BALANCING COST AND MEAN SQUARED ERROR IN RDD TELEPHONE SURVEYS: THE NATIONAL IMMUNIZATION SURVEY

K.P. Srinath, Michael P. Battaglia, Jessica Cardoni, Corinna Crawford, Richard Snyder,
Abt Associates Inc.;
Robert A. Wright, National Center for Health Statistics, Centers for Disease Control and Prevention
K.P. Srinath, Abt Associates Inc., 1110 Vermont Avenue, Washington, D.C. 20005

Key Words: Random-digit dialing, CASRO response rate, Number of call attempts, Nonresponse

1. Introduction

Nonresponse is a major problem in random-digit-dialing (RDD) surveys. Nonresponse can occur at various phases of data collection, especially if the eligible sample is identified through the selection of a larger screening sample. For example, if we are interested in obtaining data from households containing children in a specific age group, these households will have to be first identified by screening a larger sample of households. The larger sample of households is contacted through calling an even larger sample of telephone numbers. It is well known that a large proportion of telephone numbers is nonresidential or nonworking numbers. Therefore, when there is nonresponse to the initial call attempts to identify whether a selected number is residential, nonworking or nonresidential (e.g., business), there is a possibility that the number is residential but we have failed to identify the number as such (e.g., no answer to all call attempts). The ratio of the number of telephone numbers resolved as either residential, nonresidential or nonworking to the total in the sample is called the *resolution rate*. Similarly, when a sampled number is identified as residential, we may fail to establish whether the household is eligible for the survey. The ratio of the number of households deemed eligible or ineligible to the number of known households in the sample is called the *screener response rate*. Finally, there is the usual problem of nonresponse, which occurs in all surveys, when we fail to get data from a household that has been identified as eligible for the survey. The ratio of the number of eligible households that respond to the total number of eligible households is the *interview response rate*. The *overall response rate* for the survey is the product of these three rates (Ezzati-Rice et al., 2001). Therefore, it is essential that in RDD surveys we try to achieve a high response rate at each phase so that the overall response rate is reasonable.

The response rates, in addition to determining the number of completed interviews on which the survey estimates are based, also determine the sampling weights that are attached to the respondents for obtaining population-based estimates. Since at each phase, we have to adjust the sampling weights to account for nonresponse, the adjustments depend on the rates achieved. A less than desirable response rate will lead to estimates that may be biased and also have a large variance due to reduced sample size. This results in estimates with a large mean squared error. Therefore, to reduce the bias and also the mean squared error, a common practice in RDD surveys is to make several call attempts to identify the status of the sampled unit at each phase and to collect data from the eligible households. Another technique to reduce the bias due to nonresponse in RDD surveys is to make a concerted effort to convert refusals to respondents.

It is obvious that we cannot have an unlimited number of attempts to collect data, because of a ceiling on the total survey cost and due to time constraints. For example, some later attempts may result in additional completed interviews reducing the bias and the variance but may also add substantially to the cost of the survey. In other words, the size of the reduction in the mean squared error may not be worth the increase in the cost of the survey. Therefore, it is important to determine the number of attempts that minimizes the mean squared error of the estimate for a given total survey cost. In this paper, we examine one approach to comparing the mean squared error and cost at various numbers of attempts. The method suggested is a simpler version of the method of comparison used by Deming (1953).

We apply this method to the National Immunization Survey (NIS), which is conducted by the Centers for Disease Control and Prevention. The NIS is a very large RDD survey that screens for households with children aged 19-35 months. Smith et al. (2001) and Zell et al. (2000) provide further details on the NIS design. The NIS attempts to maintain a high overall response rate and therefore has a relatively high ceiling on the number of call attempts. Therefore, it is a good candidate for determining the optimum number of call attempts. Previous research on the number of call attempts focused only on minimizing response bias (Dennis et al., 1999), whereas in this paper we address both cost and mean squared error issues, as well as the reduction in bias due to conversion of refusals.

2. Description of the NIS

Independent random samples of telephone numbers are drawn each quarter from the 78

Immunization Action Plan (IAP) areas, which consist of the 50 states, the District of Columbia and 27 selected urban areas. The samples for each IAP are drawn at the beginning of the data collection quarter. Before the sampled telephone numbers are released to the telephone centers, they are processed through a business number matching process and auto-dialer to eliminate as many business and non-working telephone numbers as possible.

Resolution of the telephone number. The computer assisted telephone interviewing (CATI) system assigns each selected telephone number to the phone center and the number is dialed until it is resolved as a household, a non-residential telephone number, or a non-working number.

Screening for eligibility. The household numbers are called repeatedly in an attempt to screen them as eligible or ineligible for the NIS interview. The goal of this screening process is to identify the estimated 3.7 percent of households in the sample that have children between the ages of 19 and 35 months. A brief screening interview is administered to ascertain if the household fulfills that criterion. If the household has no children, or has children outside the age range, it is screened out, and the number is not dialed again. For those households in which there is at least one child within the eligible age range, the screening process identifies the resident who is the most knowledgeable person (MKP) concerning the vaccination records and general health of the children who are living in that household.

Data collection and consent. Once the interviewer identifies the MKP, she or he collects the shot record information for all eligible children. Specifically, the interviewer asks the MKP to locate the child's shot records and report the dates of all shots administered to the child. If a shot card cannot be located, the child's vaccination report is based on respondent recall. Demographic information about the household is then collected, such as child's race and ethnicity, race and ethnicity of the child's mother, mother's education, and information about other telephone lines in the household. After the respondent answers these demographic questions, the interview continues to the next section of the questionnaire, where information is collected about the child's vaccination providers. The respondent is asked for the name and address of those providers. The MKP is then asked for consent to contact the providers in order to release the child's shot records. After interviewers receive consent from the MKP, the providers are contacted and immunization history data are collected from them.

Currently, each case is called as necessary until it is resolved, screened, or completed. The case delivery system determines how and when the case will be delivered to an interviewer. After each call, the case is assigned a disposition code indicating what its resolution status is. These codes are updated throughout the quarter, as the case is worked. During a quarter, almost 8 percent of the sample (around 60,000 telephone numbers) is dialed with no answer. In 1999, the methodology for finalizing these cases changed, so that the case was finalized as a respondent (i.e., completed interview) or a nonrespondent or an ineligible etc.after the 15^{th} call attempt, rather than the 24^{th} call attempt. Our analysis showed that these cases have very little probability of being resolved after 15 dial attempts (Dennis et al., 1999). However, because of the desire to maintain high overall response rate, cases that are *not* non-contact cases have no limit on the number of dial attempts that they receive.

This study looks at the Record of Calls (ROC) file for 4 recent quarters (Q4/1999-Q3/2000). It lists every call and its outcome (case disposition), for every case in the sample. In one year, there are over 2 million telephone numbers released to the phone center. Each ROC file contains around 3 million records relating to call attempts and call outcomes for all the cases released during that quarter.

Our goal in this paper is to apply a statistical method for examining the impact of limiting the number of call attempts made on all cases, not only generic non-contact cases. Data were analyzed at 5 different maximum attempt levels-5, 8, 10, 12, and 15 call attempts-and compared to the final sample at the end of the data collection period. We assigned a final case disposition to each household telephone number at each of five maximum attempt levels. Data on the response rates at each maximum attempt level are also analyzed, as the study relies both on achieving all targeted number of respondents and maintaining a high overall response rate.

The response rate is calculated using the definition recommended by the Council of American Survey Research Organizations (CASRO). This is called the CASRO rate (Ezzati-Rice et al., 2001). For the NIS, the CASRO rate is calculated as the product of the three rates, which are the resolution rate, screening rate, and interview completion rate. The final CASRO response rate for the data collection period (Q4/1999-Q3/2000) was 79.28%. Table 1 reports all of the response rate components for several maximum number of call attempt levels.

In addition to examining the impact on the response rate of limiting the maximum number of attempts to 5, 8, 10, 12, or 15 attempts, refusal conversion was also analyzed. Current procedure on the NIS is that if a case is considered a refusal (respondent hangs up, respondent refuses to answer further questions, etc.) then the case is moved to a separate queue. The case delivery system allows the case to "cool off" for a period of 3 days to a week, and the case is re-assigned to a different type of interviewer, known as a Refusal Converter. This type of interviewer is more experienced and is highly skilled at converting refusals.

For Q4/1999-Q3/2000, all cases that received one or more refusal conversion attempts were identified in the final sample. We assigned a final case disposition to these sample telephone numbers as if no refusal conversion attempts had been made and assuming that these cases were finalized as various categories of refusals. We then calculated the response rate components and the overall CASRO response rate. Clearly, these rates (except the resolution rate) are lower if refusal conversion is not attempted. Table 2 illustrates this.

The main purpose of this paper is to examine bias versus cost trade-off in limiting the maximum number of call attempts. We will also examine bias reduction from refusal conversion efforts in the NIS. The NIS estimates vaccination coverage levels for the U.S., states, and IAP areas. A key vaccination outcome measure is the 4:3:1:3 series: 4 or more diphtheria and tetanus toxoids and pertussis (DTP) vaccinations, 3 or more poliovirus (polio) vaccinations, 1 or more measles-containing vaccinations (MCV), and 3 or more *Haemophilus Influenzae* type B (Hib) vaccinations.

For each maximum attempt level, weights were calculated for each child with a completed interview. The same procedure was followed for the file of completed interviews that resulted without refusal conversion. Estimates of vaccination coverage were developed for the U.S. and the 78 IAP areas. Standard errors were calculated by the Taylor series approximation method using SUDAAN.

3. Cost Data

The cost elements underlying the NIS data collection are extensive, with a mix of cost behavior types in an ever-changing environment. For purposes of this paper, the actual cost per completed case was replaced with a proxy, specifically, timing data supplied by the NIS CATI system. Rather than look at actual cost per completed case for each of the call attempt data points, the relative hours per case for each data point was calculated.

The time recorded in the CATI system is assumed to be indicative of the relative cost to collect data; ratios of hours per case derived from the CATI system timing data would also represent the ratios of total interviewer labor necessary to collect data (not all interviewer labor is recorded by the CATI system). In turn, ratios of hours per case would be an accurate proxy for the relative ratios of data collection cost per completed case. For example, if the average interviewer labor is twice as great for one case compared to another, the cost to collect data for the first case is also twice as great as the cost to collect data for the second case. It is important to note that the analysis focused on NIS-only cases, excluding optional sample supporting supplemental topics investigated on the NIS contract through its sampling frame. Given the overlap of interviewer activities on such cases, and the inability to segregate interviewer labor between the NIS and optional activities, the latter were eliminated from this analysis.

4. Optimum Number of Call Attempts

Assume that we are interested in estimating a population proportion of some characteristic of interest relating to eligible households (e.g., proportion of children who are up-to-date on their vaccinations). Let P denote the population proportion of interest. Let $j =$ 1,2, .. …..U denote the number of attempts. U denotes the maximum number of attempts to collect data before designating the household as a nonrespondent. Let p_j represent the estimate of the population proportion P based on n_j, the number of completes after j attempts. Let $E(p_j) = P_j$. Therefore, the bias in the estimate is $(P_j - P)$. We expect that estimates based on a larger number of completes to have less bias due to nonresponse than those based on a smaller number of completes. Since U is the maximum number of attempts giving the largest number of completes n_U, we assume that $P_U = P$. That is, after U attempts resulting in n_U completes, p_U the estimate based on n_U is unbiased or almost unbiased for the population proportion P. Only the estimate based on the maximum number of attempts is assumed to be unbiased. Let the variance of p_j be $V(p_j)$. Then the mean squared error of p_j is

$$V(p_j) + (P_j - P)^2 .$$

If the observed cost per completed interview after j attempts is c_j, then the total cost of the survey after j attempts is $n_j c_j$. The total cost rises with an increase in the number of attempts due to a greater number of completed cases and also a higher cost per completed case because of increased cost of later attempts to collect data from harder to reach/more resistant households. Since the variance and the bias based on later attempts should decrease, but at an increased cost, it is useful to look at the decrease in the mean squared error (variance plus squared bias) holding the cost of the survey fixed. That is, since the reduction in the mean squared error of the estimate is achieved by increasing the cost, we adjust the mean squared error for a fixed cost and then compare the adjusted mean squared errors for various numbers of call attempts. The adjustment is done as follows.

Let C denote the total budget available for the survey or for data collection. Let the minimum number of call attempts we want to make be t. Given that C is fixed and the cost per complete after t attempts is c_t, the expected number of completes that we would get after t attempts is m_t where $m_t = \frac{C}{c_t}$. The initial sample size of telephone numbers that we need to contact to get m_t completes after t attempts is obtained by dividing the number of completes by the product of the interview completion rate, eligibility rate (percent of households that have children between 19 and 35 months), the screener completion rate and the percent of telephone numbers that are known household numbers out of the total number of telephone numbers. Let this product at the end of t attempts be denoted by r_t. The initial sample of telephone numbers is $m_0 = \frac{m_t}{r_t}$. If we start with an initial sample size of m_0, then apply the response rates specified above at various numbers of attempts, we can compute the expected number of completes at the end of these numbers of attempts. If the expected number of completes at the end of $(t+k)$ attempts (where k is between 1 and $U-t$) is n_{t+k} and the cost per complete is c_{t+k}, then the total cost is $c_{t+k}\, n_{t+k} > C.$

We can also compute the expected number of completes for a given fixed cost. For example, after $(t+k)$ attempts, since the cost per complete is c_{t+k}, we can only have $m_{t+k} = \frac{C}{c_{t+k}}$ completes if we want to keep the cost fixed. We now take the ratio of the expected number of completes with increasing costs to the expected number of completes with a fixed cost, which is equal to $\frac{n_{t+k}}{m_{t+k}}$. These ratios are used to inflate the variance obtained using the number of completes with increasing costs. This inflated variance is used in computing the mean squared error with a fixed cost. For example, the ratio after $(t+k)$ attempts is $\frac{m_0 r_{t+k}}{C / c_{t+K}}$. Substituting for $m_0 = \frac{m_t}{r_t}$ and letting $C = c_t m_t$ we get the ratio as $\frac{r_{t+k}\, c_{t+k}}{r_t c_t}$. The computed ratio equals 1.00 after t attempts by definition as t is considered to be the minimum number of attempts.

We want to identify the number of attempts $t+k$ that minimizes the following alternative expression for the mean squared error containing the adjusted variance:

$$V(p_{t+k})\left(\frac{c_{t+k}\, r_{t+k}}{c_t\, r_t}\right) + (P_{t+k} - P)^2 .$$

5. Application to the NIS

As indicated in section 2, we evaluated the cost, response rates, bias and variance data from the NIS at 5 different maximum attempt levels: 5, 8, 10, 12, and 15 call attempts. Table 3 shows the ratio of the defined rates $\frac{r_{t+k}}{r_t}$ and the cost per complete $\frac{c_{t+k}}{c_t}$ for these attempts and the adjustment factors to inflate the variance. Table 4 shows the variance and bias of the estimates of 4:3:1:3 vaccination coverage rates over the four quarters (Q4/1999-Q3/2000) after each of these attempts.

In our analysis, 5 attempts are considered to be the minimum number of call attempts for data collection. The rate ratios $\frac{r_{t+k}}{r_t}$ and cost ratios $\frac{c_{t+k}}{c_t}$ are computed using the rates for 5 attempts as the base. That is, we have $t = 5$, $k = 3, 5, 7$, and $U = 15$. We assume that the estimate of the vaccination coverage after 15 attempts is unbiased.

Table 4 also shows the increased variance to offset the increase in the cost and the adjusted mean squared error of the 4:3:1:3 vaccination up-to-date estimates.

It is seen from the above table that the minimum mean squared error is at 12 attempts. The mean squared error is considerably higher at 5, 8 and 10 attempts.

6. Impact of Refusal Conversions

For Q4/1999-Q3/2000 277,964 cases were identified in the sample that received one or more refusal conversion attempts. To assess the impact of refusal conversion, we assigned a final case disposition to these sample telephone numbers assuming no refusal conversion attempts were made and that these cases were finalized as various categories of refusals. The entire weight calculation process was then carried out so those estimates of vaccination rates without the benefit of refusal conversion could be compared with the final vaccination estimates. At the national level 76.67% of children are up-to-date on all of the four key childhood vaccinations. Without refusal conversion the national estimate is 76.13%, which is lower by 0.54 percentage points

7. Conclusions

It is possible to determine the optimum number of call attempts that minimizes the mean squared error in RDD surveys provided detailed information on cost of data collection, response rates, variance and bias in the estimates are available. Of course, it is difficult to estimate the exact bias. An estimate of the bias in the estimates obtained from early attempts can be made by assuming that the estimate obtained at the end of the maximum number of attempts for the survey is approximately unbiased.

In this paper we find that roughly 12 call attempts seems to be optimum for the overall sample. But this needs more investigation, as we have not looked at the 78 IAP area estimates. The primary goal of the NIS is to estimate vaccination rates for the 78 IAP areas. Because some IAP areas tend to have a lower response rate, the application of a finding from the overall sample to such IAP areas could have a negative impact on the NIS. Therefore, we plan to further examine the issue of optimum number of call attempts for individual IAP areas or IAP area groups.

We also conclude that refusal conversion has a substantial impact on the response rates and a modest impact in most but not all IAP areas on the vaccination coverage rates.

References:

Deming, W.E. (1953). On a probability mechanism to attain economic balance between the resultant of nonresponse and the bias of nonresponse. Journal of the American Statistical Association., 48, 743-772.

J. Michael Dennis, Nancy A. Mathiowetz, Candice Saulsberry, Martin Frankel, K.P. Srinath, Ann-Sofi Rodén, Philip J. Smith, and Robert A. Wright. (1999). Analysis of RDD interviews by the number of call attempts: the National Immunization Survey. Paper presented at 1999 AAPOR Conference.

Ezzati-Rice, T.M., Frankel, M.R., Hoaglin, D.C., Loft, J.D., Coronado, V.G., and Wright, R.A. (2001). An alternative measure of response rate in random-digit-dialing surveys that screen for eligible subpopulations. *Journal of Economic and Social Measurement. In print.*

Smith, P.J., Battaglia, M.P., Huggins, V.J., Hoaglin, D.C., Roden, A.S., Khare, M., Ezzati-Rice, T.M., Wright, R.A. (2001). Overview of the sampling design and statistical methods used in the National Immunization Survey. American Journal of Preventive Medicine, 20:4 (Supplement 1): 17-24.

Zell, E.R, Ezzati-Rice, T.M., Battaglia, M.P., Wright R.A. (2000) National Immunization Survey: The Methodology of a Vaccination Surveillance System. Public Health Reports, Vol. 115: 65-67.

Table 1: Response Rates by Maximum Number of Call Attempts, Q4/1999-Q3/2000 National Immunization Survey

	Maximum Number of Attempts:					
	5	**8**	**10**	**12**	**15**	**Final Sample**
Resolution Rate	78.87%	82.81%	84.20%	85.33%	86.38%	88.29%
Screening Completion Rate	91.42%	93.56%	94.30%	94.80%	95.31%	96.15%
Interview Completion Rate	73.85%	79.02%	81.19%	83.08%	84.92%	93.38%

Maximum Number of Attempts:						
	5	**8**	**10**	**12**	**15**	**Final Sample**
CASRO Response Rate	53.25%	61.22%	64.47%	67.20%	69.92%	79.28
Percent of All Interviews	61.14%	74.64%	80.01%	83.49%	88.08%	100.0%

Table 2: Response Rates without Refusal Conversion, Q4/1999-Q3/2000 National Immunization Survey

Resolution Rate	88.29%
Screening Completion Rate	93.26%
Interview Completion Rate	86.08%
CASRO Response Rate	70.88%
Percent of All Interviews	78.22%

Table 3: Data on Response rates and Cost, Q4/1999-Q3/2000 National Immunization Survey

Maximum Number of Attempts (1)	Ratio of rates ($\frac{r_{t+k}}{r_t}$) (2)	Ratio of cost per complete ($\frac{c_{t+k}}{c_t}$) (3)	Adjustment factors for inflating the variance (2) x (3)
5	1.000	1.00	1.000
8	1.233	1.107	1.365
10	1.328	1.176	1.561
12	1.402	1.245	1.745
15	1.474	1.338	1.972

Table 4: Bias and Mean Squared Error of the Estimates, Q4/1999-Q3/2000 National Immunization Survey

Maximum Number of Attempts	Bias in the estimate p_t	Variance of the estimate	Adjusted variance	Mean Squared Error of the Estimate
5	0.62	0.3025	0.3025	0.6869
8	0.69	0.2401	0.3277	0.8038
10	0.72	0. .2304	0.3596	0.8780
12	0.15	0.2209	0.3856	0.4081
15	0.00	0.2116	0.4176	0.4176