

A MINIMALLY INTRUSIVE METHOD FOR SAMPLING PERSONS IN RANDOM DIGIT DIAL SURVEYS

LOUIS RIZZO

J. MICHAEL BRICK

INHO PARK

Westat

Introduction

To survey a population of persons such as all adults, many studies sample households as an intermediary step to reach this population because directly sampling persons requires a comprehensive list that exists only in specialized situations. In random digit dial (RDD) surveys, the sample of households is obtained by sampling telephone numbers. Within-household sampling is then performed to go from the household to the person level.

From the earliest days of conducting RDD surveys, a key research objective has been the development and evaluation of methods of within-household sampling that do not lower response rates. Some sampling methods that do not give probability samples have been proposed and used in practice. Although these methods are subject to selection bias, their use has been justified as necessary to increase response rates by simplifying and improving the flow of the interview.

We propose a new and simple method of within-household sampling that is minimally intrusive and, thus, should not have a deleterious effect on response rates. The method results in a probability sample of eligible persons within households. The method is especially simple to implement in computer-assisted telephone interview (CATI) RDD surveys. We focus on sampling one adult from each household as this is the primary application, but we also describe how the method can be extended to sample more than one person per household. This paper provides an introduction to the method and results from one study. Experimental comparisons of this approach with other methods of within-household sampling would be a useful next step.

Address correspondence to Louis Rizzo; e-mail: lourizzo@westat.com.

Review of Within-Household Sampling Methods

Putting aside issues of nonresponse, burden, and interview flow, there is little controversy that a probability sampling technique such as the Kish method is the preferred approach to selecting eligible persons in the household (Kish 1949, 1965 sec. 11.3). In this method all eligible persons in the household are rostered, and a person is sampled with equal probability from the eligible persons. If the screener or household respondent reports all the eligible household members accurately, then this method results in a probability sample at the person level.

The Kish method was originally developed for sampling in face-to-face interview surveys, and it limited the number of persons who could be listed because of space limitations on the household "cover sheet." When the method was later applied to CATI interviewing, the limit on the number of persons who could be listed was eliminated. However, many survey practitioners consider the method to be burdensome because all household members or all adults are rostered. Furthermore, asking for names of all persons in the household may be intrusive, especially since these questions are asked at the very beginning of the interview, before the interviewer has had time to establish a relationship with the household respondent. There has been conjecture that the Kish method may cause the screener respondent to refuse or to provide inaccurate information. Even short of a refusal, if the respondent feels the questions are intrusive, the method may damage the rapport between the interviewer and respondent.

Because of these concerns, a number of alternative methods have been developed over the past forty years. Trolldahl and Carter (1964) proposed random selection of the "oldest male," "youngest male," "oldest female," or "youngest female" to spread the sample by age and sex without rostering. Only questions about numbers of adults and numbers of men are needed to use this method. Hagan and Collier (1983) proposed a variant of the Trolldahl and Carter method that does not ask at all about the number of adults and the number of men but instead randomly asks the respondent to identify the oldest male, youngest male, oldest female, or youngest female, with alternative procedures in case there are too few males or females. This method avoids the question about numbers of persons. But without knowing the number of persons in each household, any estimator is biased because the probability of selection is unknown if the sampled adult does not complete the survey. Specifically, adults in households with many adults are underrepresented.

Salmon and Nichols (1983) proposed asking the respondent how many eligible persons were in the household,¹ and then asking which person had the most recent birthday. The person with the most recent birthday is sampled,

1. It is possible to sample using the birthday method without asking the number of persons in the household, but this leads to the same types of biases in the estimators as described above for the Hagan and Collier method.

and this method is called the “last-birthday” method. Another version is the “next-birthday” method, which is the same method except the person with the next birthday is selected. Either birthday method relies on the lack of correlation in the population between birth month and the eligible-person characteristics of interest as the source of randomness, and experience shows this is a reasonable assumption for most statistics for samples of adults. A concern with the birthday methods is how often the process is carried out accurately, that is, how often the respondent actually identifies the appropriate person. While, O’Rourke and Blair (1983) found a 90 percent accuracy rate in identifying the appropriate person, Lavrakas and Bauman (1993) and Lavrakas, Stasny, and Harpuder (2000) found roughly a 75 percent accuracy rate for the last-birthday method. Lind, Link, and Oldendick (2000) found an 80 percent accuracy rate with the last-birthday method and 73 percent accuracy rate with the next-birthday method in a randomized experiment with two- and three-adult households. An additional concern that we have noted in our applications of the birthday methods is that the screener respondent is selected more often than would be expected by chance. This self-selection by the screener respondent is likely to result in more women being sampled since women are more likely to answer the telephone and participate in the screening process. Note that these problems are errors in sampling and are not associated with nonresponse.

Many studies have examined the response rates when the various methods are used in RDD surveys. Some have shown that these alternative methods do succeed in yielding higher initial response rates than the Kish method before refusal conversion attempts are taken into account (Binson, Canchola, and Catania 2000; Oldendick et al. 1988; O’Rourke and Blair 1983; and Salmon and Nichols 1983). Oldendick et al. (1988) found little difference after refusal conversion. Binson, Canchola, and Catania (2000) did not indicate whether their results involved any attempt to convert refusing households. Even if these methods do increase the response rates to the screener interview, it is difficult to evaluate the effect of using a method that does not give a strict probability sample or has an intrinsically high error rate in selecting respondents. Presently, no single approach to sampling persons in RDD household surveys is widely accepted as the preferred one.

A New Within-Household Sampling Methodology

The method we propose takes advantage of the fact that about 85 percent of all households in the United States have two or fewer adults. To sample one adult in these households, this method randomly selects either the screener respondent or the other adult.² No questions about the household composition other than the number of adults are needed. In households with more than two

2. We are not aware of any other methods of within-household sampling that take advantage of this fact.

adults, the method first determines if the screener respondent is sampled and thereby offers a procedure with fewer steps than other methods.

The new sampling method is described below. Questions that the screener respondent answers are italicized. (The appendix gives the wording of the screener questions utilized to implement this procedure. All methods of within-household sampling require the first two items listed there.)

1. Ask screener respondent how many adults are in the household. Denote the number of adults N .
 If $N = 1$, then the screener respondent is selected. End selection process.
 If $N > 1$, then randomly sample the screener respondent with probability equal to $1/N$. If the screener respondent is selected, then end selection process.
 If $N = 2$ and the screener respondent is not selected, then tell the screener respondent that the other adult in the household is selected for the study and ask to speak to that person or set up a callback.³ End selection process.
2. If $N > 2$ and the screener respondent is not selected, then use the Kish method or the last-birthday method, excluding the screener respondent from the sampling procedure. For example, with the last-birthday method, *ask the screener respondent to identify the adult in the household other than the screener respondent who had the most recent birthday*. The identified person is sampled, and the selection process is ended.

This new method was first implemented in the 2002 Health Information National Trends Study (HINTS), a large RDD study sponsored by the National Cancer Institute and conducted by Westat (see, for example, Nelson et al. in press). The HINTS collects data on the ways American adults acquire health information. The sample design is a national probability sample of adults. The sample is a list-assisted sample of telephone numbers from sets of one hundred telephone numbers with at least one listed residential number (Tucker, Lepkowski, and Piekarski 2002). The response rate for the screener interview using American Association for Public Opinion Research (2000) definition RR3 was 55 percent.

In each sampled household, one adult was sampled using the new within-household sampling method, where the last-birthday method was the preferred method for sampling if the screener respondent was not selected. Once the sampled adult was selected, that person was asked to answer questions in the extended interview on the means used to acquire health information.

Even though no sampling is required in households with only one adult, we examined the percentage of one-adult households by sex to assess the effect of nonresponse to the screener in these households. Females were overrepresented by about 6 percentage points, compared to the corresponding Current Population

3. Whenever a callback is needed in any method, information such as the name and age of the sampled adult should be obtained.

Table 1. Number and Percentage of Households in which the Screener Respondent Was Sampled or Additional Sampling Questions Were Required

Number of Adults in Household	Number of Completed Screeners	Weighted % of Households	% of Households with Screener Respondent Sampled		% of Households with Additional Sampling
			HINTS	Expected	
1	2,931	29.6	100.0	100.0	0.0
2	5,385	53.8	50.6	50.0	0.0
3	1,222	11.4	34.5	33.3	65.5
4	441	4.0	25.4	25.0	74.6
5 or more	139	1.2	15.5	11.8	84.5
Total	10,118	100.0	61.4	60.9	12.2

Note.—The weighted percentages utilize the HINTS sampling weights of the households.

Survey (CPS) estimate (HINTS estimated that 67 percent of persons in one-adult households were female, while CPS estimated 61 percent). We believe this difference is due to males in one-adult households responding to the HINTS screener at a lower rate than females.

Next, we examined how well the new sampling procedure worked by comparing the percentage of sampled households in which the screener respondent was sampled to the expected percentage. Table 1 presents these percentages as well as the percentage of households in which the additional items for the birthday method were needed.

The table shows that the percentage of households in which the screener respondent was sampled is very close to the expected percentage. The percentages should only differ from the expected due to sampling error. Over all households sampled for HINTS, the screener respondent was selected 61 percent of the time. In these households the only item required for selecting the sample of adults was the number of adults in the household. The last column of the table shows the percentage of households in which additional sampling items were necessary to select an adult. In the two-adult households, no additional items are required (other than asking to speak to the other adult). Overall, additional sampling questions were asked in only 12 percent of the households in HINTS.

The additional sampling items were only asked in households with three or more adults in which the screener respondent was not selected. As noted earlier, HINTS used the last-birthday method to sample in those households where the selection was not otherwise determined. In 6 percent of the households with three adults or more, the screener respondent did not know the birthdays of all the other members; and the Kish sampling method was used. The Kish method had to be used more often as the number of adults in the household

Table 2. Estimated Distribution of Sex of Adults in Two-Adult Households in HINTS

Sex of		Percent of		
Screener Respondent	Sampled Adult	Number of Completed Screeners	Initial Household Respondent	Total
Female	Female	1,801	54.3	
Female	Male	1,516	45.7	
Female	Total	3,317	100.0	62.0
Male	Female	995	49.0	
Male	Male	1,035	50.1	
Male	Total	2,030	100.0	38.0
Total		5,347		100.0

increased (5 percent in three-adult households, 7 percent in four-adult households, and 16 percent in five-or-more adult households).

To ascertain whether the new method of sampling selects too large a percentage of women (a concern with the birthday sampling methods), we examined the outcomes from the HINTS sampling for two-adult households by gender, shown in table 2. In HINTS the sex of the sampled adult was asked, but the sex of the screener respondent was not collected (of course, the sex of the screener respondent was recorded if that person was also the sampled adult). To prepare the tabulation, we imputed the sex of the screener respondent when the data were missing using the CPS estimate of the distribution of sex in two-adult households (the two adults are of the opposite sex in 91 percent of two-adult households in the CPS).

Table 2 shows that even though 62 percent of the screener respondents are estimated to be female, the sampling method resulted in sampling females about half the time, irrespective of the sex of the screener respondent. A female was sampled in 52 percent of the two-adult households in HINTS, which is very close to the expected value of 51 percent estimated from the CPS (there are more two-adult households with two females than two males in the CPS). The HINTS outcome shows that the new sampling method avoids any self-selection bias that might cause females to be overrepresented in the sample prior to dealing with nonresponse to the extended interview.

Discussion

The proposed sampling method has a number of potential advantages over existing methods. First, it is a probability sampling approach and does not appear to result in self-selection biases. The method produces a strict probability sample when the Kish method is used for large households. However,

we believe the last-birthday method works well as the alternative method because concerns about self-selection bias are reduced in this context.⁴ Second, the method limits questions that must be asked to select a person in the screening interview. Only 12 percent of households in HINTS required asking any additional sampling questions other than the number of adults in the household. This procedure is simple and accurate, places a low burden on the household, and avoids intrusive and unexpected questions that might lower response rates.

We have presented the method as a way of sampling one adult from all adults in the household, but it can be applied in any situation in which the screener respondent is eligible by first determining if the screener respondent is sampled. If the screener respondent is not eligible, a variant of the method can be used. The interviewer first asks to speak to any eligible person in the household. It is not necessary to randomly select the eligible person to complete the screener interview since any eligible person will suffice. The sampling method can then be applied directly.

Appendix

THE HINTS SCREENER

Only the questions relevant to adult selection are included.

H-1. May I speak to a member of this household who is at least 18 years old?

H-2. Including yourself, how many people *aged 18 or older* currently live in this household?

If screener respondent is selected, go to H-6a. If two adults, and other adult is selected, go to H-6b.

H-3. The computer has randomly determined that one of the [H-2 answer minus 1] adults other than yourself should be selected for the rest of the interview. To help us select this person, do you know who has had the most recent birthday among these adults?

If no, go to H-5.

H-4. Other than yourself then, which adult has had the most recent birthday?

Go to H-6c.

H-5. So that the computer can choose someone to interview, please tell me the first the names and ages of the adults currently living in this household. Please do not include yourself.

After Kish method is implemented, go to H-6c.

H-6a. What is your first name?

H-6b. The other adult in the household has been selected to participate in the next part of the study. What is the other adult's name?

H-6c. What is that other adult's name?

4. Interviewers noted that when the new method was used a few screener respondents ignored the statement to exclude themselves from the list and self-selected themselves. This problem was easily corrected by adding a simple verification item when asking to speak with the "other" sampled adult.

References

- American Association for Public Opinion Research (AAPOR). 2000. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. Lenexa, KS: AAPOR.
- Binson, Diane, Jesse A. Canchola, and Joseph A. Catania. 2000. "Random Selection in a National Telephone Survey: A Comparison of the Kish, Next-Birthday, and Last-Birthday Methods." *Journal of Official Statistics* 16:53–59.
- Hagan, Dan E., and Charlotte M. Collier. 1983. "Must Respondent Selection Procedures for Telephone Surveys be Invasive?" *Public Opinion Quarterly* 47:547–56.
- Kish, Leslie. 1949. "A Procedure for Objective Respondent Selection within the Household." *Journal of the American Statistical Association* 44:380–87.
- . 1965. *Survey Sampling*. New York: Wiley.
- Lavrakas, Paul J., and Sandra L. Bauman. 1993. "The Last-Birthday Selection Method and Within-Unit Coverage Problems." *Proceedings of the Survey Research Methods Section, American Statistical Association*, pp. 1107–12.
- Lavrakas, Paul J., Elizabeth A. Stasny, and Brian Harpuder. 2000. "A Further Investigation of the Last-Birthday Respondent Selection Method and Within-Unit Coverage Error." *Proceedings of the Survey Research Methods Section, American Statistical Association*, pp. 890–95.
- Lind, Katherine, Michael Link, and Robert Oldendick. 2000. "A Comparison of the Accuracy of the Last Birthday Versus the Next Birthday Methods for Random Selection of Household Respondents." *Proceedings of the Survey Research Methods Section, American Statistical Association*, pp. 887–89.
- Nelson, David E., Barbara K. Rimer, Gary L. Kreps, Bradford W. Hesse, Kasisomayajula Viswanath, Robert T. Croyle, Gordon Willis, Neeraj K. Arora, and Neil Weinstein. In press. "The Health Information National Trends Survey (HINTS): Development, Design, and Dissemination. Manuscript." *Journal of Health Communication*.
- Oldendick, Robert W., George F. Bishop, Susan B. Sorenson, and Alfred J. Tuchfarber. 1988. "A Comparison of the Kish and Last-Birthday Methods of Respondent Selection in Telephone Surveys." *Journal of Official Statistics* 4:307–18.
- O'Rourke, Diane, and Johnny Blair. 1983. "Improving Random Respondent Selection in Telephone Surveys." *Journal of Marketing Research* 20:428–32.
- Salmon, Charles T., and John S. Nichols. 1983. "The Next-Birthday Method of Respondent Selection." *Public Opinion Quarterly* 47:270–76.
- Troldahl, Verling, and Roy Carter. 1964. "Random Selection of Respondents within Households in Telephone Surveys." *Journal of Marketing Research* 1:71–76.
- Tucker, Clyde, James Lepkowski, and Linda Piekarski. 2002. "The Current Efficiency of List-Assisted Telephone Sampling Designs." *Public Opinion Quarterly* 66:321–38.