

AN ANALYSIS OF TELEPHONE CALL HISTORY DATA FROM THE BEHAVIORAL RISK FACTOR SURVEILLANCE SYSTEM

Wendy M. Ahmed and William D. Kalsbeek

Survey Research Unit UNC-CH, 730 Airport Rd. Suite 103, Chapel Hill, NC 27514

Keywords: Nonresponse, Telephone surveys, BRFSS

Survey research, particularly via telephone surveys, is becoming a more common means of gathering information. Such methods offer powerful techniques for making estimates and inferences about specific populations. The capabilities to make proper estimates rely on probability sampling from a frame consisting of all members of a population. Here, each member has some known non-zero selection probability that contributes to estimates and variances. Nonresponse by sample members interferes with probability sampling, thus compromising the validity of resulting estimates, by implying unknown status for the fraction of the population which those nonrespondents represent. Even the simplest sample statistics are then biased.

Another motivation for studying nonresponse is to understand the operational realities of nonresponse in hopes of finding methods to improve completion rates. This particular research is a first step toward understanding nonresponse using a large national telephone survey. Two types of nonresponse are addressed - noncontact and refusals. We profile and compare contact and refusal rates to evaluate implications for calling schedules and differences due to different survey organizations. We also start to dissect the refusal conversion process, investigating effects of timing, interviewer characteristics, and survey organization policies. Again, this research is a first step, but we expect it will yield solid practical implications and lend insight into new ideas for future research.

BACKGROUND

The most recent and extensive collection of nonresponse research is a book by Robert Groves and Mick Couper (1998) called *Nonresponse in Household Interview Surveys*. This is an assembly of prior research and a cross-sectional study of nonresponse in several face-to-face surveys. Groves and Couper present a detailed model (1998: 30) of survey cooperation that includes operational factors, interviewer effects, and behavioral and socio-psychological dimensions. This model makes a distinction between factors that can be controlled by the researcher and those that can not, and then explains that all of these factors come to light in the actual interviewer-household interaction. What happens during this interaction governs the outcome of the attempt.

Using a number of federal in-person interviews, they matched data about interviewers and information about respondents from the actual surveys. Though there are distinct differences between telephone and in-person surveys and interviewers, they explain that the next step they "are most interested in [is] measuring the extent of applicability [of this research] to telephone surveys" (Groves and Couper, 1998: 320). This study begins to look at nonresponse in telephone surveys using much of the same thinking presented by Groves and Couper.

It is important to distinguish refusals from other types of nonresponse because people who refuse are likely to be quite different from those who do not, which increases the bias in resulting estimates (Groves and Couper, 1998: 12). This particular study focuses on operational, rather than behavioral correlates of nonresponse. Similarly, Weeks, Kulka, and Pierson (1987) studied survey protocols for call scheduling in attempts to develop an optimal call scheduling plan to increase completion rates. Their study, along with most others on the topic, supports the finding that reaching a respondent at home and completing an interview is most likely to occur during weekday evenings and weekends - when most people are home.

Due to the nature of face-to-face interviews, the focus of Groves' and Couper's (1998) research, there are also extensive discussions of the effects of interviewer characteristics on survey cooperation (see also Groves, Cialdini, and Couper, 1992). Since respondents actually see interviewers, these effects are particularly interesting to in-person survey operation staff. Using logistic regression procedures, Groves, Cialdini, and Couper found some interviewer characteristics to have significant positive effects on completion (1998: 316).

It is imperative to note an important difference here, however. Telephone surveys have the advantage of masking many characteristics of interviewers that may interfere with survey cooperation. Only those attributes that may become apparent by voice are likely problems: gender, age, and, in some cases, race (1998: 300-301). For both in-person and telephone interviewers it is well established that more experienced interviewers achieve better completion rates, perhaps because "they carry with them a larger number of combinations of behaviors proven to be effective" (Groves, et al., 1992: 478-9; 1998: 36). On the other hand, telephone interviewers generally have higher turn around rates than face-to-face interviewers, so relationships between completion

success and experience may not be as clear for telephone interviews (1998: 203).

The general purpose of this study is to begin applying these types of thinking to studying nonresponse in a large national telephone survey - The Behavioral Risk Factor Surveillance System (BRFSS). Being one of the largest surveys of this kind, it provides an excellent resource for studying various problems that plague all surveys, regardless of size or subject matter: namely non-response and refusals. This study explores some preliminary issues in order to understand the potential uses of BRFSS data and set the stage for future more detailed analyses of call history files. The most important premise of this research is that it is a starting point. The following analyses are exploratory in nature. Discussions with the designing agency, the Centers for Disease Control and Prevention, have focused on use of the research to study nonresponse as well as operational aspects specific to the BRFSS. Such discussions have resulted in a number of goals to be met by the preliminary analyses of BRFSS call history data. After a brief discussion of the data, results from two of the four study goals are presented here.

DATA

As mentioned above, the BRFSS is a large telephone survey designed by the Centers for Disease Control (CDC) to collect information on various "preventive health practices and risk behaviors that are linked to chronic diseases, injuries, and preventable infectious diseases." (CDC 1997). To implement BRFSS data collection, each state or territory is charged with collecting data from a random sample of one adult per household and delivering it to CDC monthly. For this preliminary study, the data are a convenience sample that happened to be available for twelve states over the entire year of 1997. Though the sample is one of convenience, the twelve states all vary across different characteristics such as urban/rural status, use of contractor, and region of the country, as well as having widely varying 1995 response and refusal rates. All twelve states, though, use the Ci3 system to collect the data and keep track of call histories.

These call histories include information on every call attempt made to every sample number such as the date of the call, the start and end time for each attempt, the telephone number, the interviewer identification number, and the disposition for each attempt. Additionally, each participating state calling organization was sent a short questionnaire asking about interviewers' age, experience, race and gender. Each organization also returned a questionnaire about their individual policies for handling refusals, i.e. if certain interviewers are designated and what information they are provided for dealing with refusal conversions.

So for each of the 109,346 telephone numbers, there is a string of dispositions that lead to some final disposition. From this string, we created an indicator for whether each number ended up likely to be a residential number. Some codes clearly indicate a number was non-working or a business number, but we determined certain strings of dispositions likely to be residential. For each telephone number, we also created an indicator of the result of the "cold call" or first call to that number. Various other indicators and rates were computed for more detailed analyses and are explained in the results section.

GOALS AND METHODS

Goal 1 - Determine the best times of the day and week to reach people at home.

Our first goal was to use the data to profile when BRFSS calls are made and then, when interviewers were most successful in reaching people at home on first or "cold" call attempts. We started with a descriptive profile of when someone picks up the telephone on the first "cold" call attempt. We looked at these rates by day of week and hour of day using a Pearson chi square statistic for general association to test for significant relationships. All comparisons were made for all telephone numbers, and then for those that are likely to be residential numbers.

Goal 2 - Determine levels of success in dealing with refusals.

This second goal explores refusal rates and patterns of initial refusals and conversion to completed interviews. Determining levels of refusal conversion success was examined on a number of levels. The first outcome measure here was the rate of initial refusal on the first "cold" call attempt. We developed a profile of these rates by hour, day, and month of the year and again use Pearson chi square tests for general association to evaluate significance levels. These relationships were also examined for all telephone numbers and those likely to be residential numbers.

Next we are interested in conversion rates - the percent of those initial refusals that are converted to completed interviews. More specifically, we see how those rates differ according to which call attempt the initial refusal occurred. For those numbers where a refusal occurred, we determined the percent of completed interviews thus calculating a rate of success in converting a refusal to a completed interview. One key question was whether these conversion rates are different according to the attempt on which the initial refusal occurs. Mantel-Haenzel chi square tests were used to test the significance of any associations, since there is inherent ordering in the sequence of call attempts.

The third component of understanding refusal conversions was exploring attempt level predictors of success. Using logistic regression procedures we looked at interviewer characteristics and calling environment policies and their effects on conversion rates. The data used for this section of the analysis was different from the data used for previous analyses. Each call attempt to a particular telephone number was a separate record in this dataset; and for each number, only those attempts after the initial refusal were kept. Further, only those telephone numbers for which we have interviewer information were used which left us with 2,545 attempts to use for modeling. The attrition of the sample was due to including only refusal numbers and then only attempts after the initial refusal. The outcome measure was an indicator of whether the refusal was converted on that particular call attempt (i.e. if the interview was completed).

The predictors considered for entry in the model were interviewer experience prior to BRFSS interviewing, length of specific BRFSS experience, interviewer age, and gender from the questionnaire sent to states. Also included were measures of whether that state's calling organization designates special interviewers for refusal conversion and some measure of the number and types of techniques that are implemented to equip interviewers to deal with refusals. Logistic regression procedures were done using SUDAAN to take into account stratification by state and month, and clustering by telephone number since a number of attempts will belong to one telephone number.

RESULTS

Goal 1. Of the 109,346 BRFSS calls in this set of data, most are generally made during the week (n=97,137), with the highest number (n=33,293) made on Wednesdays (calling sessions each month begin on Wednesdays). Further, there are 3 distinct calling "sessions," with most calls made in the evening session after 5pm (54,970). There are no calls made Sunday morning, and only modest numbers made Saturday and Sunday evenings (n=3,031).

Next, we use the indicator of the result of the first or "cold" call attempt to each number, concentrating mainly on those numbers that are likely to be residences (n=48,509). Monday has the highest percentage of residential "pickups" (60.5%) than any other day, while Wednesday has the lowest rate of 53.6% (see Table 1).

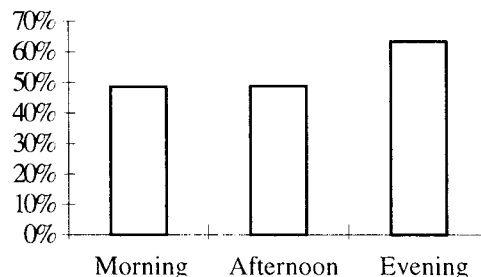
Interviewers are much more likely to reach respondents in the evening as well. Of the calls made after 5pm, 63.4% resulted in a "pickup" at a residence, while 48.8% of the afternoon calls and 48.7% of the morning calls resulted in "pickups" ($\chi^2=1051.9$, $df=2$, $p=.001$, see Figure 1).

This relationship holds for both weekdays and weekends. Weekday evenings have the highest contact rate (63.6%), while weekday morning contact rates are lowest (47%, $\chi^2=1091.48$, $df=2$, $p=.001$).

TABLE 1: PERCENT OF CALLS MADE TO LIKELY RESIDENTIAL NUMBERS RESULTING IN RESIDENCE "PICKUPS"

DAY	Residential Pickup	
	N	Percent
MONDAY	2799	60.47
TUESDAY	4125	56.41
WEDNESDAY	8725	53.61
THURSDAY	6005	56.67
FRIDAY	2192	57.07
SATURDAY	2566	55.58
SUNDAY	719	57.98
	$\chi^2=81.53$ $df=6$ $p=.001$	

FIGURE 1: RATES OF RESIDENTIAL PICKUPS FOR ATTEMPTS MADE TO LIKELY RESIDENTIAL NUMBERS BY TIME OF DAY



Goal 2. Now we look at refusal rates on the first or "cold" call attempt by month, day, and hour. First, the highest refusal rate of 6.9% occurred in March, while the lowest is in December at 5.2% ($\chi^2=24.5$, $df=11$, $p=.011$, see Figure 2).

FIGURE 2: REFUSAL RATES ON FIRST OR "COLD" CALL ATTEMPTS BY MONTH



Refusal rates are higher on weekends than during the week, and are lowest on Friday with only 4.7% of "cold" calls encountering a refusal ($\chi^2=68.8$, $df=6$, $p=.001$, see Figure 3). We also see higher refusal rates in the evening hours than throughout

the rest of the day ($\chi^2=199.5$, $df=13$, $p=.001$, Figure 4).

FIGURE 3: REFUSAL RATES ON FIRST OR "COLD" CALL ATTEMPTS BY DAY OF WEEK

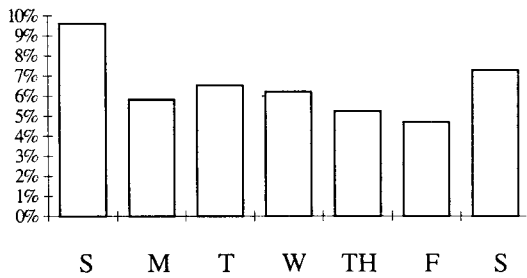
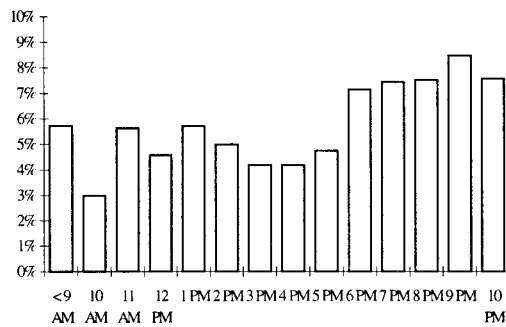
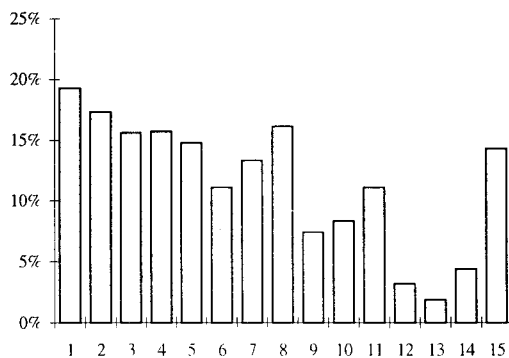


FIGURE 4: REFUSAL RATES ON "COLD" CALL ATTEMPTS BY HOUR OF THE DAY



If a string of dispositions for any telephone number includes any attempt coded as a refusal, it is given a value of 1 if the interview is completed in the end. Thus, we can calculate rates of refusal conversion (interview completion) according to the call attempt on which the initial refusal was met. From Figure 5 we can see a clear trend downward. Success in

FIGURE 5: REFUSAL CONVERSION RATES BY CALL ATTEMPT ON WHICH INITIAL REFUSAL OCCURRED



converting refusals is much lower when the initial refusal is met later in the call sequence. Also note it is important to recognize small sample sizes on the

later call attempts, as shown in Table 2, when looking to Figure 5 for evidence of relationships. Further, none of the conversion rates are greater than 20%.

TABLE 2: REFUSAL CONVERSION RATES BY CALL ATTEMPT ON WHICH THE INITIAL REFUSAL WAS ENCOUNTERED.

INITIAL REFUSAL CALL ATTEMPT	NO. OF CONVERSIONS	CONVERSION RATE	STANDARD ERROR
1	2379	19.27	.73
2	1109	17.30	1.03
3	655	15.59	1.30
4	488	15.72	1.51
5	393	14.75	1.65
6	232	11.11	1.94
7	195	13.33	2.27
8	130	16.13	2.95
9	125	7.41	2.25
10	99	8.33	2.66
11	64	11.11	3.70
12	61	3.17	2.21
13	53	1.85	1.83
14	44	4.35	3.01
15	12	14.29	9.35
16	2	0	0
17	1	0	0

Note: $\chi^2=54.014$, $df=1$, $p=.001$

Using logistic regression models generated by SUDAAN we attempt to predict the odds of completing an interview using interviewer and calling environment characteristics. The outcome is a dichotomous variable (1 if the interview is completed on that particular call attempt). The predictors in the model are interviewer and calling environment characteristics.

Interviewer characteristics obtained from the questionnaires sent to survey organizations were included in the model. Interviewer age is discussed as a contributor to refusal conversion success. It did not show significant effects here, and was removed from the model to avoid any collinearity effects with the experience measure. The measure of experience prior to BRFS calling was also removed. Most of the interviewers with more than one year BRFS experience had no prior experience. So to avoid skewed results in the model, it was removed as well.

Table 3 presents three main effects models of refusal conversion. All first-order interactions were included, but no significant relationships were found. One uses a scale measuring how many methods (of the four presented in the state questionnaire) were used to help interviewers convert refusals. The second model separated this scale and uses four dichotomous variables indicating use of each of the four methods and the third model omits these measures altogether in case

there are collinearity problems with some of the calling environment measures. The design effects attributable to stratification by state and month and clustering by phone number are noted in brackets in Table 3 as well. Some of them are quite high. Contrary to what we expected, the number of methods employed to aid refusal conversion has no association with refusal conversion success in the bivariate relationship test ($\chi^2=1.019$, $df=1$, $p=.313$). In the multivariate model it is not related to success either. Providing interviewers with call histories and giving them special training for converting refusals seems to increase success rates. However, providing reasons for initial refusals and providing interviewers with scripts to convert refusals seems to have no effect on refusal conversion success. Omitting the latter method from the model yields results identical to the remainder of Model 2,

completed by those with a script ($\chi^2=8.08$, $df=1$, $p=.004$).

Another characteristic of the calling environment included in the model is whether the outfit designates special interviewers for refusal conversion attempts. The coefficients do not test significant in either of the three models. The bivariate relationship also shows no significant association ($\chi^2=.023$, $df=1$, $p=.879$).

BRFSS experience was measured using a dichotomous variable with the value of 1 if the interviewer had over one year BRFSS experience. Unlike many telephone interviewers, these had many years of experience. This measure was used under the assumption that after the first year of interviewing, differential rates of refusal conversions are minimal. This measure proved significant in both models (though marginally in Model 2) with very similar coefficients.

Finally, being female did not show significant effects on refusal conversion rates here. We must note, though, that 97% of these call attempts were made by female interviewers. Thus, there may not be enough variation to detect a significant effect on refusal conversion rates.

TABLE 3: COEFFICIENTS FROM LOGISTIC MODELS PREDICTING REFUSAL CONVERSION

Predictor	Predicting conversion rates		
	MODEL 1 Using number of methods variable	MODEL 2 Using each method separately	MODEL 3 Omitting both methods measures
Intercept	-1.85 (.45) [1.50]	-3.16 (.39) [1.57]	-2.04 (.38) [1.23]
BRFSS exp. (months)	.50 ** (.24) [1.57]	.41 * (.25) [1.64]	.44 * (.24) [1.63]
Gender (Female=1)	.21 (.37) [1.20]	.50 (.41) [1.35]	.03 (.36) [1.17]
Designated special inter- viewers	.26 (.22) [2.13]	-.33 (.25) [2.91]	-.01 (.18) [2.91]
Number of methods	-.20 [2.62] (.16)		
Provided with call histories		1.28 ** (.39) [2.10]	
Give special training		.46 * (.25) [2.52]	
Provide scripts		.00 (.00)	
Provide reasons for initial refusals		-.50 (.41) [4.07]	

Note: (standard errors in parentheses) [design effects in brackets]

* $p < .10$

** $p < .05$

though examination of the bivariate relationship shows significant association. We found that 19.3% of the cases where no script was provided resulted in completed interviews, while only 15% were

DISCUSSION

These results suggest a number of implications for survey operations and future research. First, there is much to be learned from profiles of contact rates such as the ones presented here. Like many before us, we found that weekday evenings are the best times to reach respondents. Likewise, most BRFSS calls are made weekday evenings. Further, Fridays had considerably lower "cold" call refusal rates than any other day; however, of all the weekdays, the fewest calls are made on Fridays. Survey operations supervisors may wish to arrange calling schedules according to research findings such as these.

Refusal conversion is another worry for survey operations personnel. Organizations spend considerable time and money on training and/or special interviewers to increase refusal conversion rates. Here we found that conversion rates are higher when the refusal is met earlier in the call sequence. This may be a result of various factors and could itself be the focus of more detailed future study. The lower conversion rate may simply be the result of running out of time. One might be interested to know if these later refusals were met after a series of contacts to that number or if the refusal came on the first contact. Perhaps people got "tired" of being called and finally refused to do the interview. It would also be interesting to know how many call attempts were made after the initial refusal until the interview could be completed. Practically, though,

supervisors may consider concentrating efforts on refusal conversions when the refusal is met on an early call attempt to maximize chances of conversion success.

A distinction should be made between types of refusals: an initial refusal may be met before randomizing and choosing the adult with which to do the interview (a "household refusal") or the chosen adult may begin the interview and then refuse to finish (a "random adult refusal"). This distinction has practical implications that should be studied more specifically in subsequent work. A household refusal may be one made by some "gatekeeper", basically whoever answers the phone, who says no one is interested. On the other hand, a random adult refusal may come after the adult is selected and decides not to do the interview. One strategy to convert household refusals is to change the timing of the call attempt. If the initial refusal is from the "gatekeeper" and is met on a weekday evening, for instance, one tactic is to schedule the next attempt on a weekday morning, in hopes of reaching another eligible adult who will cooperate. A way to begin studying the effectiveness of this strategy is simply to look at the day and time an initial refusal is met and compare it to that of the completion attempt. Another possibility is to include in a model of conversion success a predictor that indicates whether the time and day of the refusal and conversion attempts are similar or different. In future work, it may be beneficial to select all the refusal attempts instead of sampling from them to ensure adequately large sample size for any such modeling.

There seem to be a few other factors supervisors should consider in attempts to improve survey cooperation rates. First, it is clear that more experienced interviewers are more successful in converting refusals. Survey organizations will also be more successful in gaining cooperation if interviewers are provided with call history information when calling a refusal telephone number. This may be a function of the same effect we see with interviewer experiences. Groves and Couper (1998) explain that experienced interviewers may be more confident, thus more successful. Along those same lines, interviewers with more information may also be more confident. They might also just have more tools with which to tailor their responses, i.e. if the interviewer knows why the respondent refused earlier and has a response to persuade them to cooperate. We see from the two logistic regression models, though, that it is less important to have numerous techniques to deal with refusals but an organization should concentrate on which techniques are more effective. Similar tests of other more specific methods organizations use to convert refusals could prove quite beneficial to survey operations personnel.

Beyond the potential operational causes of nonresponse in telephone surveys, other research must also explore behavioral dimensions of cooperation on the part of the sample member. Information from the substantive module of the interview could lend insight into respondent characteristics and interactions with interviewers. It may not suggest operational changes though. This research and what follows leaves open numerous possibilities. First, operations staff can keep in mind calling results and patterns. When scheduling calling shifts, supervisors should maintain high calling activity on weekday evenings and increase activity on Fridays, when refusal rates are lowest. When making callbacks to refusal telephone numbers, supervisors should concentrate efforts on those who refused on one of the first few attempts; they are more likely to successfully convert these. Finally, when scheduling and assigning interviewers to refusal conversion attempts, the most experienced ones should be used. When training, interviewers should be provided with and taught to understand the uses of information from prior refusal attempts to better equip them for refusal conversion interactions. These operational changes may result in higher response rates, lower refusal rates, and better overall data quality.

REFERENCES

1. Center for Disease Control and Prevention (1997). *Behavioral Risk Factor Surveillance system: Documentation for the 1984-1995 CD-ROM*. Atlanta, Georgia.
2. Groves, Robert and Couper, Mick (1998). *Nonresponse in Household Interview Surveys*. New York: Wiley.
3. Groves, R.M., and Couper, M.P. (1993b). "Multivariate Analysis of Nonresponse in Personal Visit Surveys." *Proceedings of the Section on Survey Research Methods*, American Statistical Association, pp. 514-519.
4. Groves, R., Cialdini, R., Couper, M. (1992). "Understanding the Decision to Participate in a Survey." *Public Opinion Quarterly*, 56, 4, 475-495.
5. Lessler, J.T. and Kalsbeek, W.D. (1992). *Nonsampling Error in Surveys*. New York:Wiley.
6. Lin, I-F., and Schaeffer, N.C. (1995). "Using Survey Participants to Estimate the Impact of Nonparticipation." *Public Opinion Quarterly*, 59, 2, 236-258.
7. Lengacher, J.E., Sullivan, C.M., Couper, M.P., and Groves, R.M. (1995). "Once Reluctant, Always Reluctant? Effects of Differential Incentives on Later Survey Participation in a Longitudinal Survey." *Proceedings of the Section on Survey Research Methods*, American Statistical Association, pp. 1029-1034.
8. Weeks, M.F., Kulka, R.A., and Pierson, S.A. (1987). "Optimal Call Scheduling for a Telephone Survey." *Public Opinion Quarterly*, 51, 540-549.