INVESTIGATING UNIT NON-RESPONSE IN A RDD SURVEY

Jeffery A. Stec, Paul J. Lavrakas, Elizabeth A. Stasny, The Ohio State University
Jeffery A. Stec, Center for Survey Research, The Ohio State University, 3045 Derby Hall, 154 N. Oval Mall, Columbus, OH 43210-1330

Key Words: Telephone Survey, Non-Response

I. Introduction

Among the many non-sampling errors that any given survey may contain, unit non-response error may be one of the most prevalent, most studied, yet least understood of the non-sampling errors. Its mystery lies in the fact that it requires knowledge of non-respondents' answers to the questions of the survey to overcome the potential error inherent in it. Yet, it is precisely that knowledge that non-respondents are unable, unwilling, or unavailable to give.

This has lead researchers to look at variables where there exists information on both respondents and non-respondents to examine differences between the two groups. These variables are commonly demographic variables obtained from some data source external to the survey. Examples of studies using this line of investigation are Woodburn (1991); Lavrakas, Bauman, & Merkle (1992); Groves & Couper (1993); Kojetin (1993); and Kennickell (1998).

This paper investigates unit non-response and whether it seems to lead to non-response bias by examining it in the context of a survey conducted by the Ohio State Center for Survey Research (OSU-CSR) called the Buckeye State Poll (BSP). The BSP is a Random Digit Dialing (RDD) survey. So, a contribution to the non-response literature that this paper will make is to investigate potential non-response error in a RDD setting. We note here that the conclusions must take into account the pitfalls of this method as well. Namely, the method of investigating unit non-response using external data suffers from the drawback of not having this external data for survey variables of interest. In other words, we will only have data for non-respondents for demographic variables. We will not have external data for many of the substantive survey variables. While we must temper our findings in light of these warnings, we can still contribute to the understanding of non-response for RDD studies and address whether it would be fruitful to model response probabilities and try to adjust for non-response by appropriate weighting techniques.

Section II of this paper will discuss the methods for data collection in the BSP and the means by which a data set was created for sample members. Section III will present the analyses on both a univariate and multivariate level for our external data set of respondents and non-respondents. We will examine the differences between respondents and non-respondents across variables where auxiliary information exists, while realizing that these auxiliary variables are not usually the variables of interest in a BSP. Section IV concludes our findings and presents directions for future research.

II. Methods

The Buckeye State Poll is a monthly RDD study conducted by the OSU-CSR. The survey consists of a set of economic questions followed by a host of demographic questions asked of a random sample of Ohio households. The primary sampling unit is the household with the secondary sampling unit randomly selected according to the most-recent birthday method (O'Rourke & Blair (1983), Salmon & Nichols (1983)). The economic questions ask respondents about business conditions, for their respective households as well as for the U.S. economy as a whole. These questions are identical to a set of economic questions that the University of Michigan's Survey Research Center has asked nationwide since the 1950's (Katona, 1975). A set of questions, original to the BSP, asking respondents about their households' use of credit cards and overall debt levels (Dunn, Stec, and Lavrakas (1999)) follows the Michigan questions. Finally the survey asks a set of demographic questions. Depending on how a given respondent answers certain questions, the length of the survey is approximately 7-12 minutes. Every month, on average, 500 BSP interviews are completed.

Since the survey is a RDD survey, the only information we have about a member of the sample (unless an interview is completed) is the phone number. This presents some unique challenges to gathering external data for studying non-contacts and non-respondents. Yet, this lack of initial information is common to every RDD study and has contributed to the lack of research on non-response for RDD samples. We overcome this problem by first matching the phone number to its zip code using software made by American Business Information Inc. The software, PhoneDisc®, was purchased by the OSU-CSR to generate RDD samples, but it has the additional use of providing a zip code for a majority of the RDD phone numbers in a given sample.

Once we have this external information, we can match the zip codes from the RDD phone numbers to Census information. Instead of utilizing 1990 U.S. Census information, which has obviously become quite dated, we use the Standard Demographics data base provided by Claritas, Inc. This software, purchased by the OSU, gives us current year estimates of population data based on the 1990 Census Summary tape File 1 (STF1) and Summary Tape File 3 (STF3). The combination of the phone numbers with the information in these two databases gives us information at the zip code level for members of our RDD sample.

For this study, we focus on nine months of BSP data, June 1998 through February 1999. The OSU-CSR uses RDD phone number samples that are not first "cleaned" of ineligible numbers. Roughly one half of the phone numbers sampled over the nine-month period contain ineligible or unanswered numbers. Eligible phone numbers are defined as those phone numbers that have at one point or another in their call histories led to a contact with a household.[1] The outcome of that contact could have led to anything from a completed interview to an outright refusal to participate in the survey.[2]

The cooperation rate (Groves & Lyberg (1988)) for the data examined here is 57.74% and it is the percentage of interviews to all contacted cases capable of being interviewed. We also include a measure of how well the sample was alerted to the survey. The contact rate is 75.11% and it represents the number of contacted households to the number of eligible numbers.

There is a significant proportion of non-response. However, this does not necessarily mean that there is non-response bias. It is possible that the non-respondents to the BSP do not differ from respondents across variables of interest or demographic variables to any significant degree. Non-response bias exists only if the estimate of a population statistic derived from the sample would be significantly different had non-respondents' answers also been included. It is this important issue that the remainder of this paper will address.

III. Respondents vs. Non-respondents using External Data

This section begins the analysis of non-response found in the BSP by focusing on the difference between respondents and non-respondents. We do this by utilizing the external data set that was described in the Methods section. In 76.4% of the cases, external information was available for a non-respondent. A phone number that is not listed in our PhoneDisc® database is not considered in our study because we cannot accurately obtain a zip code for that phone number.[3]

Drew, Choudhry, and Hunter (1988) found in their study of the Canadian Labour Force Survey that non-response rates were slightly higher for households with unlisted phone numbers. They attribute the higher non-response rates for the RDD component of that survey to the higher refusal rate for that component. Likewise, we find that the percentage of cases that have listed phone numbers and refuse to complete a BSP interview is 26.2%, while cases that have unlisted phone numbers and refuse to complete a BSP survey is 36.6%. The difference between these two large sample percentages is statistically significant at a greater than 99% level of confidence.

Drew *et al.*, however, found that average household size and unemployment rate for households with listed vs. unlisted phone numbers was very similar. This would seem to indicate that, while households with unlisted numbers refuse to participate in RDD surveys more than households with listed numbers, there might not be substantial differences between the two groups, at least for the variables that their study considered. Unfortunately, it is impossible for our current study to present findings on this issue since our whole methodology is predicated on cases' phone numbers being listed in our phone database. However, Drew *et al.* finding that there are not significant differences between listed and unlisted refusals at least plausibly confirms our feelings that we can gain substantial insight into the potential non-response error in the BSP survey data via our method.

[1] This leaves out some of those numbers that are ultimately resolved as non-contacts at the end of the survey period. Unfortunately, their call histories give no indication whether the phone numbers in this category are eligible households or ineligible numbers, so we do not include them in the category "non-contacts" discussed later in this paper.

[2] Specifically, a contact is any household that gives permission to call back again, any household that is eligible but not capable of or not available to complete an interview during the survey period, any household that is eligible but does not have an English-speaking respondent, any household that refuses to participate at either the household or respondent level, and any household that partially completes or successfully completes an interview with the CSR.

[3] A notable exception is completed interviews. In the majority of cases, we have a zip code, obtained from the interview, for cases where the phone number does not have a listing in our phone number database. For roughly 67% of the completed interviews, there is no zip code in the phone database. However, we have zip codes, whether from the database or the completed interview, for 99.3% of all the completions.

A case in the external data set represents a completion, refusal, or non-contact[4] household. For each case, we have zip code information that allows us to match the observation with its appropriate projected Census information. So, the external data for a given case is this information at the zip code level. Our rationale for focusing on zip code level comparisons is due to the data constraints by which we are bound. While our phone number database also gives us the name, address, and city for a given case, we have no practical way of using this information to obtain household level data. One of the appealing qualities of this study is that we can use our method to focus on very general RDD studies. In other words, the RDD sampling is not constrained to come from a "special" sampling frame for which external data readily exists for sample members. This problem has plagued many of the non-response studies that use external data sets because it does not readily generalize to less "special" sampling frames. Our study methodology could be applied to any RDD study straightforwardly.

The variables in the external data set that we investigated are: median household income, median household wealth, age, education, occupation, marital status, number of household members, a social and economic achievement score (quality score), and population density.

Table 1 presents the differences across these variables between the values for the completions, refusals, and non-contacts as well as displaying which differences are statistically significant.[5]

We find that there is a statistically significant difference between completions (n = 4,920) and refusals (n = 2,480) at the zip code level across median household income, median household wealth and quality score. The difference indicates that household refusals tend to come from zip codes with higher levels of income, wealth, and social and economic achievement when compared to the zip codes of household completions. It is not difficult to propose that zip codes with higher levels of median household income, median household wealth and quality scores will contain more households with higher levels of these variables. The converse is also true. We can then infer that a higher median income, higher median wealth, higher quality score household is more likely to refuse to complete the RDD survey considered here than is a lower median income, lower median wealth, lower quality score household. This conclusion follows by virtue of the facts that sample members who refuse to complete a survey tend to come from zip codes that have higher median income, higher median wealth, and higher quality scores. Also, households in higher median income, higher median wealth, and higher quality score zip codes are more likely to be higher income, higher wealth, and higher quality score households. This is the crux of the argument we use to explain the importance of differences between interviewed cases, refused cases, and non-contacted cases across the variables we consider here.

Table 1 also examines household completions against household non-contacts (n = 852), that is, households that were eligible to be surveyed, but that could not be resolved into completions or refusals by the end of the survey period. It appears that this comparison shows no differences between income at the zip code level for the two groups. But, there is a marginally significant difference for median household wealth. The population density variable is strongly, statistically significant. Non-contacts are from zip codes that, in general, have higher levels of economic achievement as well as being more highly populated.

The impetus for the last column in Table 1 is the fact that much of the survey research literature does not concern itself with the potential differences between sample members that refuse to participate vs. those that just are not contacted within the survey period (cf. Groves and Couper, 1998). However, if there is a difference between the two groups, this implies different post-survey adjustment models should be used to model non-response. Namely, models should be used that account for the different influences that drive refusals vs. non-contacts. However, in this instance, there do not appear to be any significant differences between sample members, at the zip code level, who refuse to participate and sample members, at the zip code level, who are not contacted during the survey period.

Table 2 summarizes the comparisons of completed interviews with refusals, completed interviews with non-contacts, and refusals with non-

[4] A non-contact, for the purposes of this study, is a case that has as its final disposition a scheduled or general callback or a busy signal/no answer/answering machine outcome that has been confirmed to be an eligible number. We do not include households that are physically or mentally unable to do the survey; households that are unavailable during the survey period; and non-English speaking households because, while they are still non-respondents, typically there is little a survey researcher can do *ex ante* to combat this type of non-response. Moreover, the number of cases where these dispositions occur is relatively small. For our study, they comprise only 5.6% of all the cases.

[5] The use of the large sample t-test for the difference between two population means requires that the two samples are randomly selected in an independent manner from each of the populations. Moreover, the sample sizes must be large enough so that the sample means are approximately normally distributed. If we assume that the sample interviews, refusals, and non-contacts are randomly and independently sampled from the population of interviews, refusals, and non-contacts respectively then the large sample t-test for the difference between two population means is valid here.

contacts for the categorical variables. The values in the table are the difference in the average of the percentages for that category from the external data set. In other words, each zip code has a percentage of people who fit a certain category, say, 18-24 year olds. The value for a given member of the sample for the 18-24 year old category is the percentage of people within that zip code who are 18-24 years of age. The difference in the average of these percentages across zip codes for select groups of sample members is the value of that category given in the table.

The table examines the differences between our three groups for statistical significance. Sampled cases where a completion occurs are more likely to have respondents from zip codes where there are younger, blue-collar residents in larger size households relative to sampled cases where a refusal or non-contact occurs. Refusals, besides having the characteristics already mentioned, are more likely to come from zip codes with higher concentrations of married and African-American residents. Again, there is little difference between refusals and non-contacts at the zip code level. The most notable is that refusals tend to reside in zip codes with higher concentrations of married individuals. Non-contacts reside in zip codes with fewer married individuals. Procedures that focus on repeated callbacks and not on refusal conversions might, therefore, lead to married households being underrepresented. Conversely, survey procedures that focus on refusal conversions and not on repeated callbacks might tend to have unmarried households underrepresented in the interviewed sample.

All these univariate conclusions must be tempered with the following note. If we control for some variables when looking at the tendency for sample members to respond (or not to respond) to the survey, then we may find that other variables no longer significantly explain differences between respondents and non-respondents. Thus, we examine completions, refusals, and non-contacts in a multivariate context with a logistic regression for each sample group. Recall from our univariate analysis that age, occupation, marital status, household size, race, household median income, household median wealth, and quality score, appeared to identify, at the zip code level, sample members who chose to respond to the survey. Similarly, age, occupation, marital status, household size, race, household median income, household median wealth, and quality score seemed to matter at the zip code level for refusals; and, age, education, occupation, household size, household median wealth, and population density helped explain non-contacts.

Table 3 provides the output from stepwise logistic regressions[6] where the dependent variables are whether a completed interview, a refusal, and a non-contact were obtained, respectively. The independent variables that were entered into the model at the first iteration are median household income, median household wealth, quality score, population density, age, education, occupation, marital status, household size, and race. The Hosmer-Lemeshow statistic for goodness of fit of the final model for completions is 7.819 with 8 degrees of freedom. This gives a p-value of 0.4514 indicating that the model appears to fit the data well (H_0: model is correct). From the table, we see that when we control for the effects of the variables simultaneously, only age, marital status, race, and quality score appear to matter when trying to explain completions. Household that tend to complete interviews come from zip codes that have larger numbers of younger, married respondents that have larger concentrations of African-American residents and that have lower quality of life scores. These results imply that what is driving the probability of a completed interview here is different from what the univariate analyses suggest. This is because once we control for all the variables simultaneously we find that some of the variables that appeared to be significant in the univariate analyses are not significant here.

We performed similar analyses both for cases where a refusal to be interviewed resulted and for cases where survey interviewers were not able to contact the respondent. Again, Table 3 has the results of the backward Wald stepwise regression procedure. The Hosmer-Lemeshow statistic for goodness of fit of the final model for refusals is 9.6590 with 8 degrees of freedom. This gives a p-value of 0.2898 indicating that the model appears to fit the data well (H_0: model is correct). According to the model, refusals are more likely to come from zip code where there are higher concentrations of older residents who are more likely to be Caucasian, have lower median household wealth and higher quality of life scores.

The last stepwise logistic regression procedure for non-contacts in Table 3 has a Hosmer-Lemeshow statistic for the final model of 11.2608 with 8 degrees of freedom. This gives a p-value of 0.1874 indicating that the model appears to fit the data well (H_0: model is correct). According to the model, eligible non-contacts are more likely to come from zip codes where there are higher concentrations of the oldest category of people,

[6] The exact stepwise method is a backward Wald method. We used a stepwise logistic regression because some of the variables might be fairly highly correlated and would rightly drop out in this type of procedure. We also used the backward stepwise likelihood ratio method that gave the same results.

who are more likely to have some high school education and whose quality of life score is higher.

Overall, the multivariate results seem to suggest that households within zip codes in which there are higher percentages of younger, married residents with lower quality of life scores would be the most likely to complete a BSP survey. Households in zip codes that have higher concentrations of older residents with higher quality of life scores tend to count for more refusals and households in zip codes that have higher percentages of the oldest age category residents with higher quality of life scores would have more non-contacts. These results, of course, control for the potential effects of other independent variables on the respective dependent variable. They also clarify the univariate results in which there appeared to be more significant influences on completions, refusals, and non-contacts. There do appear to be significant differences between respondents and non-respondents at the zip code level. This would imply that unit non-response bias is a strong possibility for the survey examined here whenever the response rate is not close to one.

IV. Conclusions

There appear to be statistically significant differences between completions, refusals, and non-contacts for a RDD survey. Thus, a RDD survey that does not take action to avoid and adjust for unit non-response may suffer from non-response bias due to the fact that the sample statistic for respondents is significantly different from the sample statistic for non-respondents for a given question. In this paper, we have found a statistically significant difference across a multitude of demographic variables by respondents' zip code (see Tables 1 through 3). By the argument we proposed at the beginning of this paper, differences in these variables at the zip code level suggest similar differences in these variables at the household level. Moreover, differences in demographic variables suggest that there may also be differences in substantive variables.

Survey operations that do not adequately try to combat refusals and non-contacts while the particular survey is in the field may suffer from non-response bias. What this study adds to that general rule of thumb is that, by focusing on either multiple callbacks or refusal conversions, a survey may still suffer large errors due to non-response. This is because there appear to be some statistically significant differences between the sample members who refuse vs. the sample members who are non-contacts.

Future research efforts will focus on examining how different hard-to-reach respondents and initial refusal respondents are from non-contacts and permanent refusers, respectively. One form of action that a survey research organization can take to combat error due to unit non-response is to make repeated callback and refusal conversion attempts. If hard-to-reach completions and refusal conversion completions are good proxies for non-respondents, then the addition of these types of respondents to a given sample should make that sample more representative of the general population. However, if these types of respondents are not good proxies for the non-respondents then research efforts that make a high number of repeated callback attempts and refusal conversion attempts a priority will be misusing their resources. This question can be answered with the external data set used in this paper.

REFERENCES

Drew, J. D., G. H. Choudhry, and L. A. Hunter. (1988). "Non-Response Issues in Government Telephone Surveys", In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nichols, and J. Waksberg (Eds.), *Telephone Survey Methodology*. New York: John Wiley & Sons, pp. 233-46.

Dunn, L. F., J. A. Stec, and P. J. Lavrakas. (1999). "New Index Tracks Consumer Debt Condition", Working Paper.

Groves, R.M. and M.P. Couper. (1998) *Nonresponse in Household Interview Surveys*. Wiley & Sons, New York, NY.

Groves, R. M. and M. P. Couper. (1993) "Multivariate Analysis of Non-Response in Personal Visit Surveys", Proceedings of the Section on Survey Research Methods, 1993 Annual Meeting of the American Statistical Association, San Francisco, CA.

Groves R.M and L.E. Lyberg. (1988) "An Overview of Non-Response Issues in Telephone Surveys", in *Telephone Survey Methodology*, Eds. R.M. Groves et al. John Wiley & Sons, New York, NY.

Katona, G. (1975) *Psychological Economics*. New York : Elsevier Scientific Pub. Co.

Kennickell, A. B. (1998) "Analysis of Non-Response Effects in the 1995 Survey of Consumer Finances", Paper presented at the 1997 Joint Statistical Meetings, Anaheim, CA.

Kojetin, B. A. (1993) "Characteristics of Non-Respondents to the Current Population Survey (CPS) and Consumer Expenditure Interview Survey (CEIS)", Proceedings of the Section on Survey Research Methods, 1993 Annual Meeting of the American Statistical Association, San Francisco, CA.

Lavrakas, P. J., S. L. Bauman, and D. M. Merkle.

(1992) "Refusal Report Forms (RRFs), Refusal Conversions, and Non-Response Bias", Paper presented at the May 1992, AAPOR Conference, St. Petersburg, FL.
O'Rourke, D., & J. Blair. (1983). "Improving Random Respondent Selection in Telephone Surveys", *Journal of Marketing Research*, 20, pp. 428-32.
Salmon, C. T., & J.S. Nichols. (1983). "The Next-Birthday Method for Respondent Selection", *Public Opinion Quarterly*, 47, pp. 270-6.
Woodburn, R. L. (1991) "Using Auxiliary Information to Investigate Non-Response Bias", Proceedings of the Section on Survey Research Methods, 1991 Annual Meeting of the American Statistical Association, Atlanta, GA.

Table 1: Differences in the average values for continuous variables by completions, refusals, and non-contacts

	Completions vs. Refusals	Completions vs. Non-Contacts	Refusals vs. Non-Contacts
Median Household Income	-$694*	-$485	$209
Median Household Wealth	-$2,025**	-$1,774†	$251
Quality Score	-0.50**	-0.48	0.01
Population Density	-0.72	-2.12*	-1.40

The difference is obtained by subtracting the second variable listed in the heading from the first variable listed in the heading. The statistical test performed here is large sample t-tests for the difference between population means. † = 90% confidence, * = 95% confidence, and ** = 99% confidence.

Table 2: Differences in the average proportions for categorical variables by completions, refusals, and non-contacts

	Completions vs. Refusals	Completions vs. Non-Contacts	Refusals vs. Non-Contacts
AGE			
18-24	0.0054**	0.0034*	-0.0020
25-34	0.0027**	0.0010	-0.0017
35-44	-0.0004	0.0003	0.0007
45-54	-0.0010†	0.0006	0.0016†
55-64	-0.0016**	-0.0015*	0.0001
65+	-0.0052**	-0.0039*	0.0014
EDUCATION			
GRAMMAR	0.0013	0.0002	-0.0011
SOME HS	0.0023	-0.0002	-0.0025
HS GRADUATE	0.0011	0.0066*	0.0055
SOME COLLEGE	-0.0011	-0.0013	-0.0002
COLLEGE DEGREE	-0.0035	-0.0056	-0.0021
OCCUPATION			
WHITE COLLAR	-0.0056†	-0.0081†	-0.0025
BLUE COLLAR	0.0056†	0.0081†	0.0025
MARITAL STATUS			
MARRIED	-0.0047*	0.0016	0.0063†
NOT MARRIED	0.0047*	-0.0016	-0.0063†
HOUSEHOLD SIZE			
1 PERSON HSHLD	-0.0021	-0.0038	-0.0017
2 PERSON HSHLD	-0.0016†	-0.0009	0.0007
3-4 PERSON HSHLD	0.0020	0.0034†	0.0015
5+ PERSON HSHLD	0.0016*	0.0012	-0.0003
RACE			
WHITE POPULATION	-0.0073†	0.0042	0.0115
BLACK POPULATION	0.0073†	-0.0042	-0.0115

The difference is obtained by subtracting the second variable listed in the heading from the first variable listed in the heading. The statistical test performed here is large sample t-tests for the difference between population means. † = 90% confidence, * = 95% confidence, and ** = 99% confidence.

Table 3: Logistic Regression for Completions, Refusals, and Non-Contacts (n = 9,464)

Variable	Completions	Refusals	Non-Contacts
CONSTANT	-1.4288**	-0.2052	-3.7305**
AGE 18-24	3.8834**	-2.7618**	
AGE 25-34	4.0641**	-4.7764**	
AGE 45-54	-	-	-6.0010**
AGE 55-64	-	-	4.9861*
MARRIED	2.3383**	-	
WHITE POP.	-0.5555**	0.2673†	
SOME HS EDUC.	-	-	2.5818*
MED. HSHLD WEALTH	-	-6.1E-06**	
QUALITY SCORE	-0.0117**	0.0123*	0.0302**

† = 90% confidence, * = 95% confidence, and ** = 99% confidence.