

PANEL ATTRITION IN A DUAL-FRAME LOCAL AREA TELEPHONE SURVEY

Judith A. Schejbal, Paul J. Lavrakas, Northwestern University Survey Laboratory
Paul J. Lavrakas, NUSL, 625 Haven Street, Evanston, IL 60208

Keywords: Telephone Surveys, Panel Attrition

The magnitude and nature of respondent attrition in local area panel surveys are of particular concern for survey researchers since they can directly effect the generalizability of survey findings. Survey researchers need to know who they have lost to attrition and how this loss effects the data gathered and the results obtained. Very few reported research studies have addressed panel attrition in telephone surveys using random-digit dialing sampling (RDD) at Wave 1. Although the exact proportion of attrition will vary in part depending upon the mobility pattern of the residents within the specific residential area, approximately 50%-60% of the original respondents in RDD samples can be found and re-interviewed after a one-year time lag in large cities (Lavrakas, 1993, pp. 87-88).

Researchers have addressed the techniques by which one can minimize the size of attrition in panel studies (cf. Lavrakas, Settersten, Maier, 1991). However, few researchers have reported how the respondents retained in a telephone panel differ from the respondents lost to attrition on substantive measures of interest or by type of sampling frame (RDD versus list-based). Most efforts have concentrated instead on the demographic correlates of respondents lost to attrition. However, attempts at reducing panel attrition might have greater success if researchers can learn a wider variety of information about the respondents who are lost as compared to those retained in the sample.

This paper presents methodological findings from a dual-frame panel telephone survey conducted in 1993 and 1994 by the Northwestern University Survey Laboratory. Wave 1 respondents retained at Wave 2 are compared to Wave 1 respondents lost at Wave 2 on a host of demographic and substantive measures. This paper also includes a discussion of any sample type (RDD versus list-based) differences found.

METHODOLOGY

At Wave 1, 2,570 adult City of Chicago residents, aged 18 years and older living in nine local community areas, were interviewed as part of a multi-year evaluation research project studying the implementation of a new "community policing" anti-crime strategy by the Chicago Police Department (Skogan and Hartnett, 1993). Interviews were conducted in English and in Spanish in Spring, 1993.

An important characteristic of this particular telephone survey is its use of a dual-frame sampling design. That is, approximately one-half of the households (n=1,292) at Wave 1 were sampled via a Haines City of Chicago reverse directory (REV sample frame) and the other one-half of each community area (n=1,278) was sampled via traditional two-stage random-digit dialing (RDD sample frame). The decision to use both frames was based on cost considerations.

The RDD approach with geographic screening was known to be significantly more expensive than the REV approach (Schejbal and Lavrakas, 1995). It was reasoned that RDD would reach a sample of households with less potential Coverage Error (i.e., it would reach households with unlisted telephone numbers), but that the proportion of households reached and needing to be screened out due to geographic ineligibility would be large and quite expensive. By using a dual-frame approach, the researchers could save money and conduct post hoc analyses to determine what differences in the substantive data, if any, were associated with the different sampling frames.

At the conclusion of the Spring 1993 interview (Wave 1), respondents were asked: "*We may want to call some people back in a year or so to see how things are in their neighborhood. May we call you back?*" If the respondent replied "Yes", the interviewer asked for the respondent's first name. Those respondents who said "No" were excluded from the Wave 2 sampling frames. Only 84 (3.3%) of the original 2,570 Wave 1 respondents did not give permission to be re-contacted at Wave 2.

At Wave 2, attempts were made to re-contact the 1,073 Wave 1 participants in four of the original nine community areas in June, 1994 (time-lag of approximately 14 months). Attempts to re-contact the 1,413 Wave 1 participants living in the other five community areas were made in September, 1994 (time-lag of approximately 17 months). It was necessary to split the Wave 2 data collection into two parts due to the sponsor's funding.

RESULTS

Respondents Who Asked Not to be Recontacted Wave 2

The 84 Wave 1 respondents who requested not to be contacted in the future were compared to respondents retained in the Wave 2 samples on all demographic and

some select dependent measures. These two groups of Wave 1 respondents were not significantly different from each other on any of the demographic factors and were significantly different on only two of the 34 substantive variables included in the comparative analyses. Only one other dependent measure was marginally significant. Overall, these two respondent groups were similar enough to conclude that the Wave 2 respondent pool was not significantly altered by the loss of the 84 Wave 1 respondents who asked not to be called back.

Differential Time-lag Comparisons

Sixty-five percent (n=700) of the original Wave 1 respondents were successfully re-interviewed after a 14-month time lag, while a significantly lower percentage of 56% (n=793) were successfully reinterviewed after a 17-month time lag. Overall, a total of 1,493 (58%) Wave 1 respondents completed telephone interviews at Wave 2.

Since the Wave 2 data were collected at two different points in time and in different community areas, it was important to address the possibility that the differential completion rates associated with the differential time-lags might have been due to area-differences, not time-differences. It was found that the completion rate for the Wave 1 respondents re-contacted after a 14-month lag (Spring, 1994) was significantly higher

($p < .01$) than the completion rate for Wave 1 respondents re-contacted after a 17-month lag (Fall, 1994), even after controlling for demographic differences among the community areas.

In Table 1, the two data collection periods for Wave 2 are compared on disposition of the samples by time-lag. Compared to the Spring sample, the Fall sample contained proportionately more non-working telephone numbers ($p < .001$), fewer answering machines ($p < .05$), more telephone numbers which reached households in which the respondents were never available ($p < .10$) and fewer households in which the respondent was away for an extended period ($p < .10$); (in addition to fewer completions than the Spring 1994 sample).

Wave 2 Sample Frame Comparisons

Turning now to the combined Wave 2 data, a significant difference was found for completion rate by sampling frame ($p < .05$), with 58% of Wave 1 RDD respondents versus 65% of Wave 1 REV respondents completing interviews at Wave 2.

The final disposition for each Wave 1 telephone number included in Wave 2 is shown in Table 2 by type of sampling frame. The Wave 2 REV sample contained proportionately more telephone numbers that reached respondents who had moved to ineligible geographic areas ($p < .01$), respondents who refused ($p < .10$), in addi-

Final Disposition	Spring	Fall
Respondent lives in ineligible geographic area	1.6%	0.9%
Non-working	5.0%	12.1%
Temporary disconnect	1.2%	1.5%
Non-residential	0.9%	0.6%
Another household has number	7.5%	8.8%
Family moved no new # available	2.1%	4.9%
No answer	1.9%	2.2%
Answering machine	2.6%	1.5%
Respondent never available	1.5%	2.7%
Non-English speaking	0.5%	1.3%
Mental/Physical disability	1.9%	1.3%
Respondent away for extended period	1.9%	0.9%
Household/Respondent refusal	3.1%	2.3%
Completed full interview	65.2%	56.1%
Completed partial interview	0.5%	0.2%
Miscellaneous	2.6%	2.5%
	100.0%	100.0%

Table 2: Final Sample Disposition by Sample Type Wave 2 Spring & Fall Combined		
Final Disposition	Sample Type	
	RDD	REV
Respondent lives in ineligible geographic area	0.6%	1.8%
Non-working	11.0%	7.0%
Temporary disconnect	1.4%	1.4%
Non-residential	1.0%	0.6%
Another household has number	10.6%	5.9%
Family moved no new # available	4.4%	2.9%
No answer	1.8%	2.4%
Answering machine	2.3%	1.6%
Respondent never available	1.9%	2.4%
Non-English speaking	0.8%	1.1%
Mental/Physical disability	1.4%	1.9%
Respondent away for extended period	1.3%	1.4%
Household/Respondent refusal	1.9%	3.3%
Completed full interview	56.4%	63.7%
Completed partial interview	0.3%	0.3%
Miscellaneous	2.9%	2.3%
	100.0%	100.0%

tion to more completions. The Wave 2 RDD sample, on the other hand, contained proportionately more telephone numbers that were non-working ($p < .01$), and ones in which another household name was associated with the telephone number ($p < .001$).

Compared to the Wave 2 REV respondents, the Wave 2 RDD respondents were significantly more likely to be female, younger, non-White, have more children living in the household, be renters, have lived in their neighborhoods for a shorter period of time. In addition, the Wave 2 RDD respondents were marginally more likely to be unmarried.

Looking at the 34 substantive dependent measure comparisons, the RDD respondents were significantly different from the Wave 2 REV respondents on only two of these variables and marginally significantly different on only two other substantive measures.

Separate logistic regressions were conducted for each type of sample frame using demographic variables to predict whether or not a Wave 1 respondent was re-interviewed at Wave 2. Those Wave 2 respondents originally sampled via RDD, and who were subsequently re-interviewed, were more likely to be home owners, female, have fewer adults living in the household, have lived in their neighborhood longer, and be married than those original RDD respondents who were not re-inter-

viewed. Those Wave 2 respondents originally sampled via REV, and who were subsequently re-interviewed, were more likely to be home owners, White, and be married than those original REV respondents who were not re-interviewed.

Table 3 presents the results of demographic variable comparisons made between Wave 1 respondents "found" at Wave 2 and Wave 1 respondents "lost" at Wave 2 by type of sampling frame. Of particular interest are the sampling frame differences associated with gender, race, and income. Significantly fewer males than females were found at Wave 2 for the RDD sampling frame. No significant gender difference was found for the REV sampling frame. Regarding race, significantly fewer Hispanics than Non-Hispanics were found at Wave 2 for both sampling frames. However, a greater proportion of Hispanics were *lost* in the REV sampling frame. Significantly more Whites than Non-Whites were found at Wave 2 for both sampling frames; however, a greater proportion of Whites were *found* in the REV sampling frame. Regarding income, fewer low income households and more high income households were found for the REV sampling frame at Wave 2, with no significant income difference found for the RDD sampling frame.

Table 4 presents the results of substantive variable

Table 3:

Demographic Variables for Wave 1 Respondents LOST versus FOUND at Wave 2
by Type of Sampling Frame (RDD $n = 1,278$, REV $n = 1,292$)

Variable	RDD Sampling Frame			REV Sampling Frame		
	LOST	FOUND	$p <$	LOST	FOUND	$p <$
% Male	43.1	32.6	.000	42.6	43.0	NS
% 60 Yrs or Older	15.3	23.0	.001	26.7	30.8	NS
% 18-29 Yrs	34.0	16.0	.000	23.7	14.1	.000
Mean Age in Yrs	39.2	45.9	.000	45.0	48.7	.001
% African American	52.2	49.6	NS	53.1	49.0	NS
% Hispanic	17.2	12.9	.036	12.8	5.3	.000
% White	24.8	31.9	.006	26.4	40.9	.000
% HS Grad, Not Coll Grad	54.7	52.3	NS	54.2	48.3	.060
% College Grad	19.3	24.8	.023	18.1	30.8	.000
% Hshld Income \$10K or <	21.9	19.1	NS	26.4	17.0	.000
% Hshld Income \$60K or >	9.1	8.5	NS	5.5	11.3	.001
% Married	37.1	42.4	.063	35.0	47.4	.000
Mean # Adults in Hshld	2.3	2.1	.029	2.1	2.2	NS
% Hshld with Child(ren)	49.7	44.7	.082	40.3	36.9	NS
% Employed Full-time	46.4	50.4	NS	42.6	46.2	NS
% Own Home	36.6	54.6	.000	40.0	62.5	.000
Mean # Yrs Lived in Nhbd	9.5	12.6	.000	11.5	14.8	.000
% With > 1 Telephone Line	21.1	14.9	.005	15.6	17.4	NS

comparisons made between Wave 1 respondents "found" at Wave 2 and Wave 1 respondents "lost" at Wave 2 by sampling frame. Significant differences were found for respondents "lost" and respondents "found" across a number of variables presented for both sampling frames. In general, regardless of sampling frame, respondents "lost" to attrition had responded to the Wave 1 crime-related substantive variables in a consistently more *negative* manner than the group of respondents "found". In addition, the respondents "found" at Wave 2 had responded to the specific Wave 1 community policing items in a consistently more *positive* manner than the group of respondents "lost".

DISCUSSION

There has been relatively little reported about telephone survey respondents who are lost to panel attrition and how they differ from respondents who are re-interviewed. Heretofore, nothing has been reported about differences associated with differential panel attrition by type of sampling frame. This paper provides information on these differences based on one local area survey utilizing a dual-frame sampling approach.

To review, our major findings were:

1) Compared to those Wave 1 respondents who

were called back, the Wave 1 respondents who asked not to be called back at Wave 2 did not differ on any of the demographic variables and were only marginally different on one substantive variable. Therefore, the removal of these respondents did not significantly alter the representativeness of the Wave 2 sample;

2) A significant difference was found for completion rate by type of *sampling frame* at Wave 2, with the REV frame having less panel attrition. Although the completion rate for the two time periods (Spring and Fall) was at, or exceeded, the expected completion rate found in previous studies (i.e., 50%-60%), the longer time-lag of the Fall 1994 data collection period contributed to significantly more panel attrition;

3) Differences by sampling frame were noted in the final disposition of the Wave 2 sample shown in Table 2 (Spring and Fall, 1994 data). Since RDD samples contain unlisted as well as listed telephone numbers, and since households with unlisted telephone numbers are more likely to be renters (c.f., Lavrakas, 1993), it is not surprising to find that the Wave 2 RDD sample contained more non-working telephone numbers and more numbers that reached households in which the original respondents were no longer residing;

4) Sampling frame differences were also found for some Wave 2 respondent demographic characteristics as

Table 4

Substantive (Dependent) Variables for Wave 1 Respondents LOST versus FOUND
at Wave 2 by Type of Sampling Frame (RDD $n = 1,278$, REV $n = 1,292$)

Variable	RDD Sampling Frame			REV Sampling Frame		
	LOST	FOUND	$p <$	LOST	FOUND	$p <$
% Dissat with Own Nhbhd	35.4	25.4	.000	28.2	25.3	NS
% Unsafe at Night in Nhbhd	50.7	47.7	NS	50.4	45.5	NS
% Big Prob Abandoned Bldgs	18.0	13.9	.052	18.4	14.7	NS
% Big Prob Graffiti in Nhbhd	23.4	20.7	NS	24.7	23.0	NS
% Big Prob Police Exc Force	12.4	7.4	.003	10.6	7.1	.041
% Big Prob Illegal Drugs	44.7	34.9	.000	37.3	32.8	NS
% Big Prob Street Gangs	31.1	24.1	.006	25.9	21.7	NS
% Big Prob Auto Theft	20.9	16.6	.056	19.6	18.8	NS
% Big Prob Home Burglary	17.0	16.6	NS	15.4	17.0	NS
% Big Prob Street Crime	25.0	18.9	.010	20.7	18.4	NS
% Big Prob Rape & Sex Crime	10.1	11.1	NS	11.1	7.2	.026
% Past Yr Burg Victim	15.5	12.5	NS	11.6	15.6	.064
% Past Yr St Crime Victim	13.5	7.7	.000	10.8	7.6	.068
% Report Crime Past Yr	28.4	27.4	.000	22.7	29.5	.014
% Aware of Nhbhd Meetings	45.8	54.9	.002	47.9	59.9	.000
% Aware Comm Policing Prog	23.8	32.1	.000	26.2	39.8	.000
% Think Police Responsive	58.8	66.1	.009	64.2	69.7	.061
% Think Police Improving	16.4	18.3	NS	15.9	15.8	NS
% Seen Police Past Wk	31.1	33.6	NS	33.5	33.1	NS
% Seen Foot Patrol Past Mo	19.9	16.9	NS	18.6	18.3	NS

noted in Table 3 and as shown by the logistic regression analyses. Most striking is the difference in the proportion of males lost for the RDD sampling frame versus the REV sampling frame. This sampling frame gender difference could be problematic for researchers using a RDD sampling frame in a panel study, in particular when their substantive measures correlate with gender;

5) Sampling frame differences were found for several substantive variables (see Table 4) for respondents lost to attrition versus respondents found at Wave 2. Furthermore, the data presented in Table 4 indicate a difference in overall outlook between the lost versus found groups, regardless of type of sampling frame. That is, the respondents "found" were generally positive when asked about crime in their neighborhood and about local policing efforts, while the respondents "lost" were relatively more negative.

In conclusion, additional methodological research is needed to determine who comprises the group of respondents "found" and who comprises the group of respondents "lost" at subsequent waves of telephone panel surveys using different sampling frames. These distinctions are important to make because they will

ultimately effect the survey errors related to different sampling frames as well as the final project costs, both of which will affect survey quality and related policy recommendations.

REFERENCES

- Lavrakas, P.J. (1993). *Telephone survey methods: sampling selection and supervision* (2nd Edition). Newbury Park: Sage Publications.
- Lavrakas, P.J., Settersten, R.A., and Maier, R.A. (1991). "RDD Panel Attrition in Two Local Area Surveys." *Survey Methodology*, Vol. 17 (2), pp. 143-152.
- Schejbal, J.A. and Lavrakas, P.J. (1995). "Coverage Error and Cost Issues in Small Area Telephone Surveys." 1994 Proceedings of the ASA/AAPOR Survey Research Methods Section.
- Skogan, W.G. and Hartnett, S.M. (1993). *Year One Feedback on the Chicago Community Policing Program*. Evanston, IL: CUAPR. (mimeo)