Using Network Sampling in a Telephone Survey to Locate Non-Telephone Households

Nadra Garas
American University

Johnny Blair
Abt Associates

KEY WORDS: network sampling, frame coverage, multiplicity estimation, telephone survey, survey methodology, non-telephone households

Reaching Non-Telephone Households

An inherent limitation of general population telephone surveys is the exclusion of non-telephone households, and the potential bias that may result. Often this bias is simply ignored, or the survey population defined to exclude non-telephone households. In other cases, statistical adjustments, based on interruptions in service or other models, are used to address this problem. But, in general, the only way to actually include non-telephone households in the survey is to locate them using a different frame and expensive screening procedures. In this research, we explored the possibility of locating non-telephone households in a telephone survey by using network sampling.

Many people without telephones may still arrange to have access to telephones for emergencies, contact with their child's school or for simple convenience. Such people, without telephones in their own households, may have access to a telephone in the home of a relative, neighbor, landlord or someone else who lives nearby. If this conjecture is correct, it may be possible via network sampling to take advantage of the linkages between these non-telephone households and the households whose telephones they use.

In most household surveys, respondents' reporting is limited to their own residence. Network sampling uses a different reporting rule. Households in the selected sample are conceptually linked to other households. Sample households are then asked to report about the households to which they are linked. However, to maintain a probability sample, the inclusion probabilities of non-telephone households identified in this manner must be known (Sudman, Sirken and Cowan 1988).

For network samples to produce unbiased results, the following conditions are necessary. First, the

linkages (called "counting rules") must be defined in such a way that probabilities of selection can be determined. Second, sampled households must be able and willing to report about households to which they are linked. Third, if the nominated households are to be interviewed, the sample household must be able and willing to provide contact information. In order to determine whether network sampling is feasible for a particular application, a pilot test is usually necessary to determine whether or not these condition hold (Blair 1990).

The probabilities of inclusion in network sampling typically are unequal. However, if the network size of each nominated household can be determined, then weights can be constructed to account for the different inclusion probabilities.

As noted, the network sample does not produce an epsem (equal probabilities of selection method) sample. There are M possible reporting households, i.e. the respondents from the selected RDD sample. Each nominated (non-telephone) household is linked to m possible nominating households, [$m = 0, 1, 2, 3....$]. Each non-telephone household has an inclusion probability $p_i = m_i / M$, and a base weight of $\omega_i = 1 / p_i$.

If the (typically remote) possibility is ignored that a non-telephone household might be nominated more than one, then a multiplicity estimator of the population total given by Thompson (1992) simplifies to:

$$T = M/n \sum_i y_i / m_i$$

Study Description

In this study, the network counting rule required telephone survey respondents to report whether, during the survey data collection period, anyone used their telephone *because they did not have telephone service themselves*. Therefore, the non-telephone

households would need to report how many different households they used for telephone service during the data collection period.

The question asked of sample households was [a]:

> In the past two months, have any friends, neighbors or relatives, outside of your household, used your telephone *because they do not have phone service themselves*? {The interviewers were instructed to emphasize the italicized phrase}
>
> If YES,
>
> How many such people have used your phone?

If more than one person from another household had used the respondent's telephone because of lack of telephone service, then the person who used their telephone most recently was selected as the nominated household.

In three surveys– two national and one statewide– conducted in1997, 1999, and 2001, about ten percent of sampled telephone households reported such use of their telephones (see Table 1).

Table 1 Households reporting that others have used their telephones		
	Number	Percent
Study A	134	13.8
Study B	108	10.8
Study C	98	9.8

Attempts to contact the identified households were less successful. However, given the typical costs to include some coverage of non-telephone households, this methodology seems worth further examination. We report the results of these three studies as well as plans for further research.

First, we report the findings for the basic network question in the three studies. Each study was a random digit dial (RDD) telephone survey of the general population of adults, age 18 or older. Each study was an omnibus survey, consisting of questions on a range of topics submitted by different sponsors. All three surveys were conducted at the Survey Research Center at the University of Maryland, College Park.

Survey A was a national omnibus conducted in 1997. A sample of 3,836 telephone numbers was generated. Of these, 1,766 were identified as households. Of these households, 55% agreed to the interview, 27% refused, 11% were noncontacts, and the remaining 8% were miscellaneous problems. (Telephone numbers never answered after 20 callbacks were assumed to be nonresidential.) Of those households contacted, 67.2% agreed to the interview. Within sample households, a respondent was chosen using the Next Birthday method. A total of 972 interviews were completed. Data were weighted for number of voice telephone lines in the household and for the number of adult residents. Post-stratification weights for sex, age, education, race and geographic region were also used.

Survey B was, like survey A, a national omnibus, and employed the same methodology as that survey. Survey B was conducted in 1999. From 2,979 telephone numbers, 1,766 were identified as households. Of these households, 57% agreed to the interview, 21% refused, 18% were noncontacts, and the remaining 4% were miscellaneous problems. The cooperation rate was 72%.

Survey C was a statewide Maryland survey, conducted in 2001, which employed the same methodology as surveys A and B. A sample of 3,073 numbers yielded 1,840 identified households. Of these 1,005 agreed to the interview for a 55% response rate. There were 22% refusals, 19% noncontacts, and 4% other problems. The cooperation rate was 71%.

Use of Respondents' Telephone By Nontelephone Households

In survey A, conducted in 1997, 13.8% of respondents (n=134) reported use of their telephone by someone from a nontelephone household. This result was replicated in the 1999 national survey (B), where 10.8% (n=108) respondents reported that someone from outside their household used their telephone.

In survey C, a state of Maryland survey, similar results were obtained, with approximately 9.8% (n=98) of respondents reporting that someone used their telephone because that person did not have telephone service in their own household.

These findings indicate that, on average; about ten percent of households are linked, by this counting rule, to nontelephone households.

If a high percentage of these nominated households can be contacted and interviewed, this would be a valuable addition to the basic telephone sample. These households could be used to improve the estimates in a given study. While it is encouraging that the network counting rule linked a nontrivial number of sampled households to nontelephone households, there was much less success in actually contacting the nominated households.

The validity of these findings is indirectly supported by the demographic characteristics of households reporting use of their telephones (Table 2). As would be expected, lower income households are much more likely to report use of their telephones than are higher income households, which seem logical. The percentage of households reporting linkage to nontelephone households is inversely related to household income. As household income of the respondent declines, the percentage of respondents reporting someone using their telephone increases.

Table 2
Income Distribution and Other Telephone Use

	Study A*		Study B		Study C#	
	N	%	N	%	N	%
Less than $12,000	11	24.4	9	22.5	3	15.8
$12,001-$20,000	19	26.0	10	14.7	7	20.6
$20,001-$30,000	27	20.5	16	12.0	8	12.5
$30,001-$50,000	22	11.3	22	10.0	23	11.9
$50,001-$75,000	13	10.2	17	9.8	16	10.1
$75,001-$100,000	9	14.8	4	4.2	7	5.6
Over $100,001	0	0.0	3	3.2	10	5.9
	101		81		74	

*p<0.01   #p<0.05

Race

Table 3
Race and Other Telephone Use

	Study A*	Study B*	Study C*
White	73.3	71.3	36.7
Black	18.8	19.4	59.2
Asian	0.0	0.0	1.0
Other	7.5	7.4	3.1
Refused	0.0	1.9	0.0

* P<0.05

Hispanic Origin

For Study B, about 22% of all Hispanic origin households reported that some other person used their telephone, compared to only 9.9% of all households reporting that they were not of Hispanic origin. For study A, only 19.6 % of all Hispanic households reported that someone else uses their telephone, while only 13.9% of all non-Hispanic households indicated that someone else used their telephone in the last two months.

For the Maryland study, the percentage of Hispanic households that reported some other person using their telephone was only 2.5% of all Hispanic households. This compares with around one quarter of the households in the national surveys. However, the percentage of the non-Hispanic households in the Maryland study was slightly higher than the national study conducted in 1999, but about three percentage points lower than the 1997 national study.

Table 4
Hispanic Origin and Other Telephone Use

	Study A	Study B*	Study C*
Hispanic	19.6	21.5	2.5
Non-Hispanic	13.9	9.9	10.2

* P<0.05

Reaching Nominated Households

In survey A, a direct approach was used. For each nominated household, the respondent was asked:

We'd also like to include [nominated person] in this survey. Would you please permit him/her to be interviewed on your phone?

We then tried to set up a time to reach the nominated person, and, if refused, we asked that the respondent leave the Survey Research Center's 1-800 number by the phone for the nominated person to use[b].

Would you please leave the Survey Research Center 1-800 number by the phone so when [nominated person] comes by next time, he/she can call us for an interview?

Please tell [nominated person] to mention your phone number when he/she calls us back.

More than half (62%) of respondents reported willingness to let the nominated person use their telephone. Over two thirds of the respondents (67%) who indicated that someone else uses their telephone and who refused to allow that person to be interviewed on their telephone, said they were willing to give the nominated person the SRC 1-800 number. Only a trivial number of nominated persons were ever reached.

In survey B (1999), a different approach was tried, in addition to the offer of the 1-800 number. Respondents nominating a nontelephone household were asked:

> People without telephones are often missed in surveys. We're making an effort to include such people in this survey. Would you mind giving me [nominated person's] mailing address?

This was even less successful than the original appeal in survey A. Less than a third, 27.6%, of respondents provided a mailing address. However, about two thirds of the respondents (67.8%) who refused to provide the nominated person's address agreed to leave the SRC toll free number by the telephone so that this person could call for an interview. And again, only a trivial number of nominated persons were ever contacted.

In survey C, a small incentive was tried to obtain nominated persons' mailing addresses:

> People without telephones are often missed in surveys. We're making an effort to include such people in this survey. So we are going to send you a 30-minute phone card and give [nominated person] another 30-minute phone card if he or she calls us to complete an interview.
>
> Can you tell me [nominee's] mailing address?

Slightly fewer than half the respondents (43.7%) were willing to provide the address of the nominated persons. However, the vast majority of the respondents who refused to provide the address of the nominated persons (98.9%) indicated that they would be willing to leave the toll free number by the telephone so that the person could call when he or she next came to the household. As was the case with the two national studies, a trivial number of nominated persons were ever contacted.

Discussion of Results and Future Research

Applicability of network sampling and multiplicity estimation to locating nontelephone households in a telephone survey may be a viable, and cost effective, approach if the practical obstacles to reaching the nominated households can be overcome.

In particular, we think that the use of monetary incentives may be effective. However, since promised incentives are generally less effective than pre-paid incentives, the amount of the incentive may have to be moderately large. Still such a cost may be justified if a high percentage of nominated households can thus be included in a telephone survey.

References

Blair, Johnny. (1990) "Improving Data Quality in Network Surveys of Rare Populations" in Gunar Liepins and V. Uppuluri eds: Data Quality Control: Theory and Pragmatics, Marcel Dekker.

Sudman, S., Sirken, M.G., and Cowan, C.D. (1988) Sampling rare and elusive populations. Science 240, 991-996.

Thompson, Steven K. (2002) Sampling, Wiley Series in Probability and Statistics. 2nd Ed. Wiley Interscience.

---

[a] This version of the question was used for simplicity of administration. Note that it does not ensure that the nominated household was without telephone service during the precise data collection period of the survey; this would need to be verified when a nominated non-telephone household was contacted.

[b] For pilot test purposes, the objective was to locate the individual person who used the survey respondent's telephone. The issues of how to implement random respondent selection within nominated households was not addressed in this research.