ESTIMATING THE NUMBER OF ELIGIBLE RESPONDENTS FOR A TELEPHONE SURVEY OF LOW-INCIDENCE HOUSEHOLDS

James M. Ellis, Survey and Evaluation Research Laboratory
Virginia Commonwealth University, 921 W. Franklin St., P.O. Box 843065, Richmond VA 23284-3065

Key words: response rate, screening, RDD

This paper describes the application of several formulas for estimating a response rate for a list-assisted RDD telephone survey about adolescent substance abuse. The survey was conducted in Virginia with adolescents aged 12-17, after obtaining a parent or guardian's permission to interview the adolescent. The sample was drawn by Genesys, Inc. from banks of 100 consecutive phone numbers with at least one known listed phone number, and was disproportionately stratified by geographic region. The survey was conducted by the Survey and Evaluation Research Laboratory (SERL) at Virginia Commonwealth University under contract with the Virginia Department of Mental Health, Mental Retardation and Substance Abuse Services (DMHMRSAS). The survey instrument was developed by North Charles Planning and Research Group (NCRPG) of Boston, MA, and was provided by NCRPG as part of their sub-contract with SERL. Funding was provided to DMHMRSAS by the Center for Substance Abuse Treatment (CSAT) as part of the Substance Abuse Needs Assessment Project (SANAP). The author thanks Genesys for supporting this research by providing the listed status data at no cost.

Two documents were used to guide the development of the response rate calculations: "On the Definition of Response Rates" (Council of American Survey Research Organizations [CASRO], 1982) and "Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for RDD Telephone Surveys and In-Person Household Surveys" (American Association for Public Opinion Research [AAPOR], 1998). Response rate formulas were also available to SERL through its work with DMHMRSAS and the Centers for Disease Control and Prevention.

Contract issues initially drove the research on response rate calculations. SERL's contract with DMHMRSAS stipulated a minimum "CASRO response rate" of 60%, but no formula was specified. The effort to settle on a formula for the adolescent substance abuse survey may have produced information that could be more generally useful in low-incidence surveys. This effort may also become useful in thinking about response rate calculations for general population surveys, as the telephone system continues to change rapidly and lose efficiency (Piekarksi, 1999). In essence, the known listed status of sampled telephone numbers may be useful in estimating the number of working residential telephone numbers among unresolved numbers.

Because only about 10% of all adult respondents in the adolescent survey said they had an adolescent aged 12-17 living in the household, call dispositions in almost every category accumulated much faster than did completed interviews, relative to a typical general population survey of all adults. Table 1 shows the final call dispositions using some basic summary categories.

Table 1: Final Call Dispositions

DISPOSITION	N	%
Complete (comp)	3,217	2.8%
Screened elig., refused (ref)	1,057	0.9%
Screen incomplete (unk)	1,672	1.5%
Screened ineligible (no elig)	39,839	34.8%
Callback, screened elig. (cb)	175	0.2%
Callback, no screen yet (unk)	319	0.3%
Other (gone)	151	0.1%
Deaf, lang. problem (lang)	395	0.3%
Non-residential (bus)	22,723	19.8%
Busy (bz)	3,072	2.7%
Mech. answering device (mad)	3,218	2.8%
Ring-no-answer (na)	5,126	4.5%
Not in service (nis)	33,512	29.3%
TOTAL	114,476	100.00%

(NOTES: Totals for non-residential and not in service include numbers purged from the sample by Genesys [Genesys-ID service]. There were 23,552 numbers purged by Genesys from the replicates that were dialed. The actual purged sample pieces were not requested by SERL and not all of the requested sample was used. Overall, approximately 80% of the purged numbers were not in service and about 20% were businesses. For the purpose of this analysis, the 23,552 numbers purged from the replicates that were dialed were assumed to comprise 80% not in service and 20% non-residential numbers. By definition, all purged numbers were not listed numbers. The data presented here will vary slightly from the call disposition and response rate data in the final report for this project, for three reasons. (1) The sample was disproportionately stratified by geographic region. When the formula chosen for reporting the project's response rate is applied to each stratum individually, the sum of the stratum-level number of estimated eligible households is less than the result obtained when the formula is applied to the state-level numbers overall. This analysis presents data from the state-level numbers overall. The final project report used the sum of the stratum-level estimates. (2) Out of 115,550 telephone numbers used in the project, 1,074 were found more than once in the sample or data files that were merged for this analysis. Sample was requested from Genesys in five separate orders over approximately 18 months, including some from sampling strata with relatively high sampling rates. For some of these duplicates, the listed status did not stay the same. In addition, some sample loading files created at SERL inadvertently included the same sample pieces in multiple files. To simplify this analysis, all records involving a duplicated phone number were excluded. The distribution of call dispositions and listed status within the 1,074 excluded numbers was very similar to that of the group that was analyzed. (3) Cases in the deaf/language barrier category were considered to be ineligible for reporting the project response rate. For this paper they are considered to have unknown eligibility to be consistent with AAPOR definitions).

Fundamentally, all of the formulas evaluated by SERL met the same basic conceptual definition of a response rate (RR):

$$RR = \frac{\text{completions}}{\text{known eligible + estimated eligible}}$$

where

estimated eligible = e * N of unresolved numbers

The term "e" stands for a proportion(s) applied to one or more pools of unresolved numbers to estimate the number of eligible numbers among the unresolved pool(s). There is no single method for estimating e (AAPOR, 1998). Readers of both the CASRO and AAPOR documents are advised to use the best information available. Differences across the response rate calculations evaluated by SERL are essentially differences in calculating e.

Given a certain distribution of call dispositions, all possible response rate estimates must fall within two extremes marked by the assumptions that all unresolved numbers are eligible (the lower bound), or no unresolved numbers are eligible (the upper bound). In a survey of the general population, the lower bound and upper bound estimates might differ by five or 10 percentage points. In a survey with a heavy screening burden (or what we might call a low "hit rate"), the divergence between the lower and upper bound can become quite large.

For example, response rate formula RR1 (AAPOR, 1998) represents the lowest of all possible lower bounds, assuming that all numbers of unknown eligibility would turn out to be eligible. For the adolescent substance abuse survey, RR1 would be calculated as follows:

RR1 = 3,217/(4,449 + 13,953) = 17.5%

The rate is low because it does not take into account the fact that only about 10% of cooperating respondents said they had adolescents living in their homes. When the estimated number of eligible cases is adjusted to represent the observed "hit rate" of approximately 10% among resolved numbers (setting e = .1005), the lower-bound assumption that all unresolved numbers are working residential numbers yields an adjusted lower bound response rate of 55.0%:

$RR1_e$ = 3,217/[4,449 + (.1005*13,953)] = 55.0%

The upper bound is clearly:

Upper bound = 3,217/4,449 = 72.3%

The wide range – even across the adjusted lower bound and the upper bound – makes the lower and upper bounds less useful in a survey with heavy screening.

Ordinarily, the data output used by SERL to calculate the CASRO response rate does not break out refusals or callbacks by eligibility status because they are all considered eligible for a general population survey. When these dispositions were not broken out by eligibility status, the in-house spreadsheet generated the following results:

$$\frac{\text{completions}}{\text{(comp+ref+ cb+gone)+ e * (na+bz+mad+lang)}}$$

where

$$e = \frac{\text{comps+ref+ cb+gone}}{\text{comps+ref+ cb+gone+NIS+BUS+ill+no elig}}$$

If it had been conducted on all final call dispositions as shown in Table 1, SERL's initial evaluation of the response rate would have yielded the following results:

(1) Unadjusted response rate (SERL)

$$e = \frac{3,217+2,729+494+151}{3,217+2,729+494+151+33,512+22,723+0+39,839} = 0.064$$

$$RR = \frac{3,217}{3,217+2,729+494+151+0.064*(5,126+3,072+3,218+395)}$$

RR = 43.8%

This formula was not satisfactory. It did not separate refusals and other uncompleted contacts by their eligibility status. For example, all refusals were considered to be eligible households with adolescents, when a substantial portion of the refusals occurred before eligibility could be established. The formula also did not seem to adjust for the screening burden in an appropriate fashion. CASRO (1982) states that the "general rule is the number of eligible reporting units equals the number that would have been obtained if there was perfect execution." SERL interpreted "perfect execution" to mean what would have been obtained from a list of households known to have adolescents. SERL operationalized this philosophy by separating the known eligible cases from the unresolved cases, then prorating the Ns for all unresolved dispositions by the overall observed rate of households with adolescents. This did seem to adjust for the screening burden and create numbers that probably would have existed "if there was perfect execution." These numbers were then handled in the traditional SERL calculation. This "prorated response rate" was used to control the field work in the first phase of the project. This approach produced a response rate that was too high because, in prorating the call dispositions (essentially reducing them by a factor of .1005) and then processing them through the standard formula (which included an overall calculation of e=.064), some numbers in the denominator were reduced twice. But the final result – 66.1% – was more useful than the initial attempt outlined above.

As part of their technical support to the CSAT-funded substance abuse research in Virginia and other states, NCRPG developed a response rate calculation worksheet. When SERL provided an interim dataset to NCRPG for evaluation, the NCRPG response rate differed markedly from the SERL prorated response rate at that time – approximately 64%. The difference was significant because the research contract stipulated a minimum "CASRO response rate" of 60% as a condition of performance.

(2) NCRPG response rate:

$$\frac{\text{completions}}{(\text{Tot \#s used} - \text{nis} - \text{unk elig} - \text{bus} - \text{inelig}) + e * \text{unk elig}}$$

where

$$e = \frac{(\text{Total \#s used} - \text{nis} - \text{unk elig} - \text{bus} - \text{inelig})}{(\text{Total \#s used} - \text{nis} - \text{unk elig} - \text{bus})}$$

Therefore:

$$e = \frac{(114,476 - 33,512 - 13,953 - 22,723 - 39,839)}{(114,476 - 33,512 - 13,953 - 22,723)} = .1005$$

$$RR = \frac{3,217}{114,476-33,512-13,953-22,723-39,839+(0.1005*13,953)}$$

RR = 55.0%

The NCRPG formula represents another form of the lower bound calculation. It provides a more conservative estimate of the response rate by calculating e for all eligible cases within all working residential numbers and applying it to all unresolved numbers, thus assuming that all numbers in the unresolved pool are working residential numbers.

Subsequent discussion and research on response rate calculations led to the third and fourth formulas examined for use with the adolescent substance abuse study. Formula 3 was SERL's attempt to process the disposition codes as closely as possible to our interpretation of CASRO (1982). The major difference between formula 3 and the initial in-house effort was to re-group the call disposition codes more finely, so as not to assume that all callbacks and refusals were eligible households. The formula to estimate e is still based on the assumption that all unresolved numbers break down in the same proportions as all resolved numbers. A single value for e is estimated.

(3) "Single e" response rate (CDC, SERL, others)

$$\frac{\text{completions}}{\text{eligible} + (\text{unresolved}*e)}$$

where

$$e = \frac{\text{eligible}}{\text{eligible+ineligible}} = \frac{4,449}{4,449+96,074} = 0.044$$

Therefore,

$$RR = \frac{3,217}{(4,449 + (13,953*(4,449/(4,449+96,074)))}$$

RR= 63.5%

The problem with this logic seems to be that the pool of unresolved numbers includes dispositions from numbers that are quite likely to be working residential numbers (such as MADs and refusals before screening is completed), but e is based on a denominator that includes a large number of non-working and business telephone numbers.

Formula 4 shows a refinement to the estimation of e that is used by the organization conducting the adult version of the Virginia substance abuse survey. It was shared with SERL by DMHMRSAS to help in the search for an appropriate response rate formula. The key difference from formula 3 is that the pool of unresolved numbers is split into two sub-pools, and different values for e are used to estimate the eligible numbers among those sub-pools. Essentially, MADs are grouped with other unresolved numbers that were answered, and they all are assumed to be working residential numbers. Thus, e for this sub-pool (e_1) is the same as from the lower bound calculation, that is, the rate of eligible residences (those with adolescents) among all known residential numbers. All NA/BZs, on the other hand, are assumed to contain working residential numbers at a rate (e_2) equal to the rate for the pool of all resolved numbers. This "split e" refinement would seem to adjust for the potential problem noted in formula 3.

(4) "Split e" (Organization X):

$$\frac{\text{completions}}{\text{eligible} + [((\text{BZ}+\text{NA})^* e_2) + \text{all other unresolved})^* e_1]}$$

where

$$e_1 = \frac{\text{eligible}}{\text{elig.} + \text{screened inelig. HHs}}$$

$$e_1 = \frac{4{,}449}{4{,}449 + 39{,}839} = 0.1005$$

and

$$e_2 = \frac{\text{Total used–nis–bus–BZ–NA}}{\text{Total used} - \text{BZ} - \text{NA}}$$

$$e_2 = \frac{114{,}476–33{,}512–22{,}723–3{,}072–5{,}126}{114{,}476\text{-}3{,}072\text{-}5{,}126} = .471$$

Therefore,

$$\text{RR} = \frac{3{,}217}{4{,}449+[(8{,}198^*0.471)+5{,}755]^*0.1005} = 59.4\%$$

DMHMRSAS, SERL and NCRPG agreed to use formula 4 to calculate the CASRO response rate for this study.

Re-evaluation of formula 4, however, led SERL to conclude that the refinement of splitting the unresolved numbers into two sub-pools may be offset by a lack of accurate data from which to estimate the applicable rates. Specifically, for the purpose of estimating e_2 (working phone rate for NA/BZ numbers), formula 4 assumes that numbers still in NA/BZ status after numerous dials are like all numbers that were resolved. This assumption is convenient and may not significantly affect response rates for surveys with little or no screening burden. But the difference in this case is more significant and the assumption bears investigation.

First, the fact that the NA/BZ numbers stayed in those categories after 10-20 dials suggests that they are different from those that were resolved. Second, some nonworking numbers simply ring indefinitely rather than triggering a message to the caller (Groves, 1989). Piekarski et. al. (1999) indicate that the percentage of working residential numbers within the unresolved pool of numbers from list-assisted RDD surveys may hover around 10% to 25%, while Shapiro et. al. (1995) and Keeter et. al. (2000) each provide estimates of about 20%. These estimates are well below the national estimate of 41.8% working residential numbers for all active banks (Piekarski et. al., 1999). Unfortunately, these estimates may be limited in their applicability to telephone systems whose characteristics may change across geographic areas and time periods.

Perhaps the presence of listed telephone numbers could be used to indicate a difference in the NA/BZ sub-pool, as well as the approximate degree of difference. This sort of technique is used in list-assisted sampling to assess the likely density of all working numbers among banks of telephone numbers. It seems to make intuitive sense here, but the validity of this technique for measuring e is not directly tested in this paper. The empirical exercise of applying it to the adolescent substance abuse survey is described below.

Sample vendors can include a variable in the sample file that indicates whether each sampled number is found in a database of listed residential numbers. Genesys, Inc., generated the samples used for this study, and in support of this paper Genesys retroactively applied the residential indicator to all pieces of sample that had been provided to SERL for this project, according to the listed status at the time the numbers were originally sampled. This information was matched to the call disposition data for each case used in the survey. The results are shown below. (See the notes for Table 1 for caveats related to this procedure.)

Table 2: Listed Rates for Various Disposition Pools

Disposition Category	N	% Listed Within Category	
Resolved (eligible)	4,449	56.6%	32.1%
Resolved (ineligible)	96,074	30.1%	
Unresolved (MAD, other)	5,755	46.6%	46.6%
Unresolved (NA/BZ only)	8,198	12.1%	12.1%

It seems clear that the sub-pool of NA/BZ numbers is significantly different from the other pools, at least in terms of the proportion of listed numbers. It seems that the NA/BZ sub-pool would contain the same proportion of working residential numbers as found in the resolved pool only if unlisted numbers were far more likely to end up as NA/BZ as opposed to MAD/other, and their absolute number was similar to or greater than the absolute number of listed numbers. This scenario seems unlikely, but it is untested in this paper.

While admittedly convenient, it does not seem unreasonable to assume that the proportions of listed residential numbers to total residential numbers will stay relatively constant across different sub-pools of numbers. If so, proportions of listed numbers can be compared across sub-pools as proxies for the proportions of all residential working numbers, that is:

$$\frac{(LR)}{(WR)} = \frac{(LB)}{(WB)}$$

where

LR = % listed in resolved pool
WR = % working residential in resolved pool
LB = % listed in NA/BZ pool
WB = % working residential in NA/BZ pool

Therefore, WB essentially represents e_2. This can be expressed as:

$$e_2 = (LB/LR) * (WR)$$

Using the data in Table 2 and the working residential rate among resolved numbers from formula 4, we can estimate the following:

$$e_2 = (.121/.321) * .471 = .1775, \text{ or } 17.75\%$$

Compare this value to.471 (WR), which was used alone in formula 4 to estimate the number of working residential numbers in the NA/BZ pool. This value compares more favorably to the values of .209, which may be derived from Shapiro et. al. (1995), .20 from Keeter et. al. (2000), and the range of .10 to .25 from Piekarski et. al. (1999).

The formula suggested by this analysis is:

$$\frac{\text{completions}}{\text{eligible}+[((BZ+NA)*e_2)+\text{all other unresolved})*e_1]}$$

where, as in formula 4,

$$e_1 = \frac{\text{eligible}}{\text{eligible + ineligible HHs}}$$

$$e_1 = \frac{4{,}449}{4{,}449 + 39{,}839} = 0.1005$$

but now

$$e_2 = (LB/LR) * (WR) = (.121/.321) * .471 = .1775$$

Therefore,

$$RR = \frac{3{,}217}{4{,}449 + [(8{,}198*0.1775) + 5{,}755]*0.1005}$$

RR = 62.2%

Note that this rate is close to the one obtained in formula 3. It may be too much to suggest that formulas 3 and 5 validate one another in this fashion. Certainly, the critical assumption $e_2 = (LB/LR) * (WR)$ should be subjected to validation research.

Assuming that the list-assisted approach is valid, however, this interesting empirical result raises a question – why use the list-assisted version of the formula if a nearly identical result can be obtained using formula 3? One possible advantage is that the list-assisted version may be sensitive to systematic changes in how call dispositions might be distributed across the NA/BZ pool versus the MAD/other pool, holding all other categories constant. Using the adolescent substance abuse survey data, we can examine what would be the upper and lower bounds of the various formulas that have been discussed so far, according to how the total number of NA/BZ and MAD/other calls could have fallen.

Clearly, the minimum number of calls in any one of the two sub-pools of interest could be zero if all of the calls across the two sub-pools fell into just one sub-pool, as shown in Table 3. Whether the extreme value occurred in the NA/BZ or MAD/other category, formula 3 and APPOR RR1 would yield unvarying response rate estimates.

The only difference between formula 4 and 5 is e_2, the rate of estimated eligible households applied to the NA/BZ pool. This fact is demonstrated in Table 3 when the projected number of cases in the NA/BZ pool is zero. At this point, the different values for e_2 disappear, having zero cases on which to operate, and the formulas yield identical estimates.

Table 3: Response Rates When NA/BZ and MAD/Other Results Vary to Extremes

Formula	NA/ BZ	MAD/ Other	Est. eligible	Resp. rate
#3: Single e (CDC, SERL, others)	13,953	0	**617.5**	**63.7%**
	8,198	**5,755**		
	0	13,953		
#4: Split e (Contractor X)	13,953	0	660.0	63.0%
	8,198	**5,755**	**965.9**	**59.4%**
	0	13,953	1,401.6	55.0%
#5: List-assisted split e	13,953	0	242.4	68.6%
	8,198	**5,755**	**724.3**	**62.2%**
	0	13,953	1,401.6	55.0%
Lowest bound (AAPOR RR1)	13,953	0	**13,953**	**17.5%**
	8,198	**5,755**		
	0	13,953		

(NOTE: For all formulas, the number of known eligible cases is 4,449. Estimated eligible cases shown here may vary slightly from calculations using data in the paper due to rounding of terms.)

The level of precision sought by some of these formulas may not be needed for many surveys, and the extra cost of obtaining the listing information and re-doing standard in-house calculations may not be justified. As the telephone system becomes less efficient, however, call dispositions in general population surveys may start to look more like dispositions from surveys with heavy screening burdens. Changes in the telephone system itself, as well as how people use the technology, may produce systematic shifts in how call dispositions are distributed across sub-pools, which would otherwise be missed by some response rate formulas.

In addition, survey research is constantly advancing into more extreme territory, so more survey researchers may find themselves looking for a new response rate calculation. As the consumers of survey research become more sophisticated, more emphasis may be placed on response rates as one measure of the aggressiveness or sophistication of the data collection effort. Finally, the growing trends toward multi-modal surveys and suggestions for the use of differential levels of effort across sub-pools of the sample (Groves, 1989) may place more pressure on the identification of appropriate formulas to assess the survey response rate. For these reasons, the list-assisted approach to calculating response rates may be useful at times.

References

American Association for Public Opinion Research. 1998. "Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for RDD Telephone Surveys and In-Person Household Surveys." Ann Arbor, MI. AAPOR.

Council of American Survey Research Organizations. 1982. "On the Definition of Response Rates," L. Frankel, chairman, Completion Rates Task Force. Fort Jefferson, NY. CASRO.

Groves, R. M. 1989. *Survey Errors and Survey Costs*, pp. 133-183. New York: Wiley.

Keeter, S.; Carolyn Miller; Andrew Kohut; Robert M. Groves; and Stanley Presser. 2000. "Consequences of Reducing Nonresponse in a National Telephone Survey." In press.

Piekarski, L.; G. Kaplan; and J. Prestegaard. 1999. "Telephony and Telephone Sampling: The Dynamics of Change." Paper presented at the annual meeting of the American Association for Public Opinion Research, St. Petersburg, FL.

Shapiro, Gary; Michael P. Battaglia; Donald P. Camburn; James T. Massey; and Linda I. Tompkins. 1995. "Calling Local Telephone Company Business Offices to Determine the Residential Status of a Wide Class of Unresolved Telephone Numbers in a Random-Digit-Dialing Sample." In *American Statistical Association 1995 Proceedings of the Section of Survey Research Methods*, pp. 975-980.