NATIONAL COUNCIL OF APPLIED ECONOMIC RESEARCH

Working Paper No.: WP 136

Coverage and Nonresponse Bias in Telephone Surveys during the COVID-19 Lockdown in India

Santanu Pramanik, Neerad Deshmukh, Bijay Chouhan, Manjistha Banerji, Reem Ashraf, Dinesh Tiwari, and Sonalde Desai

March 2022

National Council of Applied Economic Research

COVERAGE AND NONRESPONSE BIAS IN TELEPHONE SURVEYS DURING THE COVID-19 LOCKDOWN IN INDIA

NCAER Working Paper

Santanu Pramanik*, Neerad Deshmukh*, Bijay Chouhan*, Manjistha Banerji*, Reem Ashraf*, Dinesh Tiwari*, and Sonalde Desai**

Abstract

Objectives: In the wake of the COVID-19 pandemic, telephone surveys have been used extensively for carrying out studies on health knowledge, morbidity, and mortality surveillance. In order to understand the extent of different sources of non-observation errors in telephone surveys, we compare the distributions of units covered in the sampling frame and survey respondents with those who were excluded from the sampling frame and survey nonrespondents, respectively. The distributions are compared with respect to key socio-economic and demographic characteristics, which are often associated with most health outcomes for two different study designs, viz., panel surveys and repeated cross-sectional surveys.

Design: We used data from two rounds of the Delhi NCR Coronavirus Telephone Survey (DCVTS). DCVTS-1 was a panel sample that had been surveyed face-to-face a maximum of four times prior to Round 1 of the telephone survey. For DCVTS-2, we used a fresh sample using a repeated cross-sectional design as opposed to following the DCVTS-1 sample.

Setting: The target population for DCVTS is the National Capital Region (NCR), a diverse region comprising the metropolitan areas of Delhi as well as urban and rural areas of surrounding districts from the States of Haryana, Rajasthan, and Uttar Pradesh (UP).

Participants: For DCVTS-1, a sample of households selected for a pre-pandemic face-to-face survey, using a three-stage stratified cluster sampling design, was used. For DCVTS-2, we randomly selected new households from the same set of selected clusters.

Results: Following are the findings of our study:

(1) As compared to developed countries, the telephone survey response rates are higher in India, even when we did not have prior contact with the respondents, as in the case of DCVTS-2;

(2) While pre-existing rapport with the respondents and access to updated phone numbers lead to a higher response rate in a panel sample, as was observed with DCVTS-1, the cumulative attrition over panel rounds result in distributions of respondents and nonrespondents being different with respect to key background characteristics of households;

(3) The exclusion of households from the telephone survey sampling frame due to lack of access to phone or the unavailability of phone numbers, as was observed in DCVTS-2, may lead to under-representation of poor households;

(4) Nonresponse due to non-functional telephone numbers may occur at random and does not seem to be associated with the household background characteristics.

Conclusion: The above findings suggest that the representativeness of telephone survey samples may vary depending upon the study design, and the completeness and accuracy of the sampling frame..

Keywords: Delhi NCR Coronavirus Telephone Survey, Nonobservation Error, Panel Survey, Remote Mode of Data Collection, Repeated Cross-sectional Survey, Representativeness of Surveys

JEL codes: C830

---

*NCAER National Data Innovation Centre, National Council of Applied Economic Research, New Delhi, India
**Department of Sociology, University of Maryland, College Park, USA
Corresponding Author: Santanu Pramanik, Email: spramanik@ncaer.org

1. Introduction

The widespread use of mobile phones in India provided us with the opportunity of conducting surveys remotely during the COVID-induced country-wide lockdown at a time when there was a need for scientifically collected data to enable decision-making on various aspects of the pandemic. Moreover, the telephone mode of data collection coupled with computer-assisted technology satisfies the need of a quick turnaround in the absence of travel time and helps in measuring or informing policy responses in a timely manner. The challenges of obtaining representative samples in countries lacking universal access to telephones and the absence of a sampling frame of verified phone numbers have led researchers to leverage existing panels to provide rapid and useful data (Himelein et al. 2020). However, leveraging existing panels, in turn, poses different challenges for sample representativeness due to selective attrition.

In the pre-COVID era, telephone survey was not a common method for conducting household surveys in India. Hence, the effect of the use of the phone mode of data collection on survey estimates is not fully known. We compare the distributions of units covered in the sampling frame and survey respondents with those who were excluded from the sampling frame and survey nonrespondents, respectively. The distributions have been compared with respect to the characteristics that are usually associated with most health outcomes, that is, the State of residence (Delhi, Haryana, Rajasthan, or UP), area of residence (rural or urban), caste, religion, and wealth quintiles. The distributions have also been compared for two different study designs, viz., panel surveys and repeated cross-sectional surveys. Further, we also explore the reasons for non-inclusion of households in telephone samples, such as lack of access to mobile phones, phone numbers being non-functional, and unwillingness to participate in phone surveys.

2. Methods

2.1. Data and Study Design

In order to understand and quantify the early impact of the Coronavirus pandemic and the resultant lockdown, the National Council of Applied Economic Research (NCAER) conducted telephone surveys in rural and urban areas of the National Capital Region (NCR). The objectives of the first two rounds of the NCR Coronavirus Telephone Survey (DCVTS, Rounds 1 and 2) were to estimate the levels and changes over time in people's knowledge, attitudes, perceptions, and practiced behaviour with respect to COVID-19. The surveys also estimated the impact of the Coronavirus pandemic on people's livelihoods, income, access to essential items, social lives, and their coping mechanisms (NCAER National Data Innovation Centre. 2020a; 2020b).

In order to understand the extent of different sources of non-observation errors (Peytchev et al. 2011) in the DCVTS, Rounds 1 and 2 samples, we have used data from five different but related sources. The data sources, their chronological sequence, and the link between different data are presented Figure 1. As explained in the subsequent sections, the DCVTS-1 sample can be viewed as a panel sample, which has been surveyed face-to-face a maximum of four times (not all the households participated in all the rounds) prior to administration of DCVTS-1. On the other hand, DCVTS-2 sample was interviewed for the first time in the second round of the telephone survey (DCVTS-2).

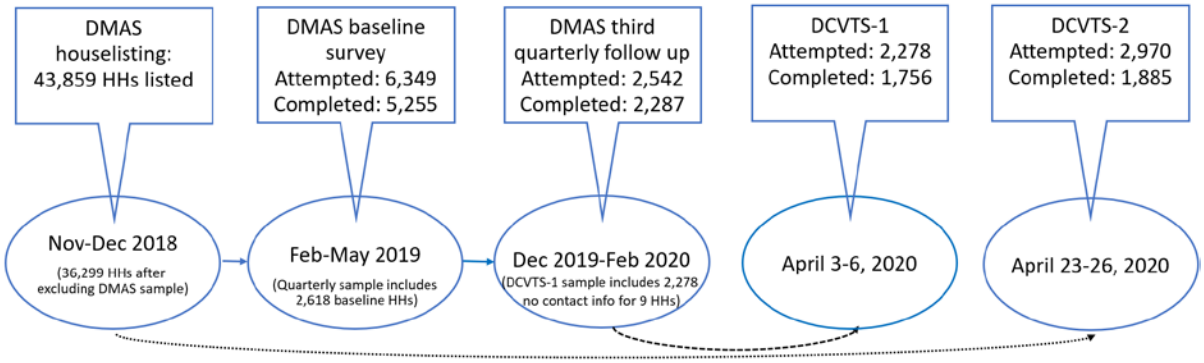
2.2. The Study Setting and Participants

The target geographical area for DCVTS is NCR, which comprises 31 districts, including 13 districts from Haryana, 9 from Delhi, 2 from Rajasthan, and 7 from UP (National Capital Region Planning Board 2017). In other words, NCR is a highly diverse region, encompassing the metropolitan areas of Delhi as well as the urban and rural areas of Haryana, Rajasthan, and UP. The sample for DCVTS was drawn from an ongoing study, namely, the Delhi Metropolitan Area Study (DMAS), a flagship NCAER study undertaken by NCAER National Data Innovation Centre for conducting methodological experiments in data collection. During the period February-May 2019, a sample of 5,255 households from NCR were interviewed face-to-face for the baseline survey of DMAS. The DMAS baseline sample was drawn using a three-stage stratified cluster sampling design. The goal was to select representative random samples for each stage of selection of the districts, clusters, and households. Clusters are defined as Census villages in rural areas and National Sample Survey (NSS) urban blocks in urban areas.

Half of the baseline households were randomly assigned to receive high-frequency quarterly follow-up surveys as a part of DMAS experiments. It may be noted that the quarterly samples include approximately half the baseline households from each of the DMAS sampled clusters. Three quarterly follow-up surveys were completed between June 2019 and February 2020 for these households. The fourth and the last quarterly follow-up started in March 2020, but the data collection was suspended in mid-March due to reports of Coronavirus cases and the subsequent lockdown in the country. Since it was not possible to conduct face-to-face surveys for the original DMAS experiments, we used the available resources for conducting telephone surveys to understand the impact of the Coronavirus pandemic on people's lives and livelihoods.

For DCVTS, Round 1 (DCVTS-1, April 3-6, 2020), the quarterly sample of DMAS that had been followed-up multiple times in the past was used. For DCVTS, Round 2 (DCVTS-2, April 23-26, 2020), we randomly selected new households from the same set of DMAS clusters using houselisting data originally collected for the purpose of constructing a DMAS baseline sampling frame, after excluding the DMAS sample from the frame. It may be noted that the DCVTS-2 sample includes households from each of the DMAS baseline sampled clusters. House-listing entails preparing a list of all households in a particular cluster along with a few background characteristics and phone numbers (when available). The houselisting helped in providing a sampling frame for selection of both the DMAS and DCVTS-2 households.

Figure 1: The Five Different but Related Data Sources Used for Analysis



2.3. Primary Outcomes and Statistical Analyses

The primary outcomes of interest are categorical or binary in nature, for example, e.g., whether the household responded to the survey or completed the survey, the household's State of residence (Delhi, Haryana, Rajasthan, or UP) and area of residence (rural or urban), and whether the household belongs to a particular socio-demographic and economic group. Among the socio-demographic groups, we have considered caste and religion. We considered the following three categories for characterising the caste of households in NCR: 1) General or Forward Caste, 2) Other Backward Caste (OBC), and 3) Scheduled Castes and Scheduled Tribes (SCs and STs). The SCs and STs are various officially designated groups, recognised in the Constitution of India as historically disadvantaged people (NCST 2006), though ST households are rarely found in NCR. For religion also, we considered three categories as follows: 1) Hindu, 2) Muslim, and 3) Other religious groups, including Sikhs, Christians, and those not following any religion.

In order to define a household's economic status, a Wealth Index was constructed for each household using the DMAS baseline data. The Wealth Index is a widely accepted measure of a household's long-term economic status (Rutstein and Johnson 2004; Filmer and Pritchett 2001). For constructing the Wealth Index, we considered variables related to housing conditions, sanitation facility available to the household, and asset possession. Each variable has been assigned a weight generated through Principal Component Analysis (PCA) and the standardised variables are multiplied by the weights and added to produce the Wealth Index (Pramanik et al. 2018). The first principal component, explaining 29.4 per cent of the total variation in the data, was considered as the Wealth Index. Based on the Wealth Index, the households were divided into quintiles as follows: poorest, poorer, middle, richer, and richest. Note that the household's economic status was available only for the panel telephone sample DCVTS-1, for both the respondent and nonrespondent groups, as it was calculated using the DMAS baseline sample, which was the sampling frame for DCVTS-1. On the other hand, the DMAS houselisting data was the sampling frame for the DCVTS-2 sample and the DMAS houselisting exercise did not collect variables that had been used to construct the Wealth Index. Hence, the household's economic status is not available for the nonrespondent households in DCVTS-2.

Elementary statistical analyses were performed to examine the representativeness of the telephone samples collected using two different study designs, viz., panel survey and repeated cross-sectional survey. All analyses are unweighted and involve calculation of proportions. The association between the response status and the

household background characteristics was tested using Chi-squared association test having a 5 per cent level of significance. If the p-value of the test turned out to be less than 0.05, the corresponding association was labelled as significant in the subsequent figures.

3. Results

3.1. Response Rates in Telephone Survey Rounds during the Lockdown

Table 1: Distribution of Different Call Disposition Codes across Two Rounds of DCVTS

	DCVTS-1 (Panel)		DCVTS-2 (Non-Panel)	
Disposition Codes	**n**	**%**	**n**	**%**
Complete interviews	1,756	77.3	1,885	63.5
Refusal	113	5.0	118	4.0
Other incomplete interviews	129	5.7	297	10.0
Non-contact*	280	12.2	670	22.6
Total Attempted Phone Numbers	**2,278**	**100**	**2,970**	**100**

Note: * Non-contact can happen due to various reasons: wrong phone number, number being out of service, phone switched off, out of coverage area, incoming facility not available on the number called, and phone rang but no one answered.

As shown in Table 1, DCVTS-1 achieved a reasonably high response rate (77 per cent) for a telephone survey perhaps because this sample consisted of households with whom we had a pre-existing rapport (Himelein et al. 2020) and also because we had updated the phone numbers collected during the follow-up visit about 3-4 months before the commencement of DCVTS-1. Overall, the response rate for DCVTS-2 is significantly lower (64 per cent) as compared to DCVTS-1 owing to a higher non-contact rate (23 per cent versus 12 per cent) due to non-working phone numbers in the sampled DCVTS-2 households, which were collected during DMAS houselisting, about 15-16 months before DCVTS-2 was carried out. It is quite common for people in many developing countries to change their phone numbers frequently to take advantage of special offers given by different service providers (Himelein et al. 2020). The same holds true for India as well.

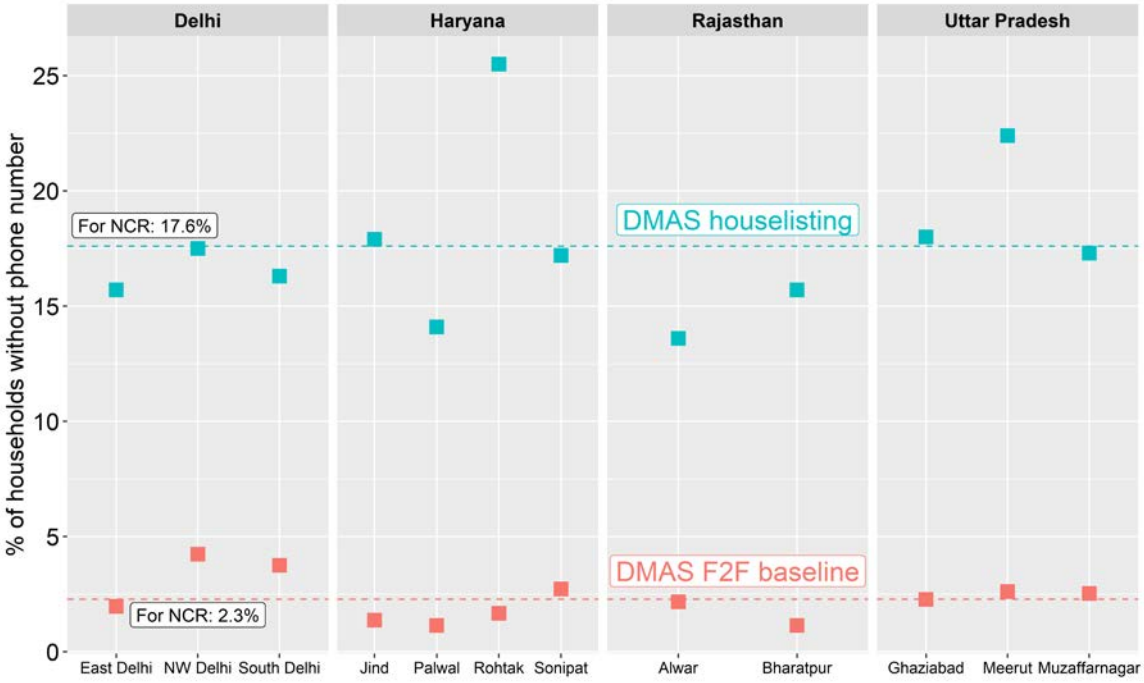
For defining the response rate, we have used a conservative approach and considered only complete interviews in the numerator. The denominator includes complete interviews, partial interviews, break-offs, refusals, and non-contacts. All the phone numbers were eligible as they belonged to the target households. Our definition of the 'response rate' conforms to the Response Rate 5 (RR5) of AAPOR standard definitions of response rates (AAPOR 2016). The response rates observed in the DCVTS rounds is much higher than that observed in telephone surveys in developed

countries in recent years (American Statistical Association 1997; Curtin et al. 2000; 2005; O'Toole et al. 2008; Bladon 2009; Dutwin and Lavrakas 2015).

Exclusion due to Lack of Phone Access or Unavailability of Phone Numbers

The selection of the original DMAS sample was not linked to households having a phone or providing a phone number since the DMAS baseline survey was conducted through face-to-face interviews. Since the DCVTS-1 sample was a follow-up of the DMAS baseline and quarterly sample, it is unlikely to exhibit any coverage bias arising from lack of access to mobile phones. For DCVTS-2, we selected the sample from the original DMAS houselisting data (after excluding the DMAS sample). Since DCVTS-2 was a phone survey, we excluded households without any phone number from the sampling frame. This exclusion may have introduced some degree of coverage bias in the DCVTS-2 estimates because we did not have phone numbers for about 18 per cent of the households in the listing data.

Figure 2: Proportion of Households without Any Phone Number across NCR Districts: DMAS House-listing Data and DMAS Baseline Face-to-face Survey



As shown in Figure 2, the proportion of households that did not have any phone number was much higher in houselisting data (the sampling frame for DCVTS-2) as compared to the DMAS baseline survey (the sampling frame for DCVTS-1). Moreover, the phone numbers collected during the face-to-face baseline survey had been updated regularly in the face-to-face quarterly follow-up surveys. These updated phone numbers were used to call the households during DCVTS-1. The absence of phone numbers in houselisting data does not necessarily reflect a household's lack of access to a telephone. There can be multiple reasons for this, such as:

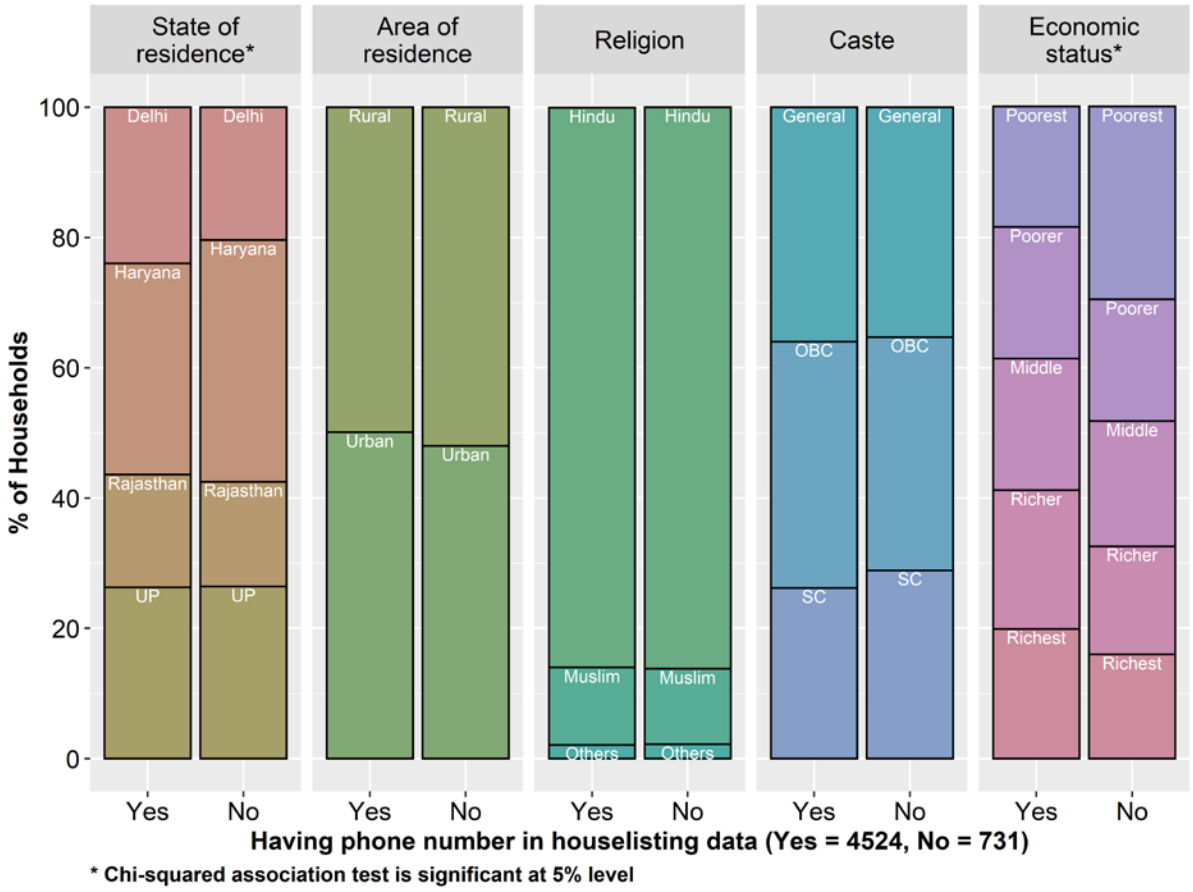
(1) *Unwillingness to share phone numbers*: House-listing was our first and rather brief interaction with the household, which perhaps explains the reluctance of some households to share their phone numbers;

(2) *Low priority to the collection of phone numbers*: Collecting phone numbers during the listing exercise was not a priority as the original plan had been to conduct face-to-face surveys;

(3) *Source of houselisting data*: In 20 per cent of the households, the sources of houselisting data were relatives, neighbours, friends, or children of the household. Among the households where the listing data was obtained from an adult household member, the proportion of missing phone numbers was relatively low, at 12 per cent. On the other hand, among households where the listing respondent was not a household adult member, the proportion of missing phone numbers was significantly higher, at 42 per cent.

The proportion of households without a phone number in the houselisting data is quite substantial and, hence, further analysis is needed to understand the potential magnitude of coverage bias in the DCVTS-2 estimates by comparing the key background characteristics of the households having a phone number with those not having any phone number.

Figure 3: Comparison of Household-level Characteristics across Households Having Phone Numbers and Those without Phone Numbers in House-listing Based on DMAS Baseline Data



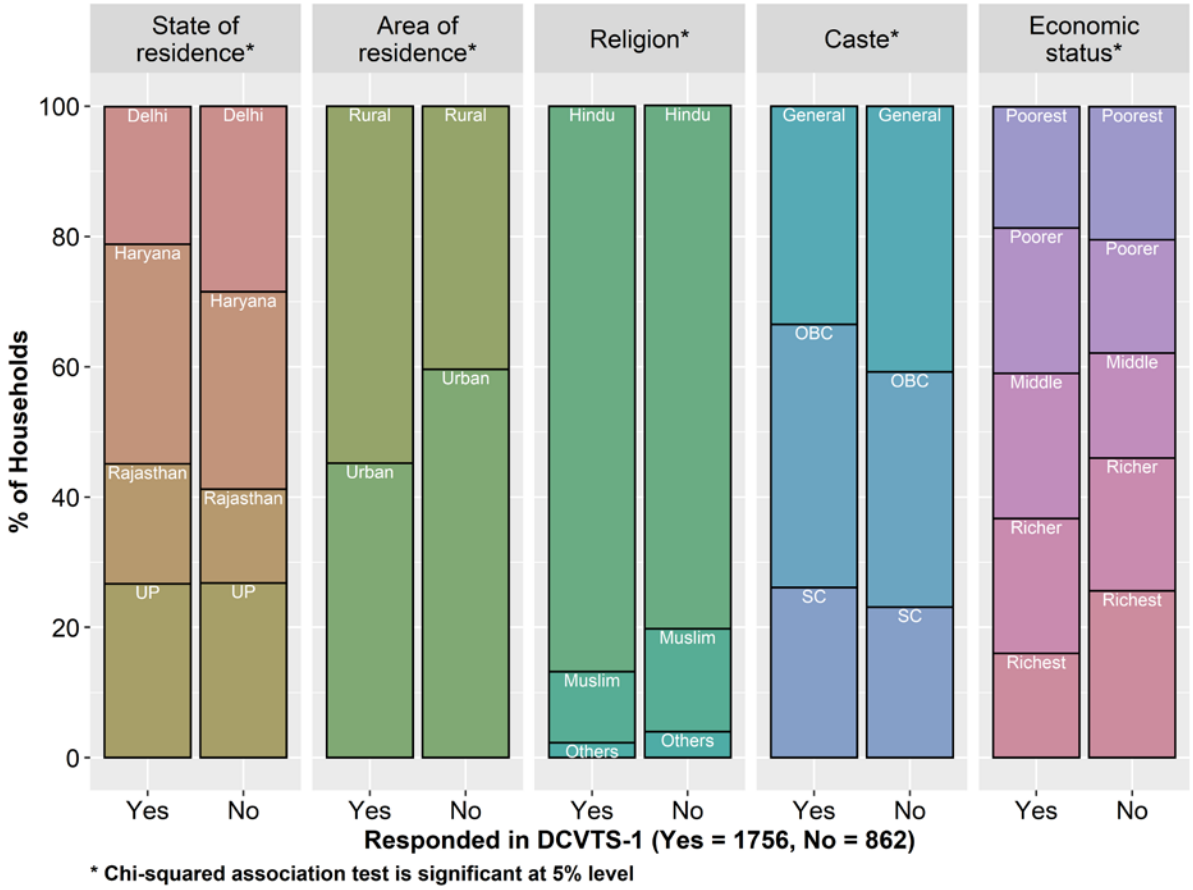
We used DMAS baseline data to compare two groups: one group includes households having phone numbers in houselisting data (the "Yes" column in Figure 3) while the

other consists of households without phone numbers (the “No” column in Figure 3). It may be noted that the latter group of households were excluded from the sampling frame before selection of the DCVTS-2 sample as they did not have the phone numbers needed to contact them for the telephone survey. Figure 3 compares the two groups based on the following five key characteristics of the households: 1) State of residence, 2) area of residence, 3) religion, 4) caste category, and 5) wealth quintile. Among the excluded households, the proportion of households residing in the NCR part of Haryana is significantly higher (but lower in Delhi) as compared to that of the group of households having phone numbers. Similarly, the proportion of households belonging to the poorest wealth quintile is significantly higher in the excluded group as compared to the group of households included in the sampling frame. The Chi-squared tests indicate that coverage of the DCVTS-2 sample is significantly associated with the State of residence and economic status of the households.

3.2. Sample Selectivity due to Cumulative Attrition across the Panel Rounds

The DMAS quarterly sample was the sampling frame for DCVTS-1. We achieved a reasonably good response rate (77 per cent) in DCVTS-1. However, because of the panel nature of the DCVTS-1 sample, some households dropped out of the panel at every prior wave, which could have made the DCVTS-1 sample less representative of NCR. Hence, it is not only the nonresponse during DCVTS-1, but also panel attrition which could contribute to sample selectivity.

Figure 4: Comparison of Household-level Characteristics across Respondents and Nonrespondents of DCVTS-1 Based on DMAS Baseline Data

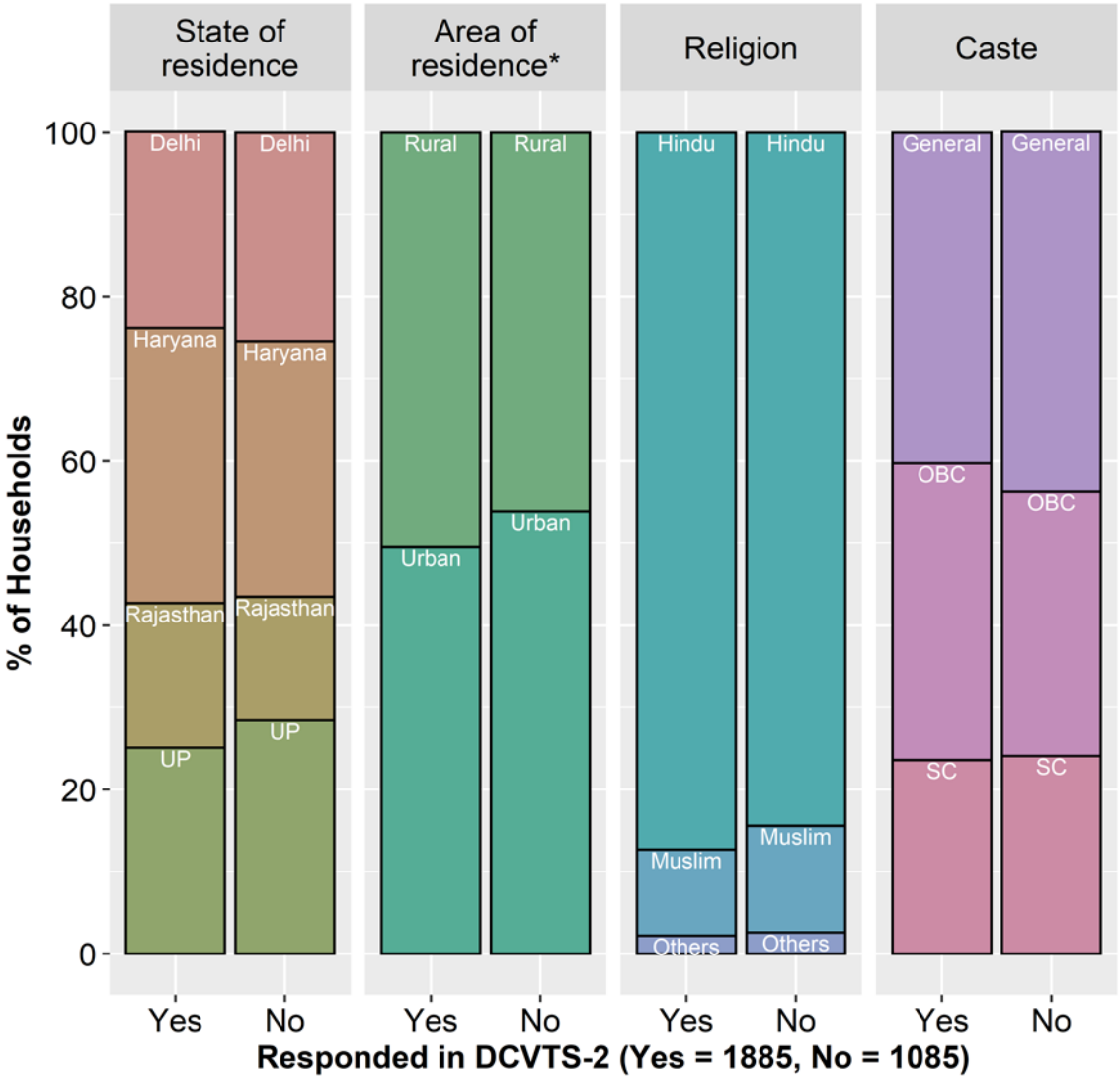


In order to understand the extent of bias due to cumulative attrition across prior waves of the DMAS quarterly survey and nonresponse in the DCVTS-1 sample, we used DMAS baseline data to compare respondents (the "Yes" column in Figure 4) and nonrespondents (the "No" column in Figure 4). The nonrespondents include the households that were part of the DMAS quarterly survey and participated in the DMAS baseline but were not part of the DCVTS-1 respondent sample. The underlying assumption behind comparing the two groups is that the DMAS quarterly sample is representative of NCR. Figure 4 demonstrates that the distributions of respondents and nonrespondents are very different across all the five background characteristics. Among the nonrespondents, the proportion of households residing in Delhi is significantly higher as compared to that of the group of households that completed DCVTS-1. Similarly, the proportion of households belonging to urban areas, Muslim religion, Forward caste, and the richest wealth quintile is significantly higher among the nonrespondents as compared to the respondents. The difference in the characteristics between the respondents and nonrespondents is most pronounced in the context of the area of residence. The rural-urban distribution for respondents is 55 per cent versus 45 per cent as compared to the corresponding figures of 40 per cent versus 60 per cent for the nonrespondents.

3.3. Sample Selectivity due to Non-response

Earlier we discussed the issue of undercoverage of the DCVTS-2 sampling frame. Here we focus on the potential bias due to nonresponse. The nonresponse rate was quite high (36 per cent) for DCVTS-2, which warrants further investigation for the potential nonresponse bias in DCVTS-2 estimates. However, 23 per cent of the nonresponse cases were due to non-contact because of non-functional phone numbers. Among those whom we could contact, the response rate was 82 per cent.

Figure 5: Comparison of Household-level Characteristics across Respondents and Nonrespondents of DCVTS-2 Based on DMAS House-listing Data



We used DMAS houselisting data to compare respondents (the "Yes" column in Figure 5) and nonrespondents (the "No" column in Figure 5). It may be noted that the DCVTS-2 nonrespondents never participated in the DMAS surveys. Hence, we resorted to the use of houselisting data for comparison. On the basis of Figure 5, we can say that the proportion of households residing in UP belonging to urban areas, Muslim religion, and the Forward caste is significantly higher among the

nonrespondents as compared to the respondents. However, only the area of residence (rural or urban) is associated with the nonresponse mechanism, at a 5 per cent level of significance as inferred from the Chi-squared association test.

4. Discussion

Sample surveys are cost-effective scientific tools for drawing inference at the population level, if designed properly. Survey estimates of unknown population parameters are often subject to different types of errors. Traditionally, survey data analysis measures the sampling error but ignores other non-observation errors such as coverage and nonresponse bias. This paper addresses this gap in the context of telephone surveys carried out in India during the COVID-19 lockdown when these two sources of non-observation errors were of grave concern.

In the context of a panel design, we explore whether the cumulative attrition over panel rounds leads to greater sample selectivity (less representativeness) in the telephone sample relative to a freshly selected telephone sample, though the latter is often subject to a relatively higher overall nonresponse rate in the absence of updated phone numbers.

The magnitude of undercoverage and nonresponse bias, often associated with a specific outcome variable, depends on the following two components: (1) Exclusion or nonresponse rate, and (2) difference between the included (respondents) and the excluded units (nonrespondents) (Groves et al. 2011). Although the estimates of (1) are often readily available, data on (2) are difficult to obtain. For example, the nonresponse rate is usually known for a survey, but the characteristics of nonrespondents may not be known. We used other available data sources to compare the characteristics of the included and excluded units in the sampling frame as well as the respondents and nonrespondents of the survey. We did not focus on specific outcome variables while comparing the two groups but instead examined the background characteristics of households which are often correlated with most outcome variables.

We had the unique opportunity to investigate the non-observation errors under two different study designs. For DCVTS-1, a panel sample of households that had been followed up multiple times in the past was used. For DCVTS-2, we used a fresh sample using a repeated cross-sectional design as opposed to following the DCVTS-1 sample. The use of a repeated cross-sectional design was meant to reduce the respondent burden (Graf 2008) as the DCVTS-1 sample had been surveyed multiple times in the past. Moreover, it precludes the possibility of the Hawthorne effect (observer effect) in the questions pertaining to the knowledge, attitude, and perceptions associated with Coronavirus (McCarney et al. 2007; McCambridge et al. 2014; Das and Leino 2011).

Because we had established a pre-existing rapport with the respondents and consequently had their updated phone numbers, we achieved a response rate of 77 per cent in DCVTS-1, which is a reasonably high percentage for a telephone survey relative to the corresponding response rates achieved in developed countries. On the other hand, a proportion of households dropped out of panel in every prior round of the DMAS quarterly survey while some did not respond to the DCVTS-1 telephone survey. This cumulative attrition over prior rounds of the DMAS quarterly survey and the nonresponse in DCVTS-1 led to differences in the key socio-economic and

demographic characteristics of households in terms of the distributions of respondents and nonrespondents. Among the nonrespondents, the proportion of households residing in Delhi and other urban areas of the districts in NCR belonging to the Muslim religion, Forward caste, and the richest wealth quintile are significantly higher as compared to the respondents. This clearly suggests that the nature of attrition and nonresponse in the panel surveys is non-random and may lead to biased estimates of the outcome of interest, if not adjusted properly (Groves et al. 2001; Cobben and Bethlehem 2005). However, in the context of studying the impact of the Coronavirus pandemic and the resultant lockdown in the target population, the characteristics of drop-outs in DCVTS-1 do not suggest exclusion of vulnerable groups, except among the Muslim community.

The DCVTS-2 sample suffers from both undercoverage error and nonresponse error. Before selection of the DCVTS-2 sample, we had to exclude 18 per cent of the households that did not have any phone number from the sampling frame. Among the excluded households, the proportion of households residing in the NCR part of Haryana and belonging to the poorest wealth quintile was significantly higher as compared to the group of households included in the sampling frame. A comparison of the DCVTS-2 respondents and nonrespondents shows that the proportions of households residing in UP, belonging to urban areas, Muslim religion, and the Forward caste are higher among nonrespondents as compared to the respondents. However, only the area of residence (rural or urban) is significantly associated with the nonresponse mechanism. Although the nonresponse rate was quite high (36 per cent) for DCVTS-2, about 23 per cent of the nonresponse cases were due to non-contact because of non-functional phone numbers. The findings indicate that nonresponse resulting from non-functional phone numbers may happen at random and may not be associated with the socio-demographic characteristics of households.

Our findings suggest that the representativeness of telephone survey respondents may vary depending upon the study design and the quality of the sampling frame. In this paper, we have only examined the issue of sample representativeness. The response quality in phone surveys vis-à-vis face-to-face surveys deserves future attention.

References

AAPOR. 2016. "Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys." Ninth Edition. Kansas, USA: The American Association for Public Opinion Research.

American Statistical Association. 1997. "Response Rates in Random Digit Dialing (RDD) Telephone Surveys." Proceedings of the American Statistical Association, Section on Survey Research Methods. Citeseer.

Bladon, T.L. 2009. "The Downward Trend of Survey Response Rates: Implications and Considerations for Evaluators." *The Canadian Journal of Program Evaluation*, 24(2): 131.

Cobben, F. and J. Bethlehem. 2005. "Adjusting Under-coverage and Non-response Bias in Telephone Surveys", *Discussion Paper 05006*, CBS, Statistics Netherlands.

Curtin, R., S. Presser, and E. Singer. 2000. "The Effects of Response Rate Changes on the Index of Consumer Sentiment." *Public Opinion Quarterly*, 64(4): 413-428.

———. 2005. "Changes in Telephone Survey Non-response over the Past Quarter Century." *Public Opinion Quarterly*, 69(1): 87-98.

Das, J. and J. Leino. 2011. "Evaluating the RSBY: Lessons from an Experimental Information Campaign." *Economic and Political Weekly*, XLVI(32): 85-93.

Dutwin, D. and P. Lavrakas. 2016. "Trends in Telephone Outcomes, 2008-2015." *Survey Practice*, 9(3): 1.

Filmer, D. and L.H. Pritchett. 2001. "Estimating Wealth Effects without Expenditure Data—or Tears: An Application to Educational Enrollments in States of India." *Demography*, 38(1): 115-132.

Graf, I. 2008. "Respondent Burden", in: P.J. Lavrakas (ed.), *Encyclopedia of Survey Research Methods*. California, USA: Sage Publications.

Groves, R.M., D.A. Dillman, J.L. Eltinge, and R.J.A. Little. 2001. *Survey Non-response*. New York: Wiley.

Groves, R.M., F.J. Fowler Jr., M.P. Couper, J.M. Lepkowski, E. Singer, and R. Tourangeau. 2011. *Survey Methodology*, Hoboken, New Jersey: John Wiley & Sons.

Himelein, K., S. Eckman, C. Lau, and D. McKenzie. 2020. "Mobile Phone Surveys for Understanding COVID-19 Impacts: Part I Sampling and Mode." *World Bank Blogs*.

McCambridge, J., J. Witton, and D.R. Elbourne. 2014. "Systematic Review of the Hawthorne Effect: New Concepts Are Needed to Study Research Participation Effects." *Journal of Clinical Epidemiology*, 67(3): 267-277.

McCarney, R., J. Warner, S. Iliffe, R. van Haselen, M. Griffin, and P. Fisher. 2007. "The Hawthorne Effect: A Randomised, Controlled Trial." *BMC Medical Research Methodology*, 7(1): 30.

National Capital Region Planning Board. 2017. Annual Report, 2016-17. New Delhi: Ministry of Housing and Urban Affairs, Government of India.

NCAER National Data Innovation Centre. 2020a. "Delhi NCR Coronavirus Telephone Survey, Round 1 (April 3-6): Preliminary Report." New Delhi: National Council of Applied Economic Research. Available at:

https://www.ncaer.org/image/userfiles/file/NDIC-TEL/DCVTS%20Results%202020-04-12-FINAL.pdf, Accessed on 10 March 2022.

——— 2020b. "NCAER National Data Innovation Centre. Delhi NCR Coronavirus Telephone Survey- Round 2 (April 23-26): Preliminary Report." New Delhi: National Council of Applied Economic Research. Available at: https://www.ncaer.org/image/userfiles/file/NDIC-TEL/Round-2/NCAER%20May%201%202020%20DCVTS-2%20Presentation.pdf, Accessed on 10 March 2022.

NCST. First Annual Report, 2004-2005 and 2005-2006. New Delhi: The National Commission for Scheduled Tribes, Government of India. Available at: http://www.ncst.gov.in/content/first-annual-report-ncst. Accessed on 10 March 2022.

O'Toole, J., M. Sinclair, and K. Leder. 2008. "Maximising Response Rates in Household Telephone Surveys." *BMC Medical Research Methodology*, 8(1): 71.

Peytchev, A., L.R. Carley-Baxter, and M.C. Black. 2011. "Multiple Sources of Nonobservation Error in Telephone Surveys: Coverage and Nonresponse." *Sociological Methods & Research*, 40(1): 138-168.

Pramanik, S., A. Ghosh, R.B. Nanda, M. de Rouw, P. Forth, and S. Albert. 2018. "Impact Evaluation of a Community Engagement Intervention in Improving Childhood Immunization Coverage: A Cluster Randomized Controlled Trial in Assam, India." *BMC Public Health*, 18(1): 534.

Rutstein, S.O. and K. Johnson. 2004. "The DHS Wealth Index: ORC Macro", MEASURE DHS.