her phone was on speaker

anytime during the survey, she was not asked the questions on household conflict. No such filter was applied to any other questions before asking them.

A major problem in extracting sensitive information is ensuring the truthfulness with which the respondent answers questions on such issues. One needs to be careful with the wording of the question to not cause any discomfort to the respondent, irrespective of whether she is alone or not. The possibility of being surrounded by others poses additional risk to the respondent of exacerbating disagreements and conflict within the household. Eliminating sensitive questions when speaker is in use and testing and accounting for potential response bias in data analysis is thus essential.

Sensitive and non-sensitive questions

In this paper, we test the impact of speaker phone on various questions. Questions on conflict were already eliminated if the phone was on speaker, however we test the impact of speaker use on responses to other questions which were not considered sensitive a priori. The respondents were asked about who decides how to spend their earnings, who decides how to spend the earnings of spouse/primary male (henceforth spouse) and to compare their earnings with those of their spouse. Similarly, questions were asked about how the respondent and her spouse divide activities of caring for household members. Involvement of women in household decisions, especially those related to income can be used to measure women's agency and to the extent that revealing agency is `risky' for women, responses to these questions might be influenced by respondent privacy. Social norms indicate that women do more household work than men, therefore, any kind of social desirability bias might make women overreport the time spent on care activities by them than by men. However, we can't predict how respondent privacy can play a role in responding to these questions. We thus hypothesize that use of speakerphone will impact the responses to these questions. However, we don't make any assumptions about the direction in which these impacts will be manifested.

We also test the impact of speakerphone on questions assumed to be non-sensitive. Respondents were asked questions about their coping strategies to deal with income loss due to COVID-19- use of savings, sale of assets, borrowing money, and use of transfers received from government or NGOs. Since there is no a priori reason to believe that the responses to these questions would differ when respondent is not alone, we hypothesize that speakerphone should not have any impact on them.

Methodology

Variable description

First, we determine what individual and region- specific characteristics govern the use of speaker phone. For this, we regress speaker use on various household & respondent characteristics. We also control for survey attributes such as call duration, time of call during the day and whether the respondent was surveyed by same enumerator in both rounds to see how trust-building might affect this behavior. A full list of independent variables included are described in the Appendix. We run separate regressions for Dang and Gujarat. Since the question on speaker phone was asked at two places in the survey- in the beginning and before the conflict module, we construct a variable that captures whether the respondent had switched on the speaker at any time during the survey:

used_speaker= 1 if respondent had put the phone on speaker either in the beginning or before conflict module, 0 otherwise.

Secondly, we determine how speaker use impacts responses to various questions discussed before. Based on the questions described earlier, we construct categorical and continuous dependent variables as described in Table 1.

Table 1: Dependent variables to study the impact of speaker on responses

Decision making about own income
1. Respondent alone decides how to spend her income=1, spouse/other family member decides=0
2. Respondent is involved in decision making about how to spend her income (that is she either decides it alone or jointly with spouse/other family member)=1, spouse/other family member decides=0
Decision making about spouse's income
3. Respondent alone decides how to spend spouse's income=1, spouse/other family member decides=0
4. Respondent is involved in decision making about how to spend her husband's income (that is, she either decides alone or jointly with spouse/other family member) =1, spouse/other family member decides=0
Household care activities
5. Caring hours spent by respondent in last 1 day (hours)
6. Caring hours spent by spouse in last 1 day (hours)
7. Caring hours spent by respondent>caring hours by spouse=1, 0 if less than spouse
Coping mechanisms used to deal with income loss
8. Respondent used savings to deal with income loss due to Covid-19=1, 0 if not used
9. Respondent sold assets to deal with income loss due to Covid-19=1, 0 if not used
10. Respondent borrowed money to deal with income loss due to Covid-19=1, 0 if not used
11. Respondent used transfers from government to deal with income loss due to Covid-19=1, 0 if not used
12. Respondent used transfers from NGO to deal with income loss due to Covid-19=1, 0 if not used

The above dependent variables are regressed on speaker use along with other household and individual attributes. As seen in table 1, we create a weaker and a stronger form of certain variables. For instance, respondent deciding how to spend their income alone is a stronger statement than respondent deciding it jointly with someone. Similarly, respondent's caring hours are strictly greater than those of spouse is a stronger statement while respondent's caring hours are at least as much as spouse's is a weaker statement.

Panel framework

We employ the technique of panel data regression using data from two rounds of surveys, with:

$$t = 1,2$$

$$i = 1,2,3 \ldots, n1 \; for \; round \; 1$$

$$i = 1,2,3 \ldots, n2 \; for \; round \; 2$$

A panel data model with random effects is chosen since it allows to control for various time invariant explanatory variables which is not possible to do under fixed effects model. The Hausman test for a model with fixed effects versus random effects was carried out for all regressions which revealed a model of random effects would provide a better fit. A least squares linear regression is estimated for the continuous dependent variables and a linear probability model is estimated for binary dependent variables. Following is the estimating equation for the panel data model:

$$Y_{it} = \beta_0 + \beta_1 X_{1,it} + \beta_2 X_{2,it} + \ldots . + \beta_k X_{k,it} + \alpha + u_{it} + \varepsilon_{it}$$

Where,

Y_{it}= one of the dependent variables as defined previously at time *t* and for respondent *i*,

$X_{k,it}$= kth demographic characteristic of the respondent as in table A1,

β_k= the coefficient of the kth covariate,

u_{it}= between entity error term for entity *i* at time *t*, and

ε_{it}= within entity error term for entity *i* at time *t*.

We dropped those observations where the respondent replied with a `don't know' or refused to answer the question. We also dropped those respondents who did not have any earnings or whose spouse/primary male did not have any earnings from the respective regressions for those dependent variables.

Propensity score matching

To supplement our analysis, we use propensity score matching to study treatment effects. As in most non-randomized experiments, matching helps to create a set of individuals in control group who are similar to those in treatment group across various observable characteristics. Then the differences in outcomes between this selected control group and treatment group reflects treatment effect (Rubin 1974; Heckman 1997). In our case, speaker use is the treatment and those who those who don't use it are categorized as control group.

We employ the technique of one-to-one nearest neighbor matching. Since propensity score matching can only be performed for a cross sectional data and does not have a direct applicability to panel data, we match speaker users with non-users for each round separately and merge the matched sample from each round for analysis. This is needed since the treatment (speaker users) and control (speaker non-users) groups change in each round so that an individual could be matched with a different individual in each round. See appendix for t-tests on differences in covariates from p-score matching.

We follow standard Rubin model to measure treatment effect. In this model, the outcome if individual *i* receives the treatment effect is $Y_i(1)$ and the outcome if individual *i* receives the control is $Y_i(0)$. The treatment effect for an individual *i* can be written as $Y_i(1)- Y_i(0)$. This individual effect cannot be estimated since we observe each individual in only 1 state- either in treatment group or control group. Instead, we look at the average treatment effect:

$$\tau = E(Y(1)- Y(0))$$

which is called average treatment effect (ATE) if it is calculated for both treatment and control groups and called average treatment effect on treated (ATT) if calculated for just treatment group. We just estimate ATT to study impact of speakerphone use. We regress dependent variables in Table 2 on speaker use and other household & individual characteristics on this matched sample in panel framework.

Results

Descriptive statistics

The sample comprised of women with an average age of 35 years in Dang and 40 years in Gujarat (Table 2). Around 46% of women in Gujarat belonged to the general caste and 38% belonged to scheduled caste/scheduled tribes. Around 34% women in Dang (Table 3) belonged to Terai Janjati, while 42% women and 26% men belonged to Chhetri caste. 88% women in Gujarat and 93% women in Dang were married. Nearly one-third women in both countries had no formal education. In Gujarat, both urban and rural areas were covered from Ahmedabad district, while for all other districts, only rural areas were covered. The Nepal sample comprised of only rural areas from the four municipalities.

Table 2: Descriptive statistics of sample households in Gujarat, India

Characteristics		Mean(sd)/ proportion
Age (years)		40.03
Caste:	General caste (%)	45.93
	Scheduled caste	27.59
	Scheduled Tribe (%)	10.37
	Other backward class (%)	16.11
Religion:	Hindu (%)	97.77
	Muslim (%)	2.23
Family size		6.27

Married (%)		87.88
Education:	No formal schooling (%)	32.38
	Less than primary school (%)	21.85
	Completed primary school (%)	11.80
	Less than secondary school	20.57
	Completed secondary school (%)	7.34
	More than secondary school (%)	6.06
Occupation:	Agriculture and/or livestock (%)	39.55
	Casual labourer (%)	16.91
	Service providers (%)	0.96
	Street vendor (%)	28.55
	Home based worker (%)	5.42
	Salaried jobs (%)	4.63
	Unemployed (%)	3.99
District:	Ahmedabad urban (%)	30.46
	Ahmedabad rural (%)	6.54
	Anand (%)	8.61
	Arvalli (%)	6.86
	Chota Udaipur (%)	8.29
	Gandhinagar (%)	7.81
	Kutch (%)	7.02
	Mehsana (%)	9.09
	Patan (%)	8.13
	Surendranagar (%)	7.18
Number of observations		627

Note: Standard deviation in parentheses

Table 3: Descriptive statistics of sample households in Dang, Nepal

Characteristics		Mean(sd)/ Proportion
Age (years)		35 (10.3)
Caste:	Brahmin (%)	9.81
	Chhetri (%)	41.48

	Dalit (%)	6.3
	Newar (%)	0.56
	Hill Janjati (%)	6.48
	Terai Janjati (%)	34.44
	Muslim (%)	0.93
Family size		5 (2.01)
Married (%)		93.15
Education:	No formal schooling (%)	33.89
	Less than primary school (%)	9.81
	Completed primary school (%)	8.33
	Less than secondary school	25.37
	Completed secondary school (%)	11.3
	More than secondary school (%)	11.3
Occupation:	Agriculture and/or livestock (%)	80.74
	Casual labourer (%)	2.04
	Work for self (enterprise or business) (%)	2.59
	Wage or salaried jobs (%)	3.15
	Unemployed (%)	11.48
Municipality:	Dangisharan (%)	7.96
	Lamahi (%)	29.26
	Raptu (%)	9.26
	Shantinagar (%)	53.52
Number of observations		540

Note: Standard deviation in parentheses

The Gujarat sample comprised of respondents from varying occupations. 40% of respondents were engaged in agriculture/livestock farming as their primary occupation at the time of the survey. 29% were street vendors selling mainly fruits and vegetables (78%), recycling paper (16%) and selling other food and household items. 17% were labourers engaged as agriculture labourers or construction workers. Around 5% of respondents were home-based workers who were mainly engaged as readymade garment workers. There were a few who were employed

in public/private jobs on salary/wage (5%) and service providers (1%). 4% of the sample was unemployed at the time of survey. The occupations were area-specific, with street vendors mostly concentrated in urban areas, and others like farmers and casual laborers in rural areas. In Dang, majority of respondents (81%) were engaged in agriculture/livestock farming as their primary occupation at the time of the survey and very few were engaged as casual laborers, self-employed or on wage/salary basis. 11% also reported to be unemployed at the time of the survey.

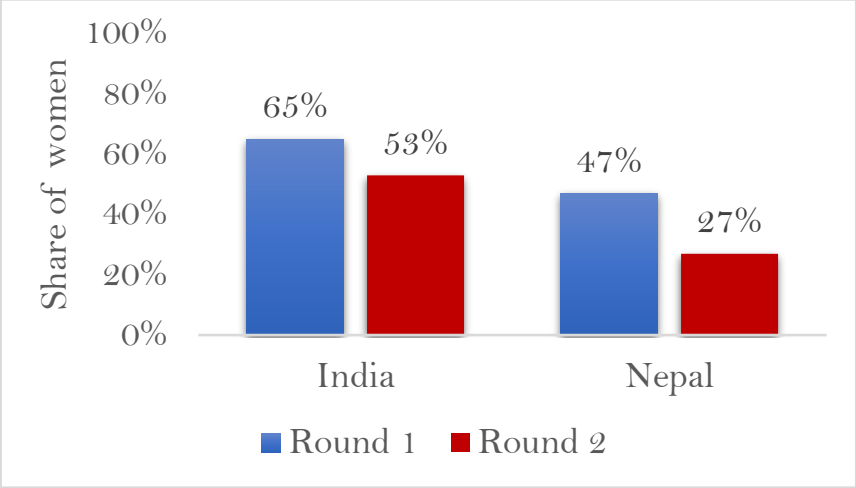
Majority of the respondents had access to piped water supply (78%) while a few relied on borewell/tubewell (16%) as well in Gujarat. In Nepal, roughly equal proportion of respondents used piped drinking water (50%) and ground water sources such as dug well, borewell and water pump (45%). Likewise, majority of respondents had access to a house toilet in both countries- 70% in Gujarat and 97% in Dang.

Only around one-fifth women in both countries reported themselves to be the household heads in round 1. In Gujarat, by default, for 86-88% of respondents, the primary male was husband, while 2% reported there was no primary male decision maker in the household. 66% of rural respondents in Gujarat said they owned agricultural land in round 1 with an average land owned of 4 hectares. In Dang, over 90% households reported owning agricultural land with an average land owned of 11 hectares. 44% respondents in Gujarat reported that their household owned a livestock animal which was mostly a cow, buffalo or poultry. In Nepal, 94% households reported to own a livestock which was mostly goats, poultry, and cattle. 63% of urban dwellers in Gujarat reported to be living in own home while the rest reported to be living in a rented home.

Out surveys revealed that 65% women in Gujarat had put their phones on speaker in round 1, either in the beginning of the survey or later before the conflict section, which got slightly

reduced to 53% in 2nd round. Speaker use in Nepal was comparatively lower in round 1, which further got almost halved by the second round (Figure 4).

Figure 4: Use of speaker phone during the survey



Source: Author's calculations

Intra-household decision making and care activities

In Gujarat, in both rounds, roughly, 1 in 3 women responded that she alone makes decisions on how to spend her earnings. Likewise, one-third women said their husband took this decision for them. In contrast, in round 1, only 19% women said they decided alone on how to spend the earnings of spouse/primary male and about 35% of the respondents said the male alone decided how to spend his income. But in round 2, 26% women reported to be deciding independently about husband/primary male's income and 28% reported that he decides it alone. In Dang, nearly 29% women in both rounds reported they alone made decision on how to spend own earnings, over 40% in both rounds said they do it jointly with their husbands, and 9% in round 1 and 12% in round 2 said their spouse does that alone. About 15% respondents in round 1 and 17% in round 2 said they alone made decision on how to spend spouse's earnings, 15% in round 1 and 18% in round 2 said their spouse does that alone, and more than 50% in both rounds said they do it jointly with their husband.

On average, the respondents in Gujarat had spent 7 hours on household care activities in the last 24 hours, while their spouse/primary male had spent 6 hours on the same immediately following the lockdown. These fell to 5 hours for both in round 2. However, in Dang, women reported spending 3 hours on household care in both rounds while their spouse spent 2 hours in both rounds. About 77-78% respondents reported that the amount of time spent on this activity has increased during the lockdown for themselves and the primary male, which fell to 54-60% by round 2. Likewise, over 36-37% women reported care time to have increased by both themselves and their spouse in round 1 which dropped to one-fourth in round 2.

Speaker phone and response bias

As shown in Table 4 for Gujarat, the responses to questions about who decides how to spend your own income differ by use of speaker in both rounds, with women underreporting making decision independently or jointly with spouse/primary male when speaker is on compared to when it is off. The t-tests don't show any significant difference in responses to who decides how to control income of spouse/primary male in round 1, but significant underreporting is observed in round 2 in this case as well. In Dang as shown in Table 5, there is significant difference between the 2 groups in proportion of woman reporting they control own income alone in only round 1, while for spouse's income, there is significant difference in proportion of woman reporting that they control their husband's income but only in round 2. No significant difference is observed in responses to questions on care activities and coping mechanisms between the two groups in both countries.

Table 4 : T-tests for impact of speaker use on outcome variables in Gujarat

	Round 1				Round 2			
	(1)	(2)	(3)	t-test	(1)	(2)	(3)	t-test
	speaker off	speaker on	Total	Difference	speaker off	speaker on	Total	Difference
Variable	Mean/SE	Mean/SE	Mean/SE	(1)-(2)	Mean/SE	Mean/SE	Mean/SE	(1)-(2)
Woman alone decides how to spend her income	0.573 (0.042)	0.458 (0.031)	0.499 (0.025)	0.115**	0.552 (0.037)	0.412 (0.037)	0.483 (0.026)	0.140***
Woman is involved in decision making about her income	0.712 (0.031)	0.639 (0.024)	0.664 (0.019)	0.073*	0.687 (0.029)	0.645 (0.028)	0.665 (0.020)	0.042
Woman alone decides how to spend husband's income	0.277 (0.042)	0.353 (0.031)	0.329 (0.025)	-0.076	0.551 (0.040)	0.340 (0.037)	0.443 (0.028)	0.212***
Woman is involved in decision making about husband's income	0.574 (0.036)	0.557 (0.027)	0.563 (0.021)	0.016	0.697 (0.030)	0.609 (0.030)	0.650 (0.021)	0.087**
Caring hours by woman	7.575 (0.449)	7.255 (0.262)	7.368 (0.232)	0.320	5.406 (0.243)	5.071 (0.223)	5.228 (0.164)	0.335
Caring hours by husband	6.110 (0.467)	6.534 (0.263)	6.387 (0.236)	-0.425	5.102 (0.302)	4.325 (0.214)	4.689 (0.182)	0.777**
Caring hours by woman> caring hours by husband	0.439 (0.033)	0.374 (0.024)	0.397 (0.020)	0.065	0.410 (0.030)	0.432 (0.029)	0.422 (0.021)	-0.022
Used saving to deal with income loss	0.466 (0.034)	0.502 (0.025)	0.490 (0.020)	-0.036	0.410 (0.030)	0.399 (0.028)	0.404 (0.021)	0.011

Used assets to deal with income loss	0.054 (0.015)	0.067 (0.012)	0.062 (0.010)	-0.012	0.038 (0.012)	0.040 (0.011)	0.039 (0.008)	-0.002
Used borrowings to deal with income loss	0.466 (0.034)	0.461 (0.025)	0.463 (0.020)	0.005	0.380 (0.030)	0.302 (0.027)	0.339 (0.020)	0.077*
Used government transfers to deal with income loss	0.425 (0.033)	0.340 (0.024)	0.370 (0.019)	0.085**	0.060 (0.015)	0.093 (0.017)	0.078 (0.011)	-0.033
Used NGO transfers to deal with income loss	0.041 (0.013)	0.047 (0.010)	0.045 (0.008)	-0.006	0.244 (0.026)	0.216 (0.024)	0.229 (0.018)	0.028

The value displayed for t-tests are the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table 5 : T-tests for impact of speaker use on outcome variables in Dang

	Round 1				Round 2			
	(1)	(2)	(3)	t-test	(1)	(2)	(3)	t-test
	speaker off	speaker on	Total	Difference	speaker off	speaker on	Total	Difference
Variable	Mean/SE	Mean/SE	Mean/SE	(1)-(2)	Mean/SE	Mean/SE	Mean/SE	(1)-(2)
Woman alone decides how to spend her income	0.939 (0.013)	0.875 (0.020)	0.910 (0.012)	0.064***	0.883 (0.016)	0.895 (0.025)	0.887 (0.013)	-0.012
Woman is involved in decision making about her income	0.656 (0.036)	0.603 (0.037)	0.630 (0.026)	0.052	0.613 (0.031)	0.682 (0.051)	0.631 (0.027)	-0.069
Woman alone decides how to spend husband's income	0.862 (0.020)	0.813 (0.024)	0.839 (0.015)	0.049	0.807 (0.021)	0.862 (0.029)	0.823 (0.017)	-0.055

Woman is involved in decision making about husband's income	0.430 (0.045)	0.358 (0.044)	0.394 (0.032)	0.071	0.356 (0.036)	0.491 (0.068)	0.388 (0.032)	-0.135*
Caring hours by woman	2.784 (0.134)	2.556 (0.128)	2.679 (0.093)	0.228	2.865 (0.122)	2.542 (0.196)	2.782 (0.104)	0.323
Caring hours by husband	-17.431 (2.062)	-12.848 (1.987)	-15.319 (1.442)	-4.583	-17.104 (1.842)	-19.941 (3.235)	-17.844 (1.601)	2.836
Caring hours by woman> caring hours by husband	0.453 (0.025)	0.457 (0.027)	0.455 (0.018)	-0.004	0.449 (0.022)	0.416 (0.037)	0.441 (0.019)	0.033
Used assets to deal with income loss	0.603 (0.024)	0.649 (0.026)	0.625 (0.018)	-0.046	0.570 (0.022)	0.517 (0.038)	0.557 (0.019)	0.053
Used borrowings to deal with income loss	0.100 (0.015)	0.057 (0.012)	0.080 (0.010)	0.042**	0.189 (0.017)	0.140 (0.026)	0.177 (0.015)	0.049
Used government transfers to deal with income loss	0.440 (0.025)	0.448 (0.027)	0.444 (0.018)	-0.008	0.434 (0.022)	0.489 (0.038)	0.448 (0.019)	-0.055
Used NGO transfers to deal with income loss	0.105 (0.015)	0.132 (0.018)	0.117 (0.012)	-0.028	0.025 (0.007)	0.034 (0.014)	0.028 (0.006)	-0.008
Caring hours by woman> caring hours by husband	0.010 (0.005)	0.009 (0.005)	0.009 (0.003)	0.001	0.002 (0.002)	0.011 (0.008)	0.004 (0.003)	-0.009

The value displayed for t-tests are the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Results from regressions

Table 6 gives the results from the regression for determinants of speaker phone for Gujarat. We produce results from separate regressions for each round, and also a panel regression taking both rounds together. The variable capturing whether the respondent was surveyed by same enumerator in round 2 takes the value 1 if yes, and 0 otherwise, but only 0 in round 1, and is hence excluded from round 1 regression. We observe that having same enumerator, as we move from round 1 to round 2 reduces the likelihood of putting the phone on speaker. We also find significant negative impacts of survey time. We don't see any strong impact of other respondent's attributes such respondent's education, age, caste, marital status and family size on speaker phone behavior.

Table 6: Speaker use determinants- Gujarat

	Round 1	Round 2	Both Rounds
Same enumerator between	-	-0.237***	-0.181***
Round 1 & 2		(0.048)	(0.044)
Total survey time (mins)	-0.016***	-0.005**	-0.010***
	(0.002)	(0.002)	(0.002)
Time of call: Afternoon=1	0.025	0.106	0.063
	(0.059)	(0.065)	(0.043)
Time of call: Evening=1	-0.054	0.173***	0.058
	(0.058)	(0.065)	(0.043)
Respondent is household	0.105**	-0.062	0.012
head=1	(0.047)	(0.052)	(0.036)
Occupation of respondent:			
Labor=1	-0.091	-0.256***	-0.180***
	(0.057)	(0.063)	(0.045)
Vendor=1	0.168	0.063	0.094
	(0.111)	(0.104)	(0.083)
Other job=1	-0.097	-0.068	-0.076*
	(0.065)	(0.067)	(0.045)
Observations	621	559	1180

rho	0.152

Note: Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. We also control for urban, whether respondent uses digital payment methods, total livestock owned, family size, religion, caste fixed effects, marital status, educational background, age and round dummy in panel regression.

Table 7 gives results from similar regression for determinants of speaker phone for Dang. We find no impact of being surveyed by same enumerator on speaker use. Instead, we find strong negative impact of education, as seen by negative significant coefficients for education dummies. We also find that upper caste women have a lower likelihood of using speaker compared to lower castes. Like Gujarat, increased survey duration reduces the likelihood of using speaker phone. Other attributes like household head status, marital status, age, family size is not found to impact speaker phone use.

Table 7: Speaker phone use determinants- Dang

	Round 1	Round 2	Both Rounds
Same enumerator between	-	0.034	0.025
Round 1 & 2		(0.052)	(0.043)
Total survey time (mins)	-0.011***	-0.009***	-0.010***
	(0.003)	(0.002)	(0.002)
Time of call: Afternoon=1	-0.003	-0.063	-0.038
	(0.054)	(0.047)	(0.035)
Time of call: Evening=1	0.083	-0.014	0.056
	(0.051)	(0.056)	(0.038)
Respondent is household	0.015	-0.020	0.002
head=1	(0.056)	(0.054)	(0.040)
Upper caste=1	-0.108**	-0.081*	-0.096***
	(0.049)	(0.047)	(0.036)
Education status:			
Completed primary or less=1	-0.096	-0.152**	-0.116**
	(0.065)	(0.059)	(0.047)
Completed 2ndary or less=1	-0.078	-0.049	-0.061
	(0.065)	(0.064)	(0.046)
Completed more than 2ndary=1	-0.222**	-0.258***	-0.232***

	(0.087)	(0.070)	(0.056)
Respondent occupation:			
Unemployed=1	-0.055	-0.111*	-0.091*
	(0.070)	(0.063)	(0.051)
All others=1	0.048	-0.051	-0.004
	(0.081)	(0.075)	(0.056)
Observations	540	475	1015
rho			0.172

Note: Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. We also control for total livestock owned, family size, marital status, age, municipality fixed effects and round dummy in panel regression.

Table 8 reports results from regressions for earnings of the respondent for both full sample and matched sample for Gujarat and Dang. While we control for various individual and household characteristics in these and subsequent regressions, only the coefficient of speaker use is displayed henceforth for brevity. Full results from all the regressions can be found in the appendix. In Gujarat, we observe that having the speaker phone on during the surveys makes the woman less likely to report that she makes the decision about how to spend her income either independently or jointly with the spouse/primary male. Similar result is observed for matched sample. In Dang, we see impact of speaker in 2 outcomes- speaker use makes the women less likely to report that she decides how to spend her income versus anyone else deciding it, and less likely to report that she is involved in joint decision making with other family member versus other family member deciding it. does not impact respondent's agency over own income except in one case where the woman reports she makes this decision jointly with other family member.

Table 9 gives the regression results for earnings of husband/primary male for Gujarat and Dang. Even here, speaker phone reduces the likelihood of women reporting any involvement in decision-making on the earnings of the primary male/husband independently or jointly in Gujarat, but no concurrent impact is seen in Dang.

Table 10 and Table 11 gives the estimates for regressions on household care activities for Gujarat and Dang respectively. We don't see any impact of speaker phone on the amount of time spent on care activities by women or their spouse. Also, having the speaker phone on does not impact the likelihood of women reporting more care time than their spouse. This is true across both full sample and matched sample regressions.

Table 12 and Table 13 report results on regression for coping mechanisms to deal with income loss in COVID-19 for Gujarat and Dang respectively. As hypothesized, speaker phone does not impact the reporting of use of various coping mechanisms for both full sample and matched sample regressions in both countries.

Discussion and conclusion

Putting the phone on speaker while talking is a common practice that we observe around us all the time. That an innocuous act such as this could bias survey responses, to the best of our knowledge, has never been evaluated. As phone surveys continue to rise in popularity, there is also the larger ethical question of respondents' breach of privacy in the interview process that merits closer attention. Using phone survey data from India and Nepal, we test the determinants of speaker use and the impact of speaker use on women's response to sensitive and non-sensitive questions.

Our results show weak evidence that speaker use can be explained partially by trust-building, and as respondents develop trust with the surveyors overtime, speakerphone use reduces in subsequent rounds. However, this is observed only in India and not in Nepal. There are many reasons why this might be true. Firstly, in India, we were forewarned by SEWA staff about women using speaker at the behest of their family, while recruiting respondents for surveys, implying some level of coercion. Moreover, early signs of intrusion by family members in

Table 8: Own income regressions

Who decides how to spend earnings of the woman? 0= Spouse/Primary male/ Other family member decides								
	Gujarat				Dang			
	Woman alone decides=1		Woman is involved in decision making process=1		Woman alone decides=1		Woman is involved in decision making process=1	
	Full sample	Matched sample	Full sample	Matched sample	Full sample	Matched sample	Full sample	Matched sample
Used speaker=1	-0.118***	-0.116***	-0.054**	-0.049*	-0.046*	-0.029	-0.048	0.002
	-0.032	-0.033	-0.027	-0.028	-0.027	-0.02	-0.05	-0.042
N	735	717	1128	1101	716	988	407	560
Rho	0.218	0.215	0.184	0.17	0.23	0.199	0.407	0.416
R-squared	0.286	0.285	0.122	0.118	0.052	0.051	0.244	0.025

Note: Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. We also control for whether respondent is household head, marital status, age, education, caste, family size, respondent & spouse's occupation, municipality Fes, land owned and livestock size of the household.

Table 9: husband's income regression

Who decides how to spend earnings of the husband/primary male? 0= Spouse/Primary male/Other family member decides				
	Gujarat		Dang	
	Woman alone decides=1	Woman is involved in decision making process=1	Woman alone decides=1	Woman is involved in decision making process=1

	Full sample	Matched sample	Full sample	Matched sample	Full sample	Matched sample	Full sample	Matched sample
Used speaker=1	-0.073**	-0.070**	-0.046	-0.044	-0.002	-0.007	-0.008	0.006
	-0.033	-0.035	-0.03	-0.03	-0.03	-0.025	-0.051	-0.048
N	666	646	999	976	724	961	356	421
Rho	0.24	0.22	0.168	0.162	0.186	0.207	0.467	0.503
R squared	0.297	0.297	0.131	0.126	0.06	0.062	0.271	0.283

Note: Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. We also control for whether respondent is household head, marital status, age, education, caste, family size, respondent & spouse's occupation, municipality Fes, land owned and livestock size of the household.

Table 10: Caring hours regression- Gujarat

	Full sample			Matched sample		
	Caring hours by woman	Caring hours by husband/primary male	Caring hours by woman> caring hours by husband/primary male	Caring hours by woman	Caring hours by husband/primary male	Caring hours by woman> caring hours by husband/primary male
Used speaker=1	-0.394	-0.364	-0.006	-0.403	-0.392	-0.002
	-0.328	-0.344	-0.029	-0.332	-0.349	-0.03
N	1158	1156	1162	1131	1129	1135
rho	0.099	0.173	0.177	0.073	0.149	0.169
R squared	0.11	0.13	0.074	0.103	0.125	0.072

Note: Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. We also control for whether respondent is household head, marital status, age, education, caste, family size, respondent & spouse's occupation, municipality Fes, land owned and livestock size of the household.

Table 11: Caring hours regressions- Dang

	Full sample			Matched sample		
	Caring hours by woman	Caring hours by husband	Caring hours by woman> caring hours by husband	Caring hours by woman	Caring hours by husband	Caring hours by woman> caring hours by husband
Used speaker=1	-0.207	-1.162	-0.022	-0.123	-1.218	-0.003
	-0.175	-2.159	-0.032	-0.147	-1.623	-0.026
N	869	869	869	1217	1217	1217
rho	0.244	0.521	0.161	0.166	0.511	0.236
R squared	0.122	0.39	0.159	0.127	0.387	0.162

Note: Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. We also control for whether respondent is household head, marital status, age, education, caste, family size, respondent & spouse's occupation, municipality Fes, land owned and livestock size of the household.

Table 12: Coping mechanism regressions- Gujarat

	Full sample					Matched sample				
	Used saving	Used assets	Used borrowings	Used govt transfers	Used NGO transfers	Used saving	Used assets	Used borrowings	Used govt transfers	Used NGO transfers
Used speaker=1	0.01	0.003	-0.017	-0.018	-0.012	0.004	0.003	-0.014	-0.019	-0.012
	-0.03	-0.014	-0.023	-0.025	-0.019	-0.031	-0.014	-0.023	-0.025	-0.019
N	1162	1162	1162	1162	1162	1135	1135	1135	1135	1135

rho	0.194	0	0.151	0.117	0	0.202	0	0.144	0.111	0
R squared	0.082	.048	0.424	0.163	0.185	0.079	0.054	0.421	0.16	0.189

Note: Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. We also control for whether respondent is household head, marital status, age, education, caste, family size, respondent & spouse's occupation, municipality Fes, land owned and livestock size of the household.

Table 13: Coping mechanisms regressions- Dang

	Full sample					Matched sample				
	Used saving	Used assets	Used borrowings	Used govt transfers	Used NGO transfers	Used saving	Used assets	Used borrowings	Used govt transfers	Used NGO transfers
Used speaker=1	-0.011	-0.034	0.011	0.018	0.006	-0.004	-0.039**	0.018	0.019	0.003
	-0.036	-0.023	-0.036	-0.02	-0.005	-0.03	-0.019	-0.03	-0.017	-0.006
N	869	869	869	869	869	1217	1217	1217	1217	1217
rho	0.21	0	0.294	0.014	0	0.24	0	0.297	0.019	0.045
R squared	0.061	0.071	0.05	0.09	0.04	0.063	0.073	0.054	0.096	0.042

Note: Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. We also control for whether respondent is household head, marital status, age, education, caste, family size, respondent & spouse's occupation, municipality Fes, land owned and livestock size of the household.

surveys were also detected by SEWA staff and reported to the survey team. However, in Nepal, we did not have any such prior evidence on women using speaker phone during surveys. Secondly, it is possible that the affiliation of enumerators with SEWA in Gujarat was helpful in engaging with respondents as they were otherwise completely new to the respondents. In contrast, in Dang, respondents already knew the enumerators since they had been previously contacted during an in-person household listing exercise. This meant that there was a significant difference in the familiarity with enumerators between the two regions. As a result, developing trust with enumerator is more significant in India than in Nepal in reducing speaker use. This points out to the role played by study context and methods of obtaining phone numbers in determining speaker behavior.

Our regressions also reveal the commonality in certain findings between the two regions. For instance, women who are household heads and/or older are more likely to control own and husband's income in both countries. There is also a role played by caste, household size, education and respondent's occupation in responses to most of these questions. Since these covariates don't show as significant in the regressions to determine speaker use behavior, we conclude there is no issue of endogeneity in estimating impact of speaker on various dependent variables along with these covariates. This also shows that there are other factors that determine speaker use behavior which might be out of scope for this study. There is also strong negative impact of survey duration on likelihood to use speaker. While there is no direct evidence to support this, but relating this to third party presence, it might indicate the fact that it is easier to overhear and intrude in shorter conversations.

We also see that speaker phone biases responses to questions on household and income decision-making causing women to under-report their agency over their own income and the income of their husband/primary male earning member in the house. Again, the results are

stronger for India than Nepal. On the time spent on household care activities and objective questions of crisis coping strategies, we find no impact of speaker phone on responses in the panel framework.

Ensuring privacy of the respondent during phone surveys is important and asking indirectly about speaker phone can be a viable proxy measure. Without more careful framing of questions, it is hard to differentiate whether women are being forced by others to have their phones on speaker or they do so of their own volition. Regardless, we demonstrate that it is important for researchers and practitioners to have information on whether respondents are answering phone survey questions in private and to account for it when analyzing data. How such bias can be corrected and accounted for is beyond the scope of this paper but remains an area for further research and enquiry.

References

Anderson, B. A., & Silver, B. (1987). The Validity of Survey Response: Insights from Interviews of Married Couples in a Survey of Soviet Emigrants. *Social Forces*, 66, 537-554. doi:http://dx.doi.org/10.1093/sf/66.2.537

Aquilino, W. (1993). Spouse presence during the interview on survey responses concerning marriage. *Public Opinion Quarterly*, 57, 358-76.

Aquilino, W. S. (1997). Privacy Effects on Self-Reported Drug Use: Interactions with Survey Mode and Respondent Characteristics. In L. Harrison, & A. Hughes, National Institute on Drug Abuse Research Monograph Series: The Validity of Self-Reported Drug Use. Improving the Accuracy of Survey Estimates (pp. 383-415). Washington D. C.: U.S. Department of Health and Human Services, National Institutes of Health.

Aquilino, W., Wright, D., & Supple, A. (2000). Response effects due to bystander presence in CASI and paper-and-pencil surveys of drug use and alcohol use. *Substance Use and Misuse*, 35, 845-967. doi:http://dx.doi.org/10.3109/10826080009148424

Austrian, K., Pinchoff, J., Tidwell, J. B., White, C., Abuya, T., Kangwana, B., ... & Mwanga, D. (2020). COVID-19 related knowledge, attitudes, practices and needs of households in informal settlements in Nairobi, Kenya. *Bulletin of the World Health Organization*. doi:

http://dx.doi.org/10.2471/BLT.20.260281

Blaydes, L., & Gillum, R. (2013). Religiosity-of-Interviewer Effects: Assessing the Impact of Veiled Enumerators on Survey Response in Egypt. *Politics and Religion*, 6(3), 459-482. doi:10.1017/S1755048312000557

Bradburn, N., Seymour, S., Blair, E., Locander, W., Miles, C., Singer, E., & Stocking, C. (1979). Improving interview method and questionnaire design: Response effects to threatening questions in survey research. San Fransisco: Jossey-Bass.

Brick, J.M., Brick, P.D., Dipko, S., Presser, S., Tucker, C., Yuan, Y. (2007). Cell Phone Survey Feasibility in the U.S.: Sampling and Calling Cell Numbers versus Landline Numbers. *Public Opinion Quarterly*. 71, 23–29.

Casterline, J., & Chidambaram, V. (1984). The presence of others during the interview and the reporting of contraceptive knowledge and use. In J. Ross, & R. McNamara, Survey Analysis for the Guidance of Family Planning Programs (pp. 267-298). Leige: Ordina Editions.

Ceballos, F., Kannan, S., & Kramer, B. (2020). Impacts of a national lockdown on smallholder farmers' income and food security: Empirical evidence from two states in India. *World Development*, 136, 105069.

Cilliers, J., Dube, O., & Siddiqi, B. (2015). The white-man effect: How foreigner presence affects behavior in experiments. *Journal of Economic Behavior & Organization*, 118, 397-414.

Demeke, E., Hardy, M., Kagy, G., Meyer, C. J., & Witte, M. (2020). The Impact of COVID-19 on the Lives of Women in the Garment Industry: Evidence from Ethiopia.

Dillman, D. A., & Christian, L. M. (2005). Survey mode as a source of instability in responses across surveys. *Field methods*, 17(1), 30-52.

Diop, A., Le, K., & Traugott, M. (2015). Third party presence effect with propensity score matching. *Journal of Survey Statistics and Methodology*, 3(2), 193-215. doi:10.1093/jssam/smv005

Dyer, J., Wilson, K., Badia, J., Agot, K., Neary, J., Njuguna, I., ... & Kohler, P. (2020). The Psychosocial Effects of the COVID-19 Pandemic on Youth Living with HIV in Western Kenya. *AIDS and Behavior*, 1-5.

Edwards, S. L., Slattery, M., & Ma, K. (1998). Measurement Errors Stemming from Nonrespondents Present at In-Person Interviews. *Annals of Emidemiology*, 8(4), 272-277.

English, Kevin C. DrPH; Espinoza, Judith MPH; Pete, Dornell MPH; Tjemsland, Amanda MPH. (2019). A Comparative Analysis of Telephone and In-Person Survey Administration for Public Health Surveillance in Rural American Indian Communities. *Journal of Public Health Management and Practice*, 25, 70-76.

Gfroerer, J. (1985). Underreporting of drug use by youths resulting from lack of privacy in household interviews. In B. Rouse, N. Kozel, & L. Richards, Self-Report Methods for Estimating Drug Use: Meeting Current Challenges (pp. 22-30). Washington D. C.: National Institue on Drug Abuse.

Grimm, P. (2010). Social desirability bias. Wiley international encyclopedia of marketing.

Groves, R. M., & Kahn, R. L. (1979). Surveys by telephone: A national comparison with personal interviews. New York, NY: Academic Press.

Hader, M. (2011). Data quality in telephone surveys via mobile and landine phones. In S. Hader, M. Hader, & M. Kuhne, Telephone surveys in Europe (pp. 247-262). Berlin, Heidelberg: Springer.

Harris, J., Depenbusch, L., Pal, A. A., Nair, R. M., & Ramasamy, S. (2020). Food system disruption: initial livelihood and dietary effects of COVID-19 on vegetable producers in India. *Food Security*, 12(4), 841-851.

Hartmann, P. (1995). Response behavior in interview settings of limited privacy. *International Journal of Public Opinion Research*, 7, 383– 390.

Heckman, J. (1997). Instrumental variables: A study of implicit behavioral assumptions used in making program evaluations. *Journal of human resources*, 441-462.

Hirvonen, K. (2020). Economic impacts of COVID-19 pandemic in Ethiopia: A review of phone survey evidence (Vol. 151). International Food Policy Research Institute.

Holbrook, A. L., Green, M. C. and Krosnick, J. A. (2003). Telephone vs. face-to-face interviewing of national probability samples with long questionnaires: comparisons of respondent satisficing and social desirability response bias. *Public Opinion Quarterly* 67, 79–125.

Hoyt, G., & Chaloupka, F. (1994). Effect of Survey Conditions on Self-Reported Substance Use. *Contemporary Economic Policy*, 7, 109-121. doi:http://dx.doi.org/10.1111/j.1465-7287.1994.tb00439.x

Jordan, L. A., Marcus, A. C., & Reeder, L. G. (1980). Response styles in telephone and household interviewing: A field experiment. *Public Opinion Quarterly*, 44, 210-222.

Kempf, A. M., & Remington, P. L. (2007). New challenges for telephone survey research in the twenty-first century. *Annual Reviews of Public Health*, 28, 113-126.

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, 5, 213-236.

Krumpal, I. (2011). Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & Quantity*, 47(4), 2025–2047. doi:10.1007/s11135-011-9640-9

Krysan, M.: Privacy and the expression of white racial attitudes—a comparison across three contexts. *Public Opinion Quarterly*, 62, 506–544 (1998)

Kuang, J., Ashraf, S., Das, U., & Bicchieri, C. (2020). Awareness, risk perception, and stress during the COVID-19 pandemic in communities of Tamil Nadu, India.

Lee, R.M., Renzetti, C.M. (1993). The Problems of Researching Sensitive Topics: An Overview and Introduction. In: Renzetti, C.M., Lee, R.M. (eds.) Researching Sensitive Topics, Sage, London.

Lee, R.M. (1993). Doing Research on Sensitive Topics. Sage, London.

Lensvelt-Mulders, G.J.L.M. (2008). Surveying sensitive topics. In: De Leeuw, E.D., Hox, J.J., Dillman, D.A. (eds.) The international Handbook of Survey Methodology, Erlbaum/Taylor & Francis, New York/London.

Lynn, P., & Kaminska, O. (2011). Factors affecting measurement error in mobile phone interviews. In M. H. S. Hader, Telephone surveys in Europe (pp. 211-228). Berlin, Heidelberg: Springer.

Mavletova, A., & Couper, M. (2013). Sensitive Topics in PC Web and Mobile Web Surveys: Is there a difference? *Survey Research Methods*, 7(3), 191-205.

Mneimneh, Z., Tourangeau, R., Pennell, B.-E., Heeringa, S., & Elliott, M. (2015). Cultural Variations in the Effect of Interview Privacy and the Need for Social Conformity on Reporting Sensitive Information. *Journal of Official Statistics*, 31(4), 673-697. doi:http://dx.doi.org/10.1515/JOS-2015-0040

Mneimneh, Z.N. 2012. "Interview Privacy and Social Conformity Effects on Socially Desirable Reporting Behavior: Importance of Cultural, Individual, Question Design and Implementation Factors." Available at: http://deepblue.lib.umich.edu/handle/2027.42/96051

Moskowitz, J. (2004). Assessment of Cigarette Smoking and Smoking Susceptibility Among Youth: Telephone Computer-Assisted Self-Interviews Versus Computer Assisted Telephone Interviews. *Public Opinion Quarterly*, 68, 565-587. doi:http://dx.doi.org/10.1093/poq/nfh040

Pollner, M., & Adams, R. (1994). The interpersonal context of mental health interviews. *Journal of Health and Social Behaviour*, 35, 283-290.

Reuband, K. (1992). On Third Persons in the Interview Situation and Their Impact on Responses. *International Journal of Public Opinion Research*, 4, 269-274.

Roth, S., Montaquila , J., & Brick, J. (2001). Effects of telephone technologies and call screening devices on sampling, weighting and cooperation in a random digit dialing (RDD) survey. Annual Conference of the American Association for Public Opinion Research. Montreal, Quebec, Canada.

Rubin, D. B. (1974), "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies," *Journal of Educational Psychology*, 66, 688–701.

Silver, B., Abramson, P., & Anderson, B. (1986). The presence of others and overreporting of voting in American national elections. *Public Opinion Quarterly*, 50(2), 228-239.

Smith, T. (1997). The impact of the presence of others on a respondent's answers to questions. *International Journal of Public Opinion Research*, 9(1), 33-37. doi:10.1093/ijpor/9.1.33

Stocké, V. (2004). The interdependence of determinants for the strength and direction of social desirability bias in racial attitude surveys.

Szolnoki, G., & Hoffmann, D. (2013). Online, face-to-face and telephone surveys—Comparing different sampling methods in wine consumer research. *Wine Economics and Policy*, 2(2), 57-66.

Tourangeau, R. (1984). Cognitive sciences and survey methods. In T. Jabine, M. Straf, J. Tanur, & R. Tourangeau (Eds.), Cognitive aspects of survey methodology: Building a bridge between disciplines. Washington, DC: National Academy Press

Tourangeau, R., & Smith, T. (1996). Asking Sensitive Questions: The Impact of Data Collection Mode, Question Format, and Question Context. *Public Opinion Quarterly*, 60, 275-304.

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, 133(5), 859-883.

Tsuchiya, T., Hirai, Y., & Ono, S. (2007). A study of the properties of the item count technique. *Public Opinion Quarterly*, 71(2), 253-272.

Turner, C., Ku, L., Rogers, S., Lindberg, L., Pleck, J., & Sonenstein, F. (1998). Adolescent sexual behavior. drug use and violence: New survey technology detects elevated prevalence among U.S. males. *Science*, 280, 867-873.

Appendix

Table A.1: List of independent variables used in regressions

Variable	Description
Used Speaker=1	=1 if speaker was switched on either in the beginning of the survey or just before conflict module, 0 otherwise
Total survey time	Total time taken in the survey in minutes
Survey timing	=0 if the survey was conducted in morning (before 12 am), =1 if survey was conducted in afternoon (between 12 am and 6 pm), =2 if survey was conducted in evening (after 6 pm). Morning is used as base category and separate dummy variables are inserted for Afternoon and Evening.
Household head is respondent	=1 if respondent is the household head, 0 otherwise
Urban	=1 if the respondent belongs to urban region, 0 otherwise
Household size	Total number of household members living in the family
Age	Age of the respondent in years
Age squared	Square of age of the respondent
Education dummies	=1 if the respondent has attained some form of education (>= primary education), 0 if no formal schooling attained
Own occupation	Fixed effects of respondent's primary occupation. The reference category in each regression is Agriculture/livestock farmers. A separate dummy variable is included for the rest of occupations: For Gujarat- casual labor, street vendor and all others (combining service provider home based worker, wage/salary job person and unemployed) For Dang: Unemployed and all others (combining laborers, self-employed and others)
Spouse occupation	Fixed effects of husband's/primary male's primary occupation. The reference category in each regression is Agriculture/livestock farmers. A separate dummy variable is included for the rest of occupations namely casual labor, service provider, street vendor, home based worker, wage/salary job person, unemployed and others.
Caste	Caste fixed effects. For India: The reference category is general caste. A separate dummy variable is included for Scheduled caste/Scheduled tribe and Other backward Classes. For Nepal: Reference category is lower caste and dummy is included for upper castes.

Religion	Only for India: Religion fixed effects. Base category is Muslims. A dummy variable is included for Hindu.
Agricultural land	Agricultural land owned by the household (hectares)
Agricultural land squared	Square of agricultural land owned by household (hectares)
Municipality	Nepal: The reference category is Dangisharan. Dummies are included each for Lamahi, Rapti, Shantinagar.
Married	=1 if respondent is married, 0 otherwise
Total animal	Total number of animals owned by the household

Link to access full regression results:

https://www.dropbox.com/sh/r2o2ovzzcp0c1hp/AAANg4L-0CAf-bQpz0Js50sza?dl=0