AN EVALUATION OF RANDOM DIGIT DIALING HOUSEHOLD SURVEY

Wen-Fu P. Shih, University of Florida

ABSTRACT

Telephone surveys based on the random digit dialing (RDD) technique have gained general acceptance among survey researchers. The technique is used because a fairly large percentage of residences' telephone numbers are not listed in the current telephone book, either due to customer request or the number being too new to be listed. The RDD procedure offers each number equal probability of being selected in a sample whether the number is listed or not listed.

The RDD household survey was conducted in Sarasota County, Florida in January, 1980. In order to decide whether random digit dialing is really necessary in a household survey, the analysis was performed both on the group of households with telephone numbers listed (LS) in the current telephone book, and on the overall households. The result indicated that most of the demographic characteristics did not show significant differences between LS and the overall RDD respondents. This suggests that in an area with a low proportion of unlisted households, the telephone survey based on the telephone book will perform as well as the RDD sample survey to analyze general demographic characteristics.

INTRODUCTION

As the cost of personal interview surveys continue to increase, investigators are forced to consider alternative methods of data collection in survey research. As a result, telephone surveys have recently gained much attention (Kahn and Grove, 1977; Glasser and Metzger, 1972; Waksberg, 1978). Telephone surveys offer several advantages, including high respondent rate, lower cost, and speed of survey operation (Frankel and Frankel 1977; Sudman 1966). The major critique of a telephone survey is that telephone subscribers are not necessary representative populations. Not every household has a telephone. Fortunately, in recent years, the percentage of U.S. households with telephone has increased to almost 93 percent (U.S. Bureau of the Census 1976). However, the number of telephone subscribers who request that their name not be listed in telephone directories has also been increasing (Rich, 1977), and new numbers assigned after the current directory is published are also not listed in the directory.

Studies comparing demographic characteristics of listed subscribers (LS) and unlisted subscribers (ULS) can be found in the literature (Brunner & Brunner, 1971, Glasser and Metzger, 1972; Leuthold and Scheele, 1971). These researchers indicated that a telephone survey based on the telephone directory (TSD) which includes only LS will present bias results due to the differences between LS and ULS.

A random digit dialing (RDD) sampling procedure can be used to include unlisted telephone subscribers as well as listed subscribers. Thus, RDD survey provides overall telephone subscribers (OS) equal probabilities of being selected in a sample.

One problem of the RDD survey is that the majority of randomly generated telephone numbers are not working household numbers (Deltor, 1980; Shih, 1980). Dialing a majority of nonworking telephone numbers not only discourages the interviewers but also prolongs the field operations.

An improved RDD survey called 2-stage random digit dialing procedure was introduced by Wakersberg (1978) to overcome the problem. This technique has increased the working telephone numbers sharply (Klecha and Tuchfarber, 1978; Shih 1980). However, 2-stage RDD procedure involves complicated administrative problems. The respondents within each primary sample must be equal to a pre-defined number k. In an actual survey, it is very difficult to locate exactly k working hosuehold telephone numbers.

While the population of RDD survey is overall subscribers. TSD covers listed subscribers only. Therefore, to evaluate the accuracy of RDD survey verses TSD, one should compare on the basis of OS and LS rather than ULS and LS. When the ULS population is much smaller than LS, the sample drawn from telephone directories has virtually the same demographic characteristics as overall subscribers (Rich, 1977).

Based on previous surveys conducted by Pacific Telephone Company and the Bell System, Rich (1977) suggested that even though there are some significant differences exist between the demographic characteristics of listed subscribers and unlisted subscribers, the differences between the listed subscribers and the overall subscribers (which includes ULS) are very small. Unfortunately, Rich did not provide any statistical proof to support his argument.

The purpose of this study are twofolds: first, based on Chi-square statistics, χ^2, to compare demographic characteristics of LS and ULS; second, to modify χ^2 statistics by adding weighting factors from LS and ULS, in order to determine whether the responses of LS are different from OS. Moreover, this study proposes to evaluate when it is necessary to use the RDD survey technique. The data of the household survey in Sarasota County, Florida was used in the study.

METHODOLOGY

Chi-square statistics, χ^2, and weighted chi-square statistics χ^2_w.

Chi-square is a test of statistical significance. It determines whether a systematic relationship exists between LS and ULS. This is done by computing the cell frequencies based on the assumption that there is no relationship between responses of LS and ULS. The expected cell frequencies are then compared to the actual observed values.

$$\chi^2 = \sum_{i=1}^{2} \sum_{j=1}^{k} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \qquad (1)$$

$$\text{and} \quad E_{ij} = \frac{O_{i.} * O_{.j}}{O_{..}} \qquad (2)$$

where O_{1j} : observed frequency of LS in j^{th} category.

O_{2j} : observed frequency of ULS in j^{th} category.

E_{1j} : expected frequency of LS in j^{th} category.

E_{2j} : expected frequency of ULS in j^{th} category.

$O_{i.}$: marginal frequency for LS for i=1, marginal frequency for ULS for i=2.

$O_{.j}$: observed frequency for OS for j^{th} category.

$O_{..}$: total number of overall subscribers.

If the computed χ^2 value is larger than the tabulated chi-square value at significance level α with k-1 degrees of freedom, then the null hypothesis of no difference between LS and ULS will be rejected. However, one should not conclude, based on this rejection, that a telephone survey which covers LS only will produce biased results.

In this computation of χ^2 the proportion of LS and ULS in overall telephone subscribers is neglected. In the actual survey, the population of LS is much larger than ULS, hence $O_{1.}$, O_{1j} and E_{1j} are larger than $O_{2.}$, O_{2j} and E_{2j} respectively. Thus, proportion terms of LS in χ^2 are less than or equal to those of ULS, i.e.:

$$\frac{(O_{1j} - E_{1j})^2}{E_{1j}} \leq \frac{(O_{2j} - E_{2j})^2}{E_{2j}} \quad \text{for all } j \qquad (3)$$

In other words, in Eq 1, the proportion terms of LS are not weighted more heavily than the ULS.

To compare overall subscribers in RDD with listed subscribers in TDS, the computation of χ^2 has to be modified by adding the weighting factors W_i's of LS and ULS, such that

$$\chi^2_w = \sum_{i=1} W_i^2 \sum_{j=1}^{k} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \qquad (4)$$

where $W_i = \frac{O_{i.}}{O_{..}}$, i = 1, 2 and $\sum_{i=1}^{2} W_i = 1$.

The resulting χ^2_w is modified by the proportion weights of LS and ULS in a population. If χ^2_w is larger than the tabulated χ^2, we can conclude that the responses of OS are different from the LS, hence the RDD survey will provide more valid data. Otherwise the less costly telephone survey based on the directory can collect data which is as satisfactory as data collected in a RDD survey.

DATA

The household survey in Sarasota County, Florida was conducted in January, 1980 using the 2-stage RDD procedure (Shih, 1980). At the end of the telephone interview, the respondents were asked whether their telephone numbers were listed in current telephone books, too new to be listed or voluntary unlisted. Of the 517 respondents, 9.5 percent of the respondents said the numbers were too new to be listed, 7.5 percent said unlisted and 3.5 percent did not answer the question. In other words, among those 499 households responding to the question, more than 82 percent of the households were listed subscribers and 18 percent were unlisted subscribers (including new subscribers and voluntary unlisted subscribers).

RESULTS OF ANALYSIS

The distributions of sample characteristics of listed telephone subscribers, unlisted subscribers and overall subscribers are shown in Table 1. Chi-square statistics, χ^2, of LS vs. ULS, and weighted chi-square statistics, χ^2_w, are presented in the table. As can be seen in Table 1, for household size and income, respondents' sex, education, reason moved to Florida, and month away from home in 1979, the response distributions of LS did not show any significant differences from ULS. Consequently weighted chi-squares, χ^2_w, were also nonsignificant for those characteristics. The results show that the average household age of ULS was significantly younger than LS; for those households not living in a single-family house, ULS were more likely to live in an apartment, while LS were more likely to live in a condominium; ULS had a higher percentage of respondents (8.4%) who were never married than LS (3.3%); there were more newcomers (20%) who had been in Florida less than one year among unlisted subscribers than listed subscribers (1%). Moreover, ULS were more likely to have moved from New England or the South Atlantic area, while almost 70 percent of LS moved from the Middle Atlantic or East North Central areas. However, because the majority of the respondents were LS (83%), the weighted chi-square χ^2_w showed significance in only two characteristics the average household age and the length of time lived in Florida.

In this household survey in Sarasota County, Florida, with 83 percent LS, the RDD survey did provide more accurate information about household members' age and the length in Florida than TDS. A regular telephone survey based on the current directory would have collected information similar to that collected by the RDD survey for all other demographic characteristics at a considerably lower cost.

CONCLUSION

Random Digit Dialing Telephone survey has been suggested by its wide coverage of sampling frame, since it covers all telephone subscribers, listed or unlisted. However, some demographic characteristics do not show significant differences between ULS and LS, e.g. sex and education of respondents, household size and household income. For research focusing on those characteristics, the RDD survey will not be necessary, considering its higher interview cost compared to the ordinary telephone survey based on the directory.

Even though there are significant differences between LS and ULS for some characteristics, when the differences are small and when 83 percent of the respondents are LS, the weighted chi-square, χ^2_w, showed some characteristics to be nonsignifi-

cant, such as type of housing unit and respondents' marital status. The distribution of average household age and the length in Florida were the only two characteristics which showed listed subscribers as being significantly different from overall subscribers.

In geographic areas where the proportion of unlisted subscribers is high, or where certain demographic profiles, such as household age, are known to be important factors in differentiating telephone subscribers, use of the RDD procedure is suggested.

REFERENCES

Brunner, J.A., and Brunner G.A. (1971), "Are Voluntarily Unlisted Telephone Subscribers Really Different?" Journal of Marketing Research, 121-124.

Conover, W.J. (1971), Practical Nonparametric Statistics, New York: John Wiley & Sons Inc., 186-196.

Delton R.J. (1980), "Improving Random Digit Dialing: Experiences with Two RDD Alternations," Annual Meeting of the Southwestern Political Science Association at Houston, Texas.

Frankel, M.R., and Frankel L.R. (1977), "Some Recent Development in Sample Survey Design," Journal of Marketing Research, 14, 280-293.

Glasser, G.J., and Metzger, G.D. (1972), "Random Digit Dialing as a Method of Telephone Sampling," Journal of Marketing Research, 9, 59-64.

Kahn, R.L., and Grove R.M. (1977), "Comparing Telephone and Personal Interview Survey," Survey Research Center at the University of Michigan, unpublished report submitted to NSF SOC 76-07519.

Klechka, W.R., and Tuchfarber, A.J. (1978), "Random Digit Dialing: A comparison to Personal Surveys," Public Opinion Quarterly, 42, 104-114.

Leuthold, D.A., and Scheele, R. (1971), "Patterns of Bias in Samples Based on Telephone Directories," Public Opinion Quarterly, 35, 245-257.

Rich, C.L. (1977), "Is Random Digit Dialing Really Necessary?" Journal of Marketing Research, 14, 300-325.

Shih, W.F. (1980), "Random Digit Dialing Household Survey in Florida," Annual Meeting of the American Statistical Association, Florida Chapter.

Sudman, S. (1966), "New Uses of Telephone Methods in Survey Sampling," Journal of Marketing Research, 10, 204-207.

Waksberg, J. (1978), "Sampling Method for Random Digit Dialing," Journal of the American Statistical Association, 73, 40-46.

Table 1
Characteristics of listed telphone subscribers (LS), unlisted telephone subscribers (ULS) and overall subscribers (OS)

	Percentage (total N) (Weight W_i)			χ^2	χ^2_w
	Unlisted subscribers	Listed subscribers	Overall subscribers	ULS vs. LS	LS vs. OS
Average household age					
Less than 18 years	0.0	0.2	0.2	16.60**	7.33*
18-44	36.8	22.7	25.2		
44-64	35.6	26.1	27.8		
65+	27.6	51.0	46.9		
TOTAL	(87)	(406)	(493)		
W_i	(0.18)	(0.82)	(1.00)		
Household size					
1 person	10.2	15.6	14.6	0.94	0.56
2	56.8	58.2	57.9		
3	13.6	12.9	13.0		
4 or more	19.3	13.4	14.4		
TOTAL	(88)	(411)	(499)		
W_i	(0.18)	(0.82)	(1.00)		
Type of housing unit					
Single-family	59.0	63.8	62.9	10.69*	2.36
Apartment	19.3	7.9	9.9		
Condominium	12.0	18.1	17.1		
Mobile home	9.6	10.2	10.1		
TOTAL	(83)	(392)	(475)		
W_i	(0.17)	(0.83)	(1.00)		
Sex					
Male	39.8	32.8	34.1		
Female	60.2	67.2	65.9	1.25	0.38
TOTAL	(88)	(411)	(499)		
W_i	(0.18)	(0.82)	(1.00)		

Table 1
Characteristics of listed telphone subscribers (LS), unlisted telephone subscribers (ULS) and overall subscribers (OS) (continued)

	Percentage (total N) (Weight W_i)			χ^2	χ^2_w
	Unlisted subscribers	Listed subscribers	Overall subscribers	ULS vs. LS	LS vs. OS
Marital status					
Married	74.7	79.4	78.6		
Widow	10.8	13.6	13.1	9.79*	2.52
Divorced	3.6	3.3	3.4		
Separated	2.4	0.3	0.6		
Never married	8.4	3.3	4.2		
TOTAL	(83)	(389)	(472)		
W_i	(0.18)	(0.82)	(1.00)		
Education					
High school or less	44.8	48.2	47.6		
Some college	29.9	23.2	24.4	1.75	0.51
College and post college	25.3	28.6	28.0		
TOTAL	(87)	(409)	(496)		
W_i	(0.18)	(0.82)	(1.00)		
Income					
$10,000 or less	14.9	18.4	17.8		
10,001-18,000	32.4	32.0	32.1	0.77	0.32
18,001-25,000	25.7	22.2	22.8		
25,001 or more	17.4	27.4	27.3		
TOTAL	(74)	(347)	(421)		
W_i	(0.18)	(0.82)	(1.00)		
Length in Florida					
Less than 1 year	20.0	1.0	4.2	65.82**	18.19**
1-5 years	40.0	31.5	32.9		
6-10 years	11.0	32.7	30.6		
over 10 years	20.0	34.8	32.3		
TOTAL	(80)	(391)	(471)		
W_i	(0.17)	(0.83)	(1.00)		
Previous residence by region					
New England	14.3	9.1	10.0	8.46*	2.46
Middle Atlantic	23.8	30.4	29.2		
South Atlantic	19.0	13.9	14.8		
East North Central	27.4	37.2	35.5		
Other	15.5	9.4	10.4		
TOTAL	(84)	(395)	(479)		
W_i	(0.18)	(0.82)	(1.00)		
Reason moved to Florida					
Employment	12.7	10.8	11.2	0.52	0.15
Climate-retirement	74.7	73.8	74.0		
Family	12.7	15.3	14.9		
TOTAL	(79)	(378)	(457)		
W_i	(0.17)	(0.83)	(1.00)		
Month away from home in 1979					
Less than a month	84.1	78.8	79.8	2.86	0.64
1-3 months	4.5	7.8	7.2		
4-6 months	3.4	6.8	6.2		
7-12 months	8.0	6.6	6.8		
TOTAL	(88)	(411)	(499)		
W_i	(0.18)	(0.82)	(1.00)		

*significant at $\alpha = 0.5$,
** significant at $\alpha = 0.01$