

## *Calling Effort and Nonresponse for Telephone Panel Surveys*

Roberta L. Sangster

Bureau of Labor Statistics, Washington, DC, United States of America

Earlier Version Presented for the 13<sup>th</sup> International Workshop on Household Survey Nonresponse

Copenhagen Aug. 29-31, 2002; Revised July 10, 2003

Over several years, I have been working with issues related to RDD (Random Digit Dialed) generated telephone survey panels and attrition rate calculations. For this study, I examine how nonworking numbers, noncontacts (basically answering machines, scheduled callbacks, and ring-no-answer numbers impact the number of completed interviews across time and also the amount of effort involved to contact the telephone numbers based on the number of call attempts. The analysis presented in this paper is only descriptive, no statistical tests are used. Opinions and errors are the sole responsibility of the author and do not represent opinions of the Bureau of Labor Statistics.

### **The Study**

The data come from the Telephone Point of Purchase Survey (TPOPS). The Bureau of Labor Statistics (BLS) conducts the TPOPS to create the sample frame used in pricing goods and services for the Consumer Price Index. The TPOPS collects the name and address of the outlets (e.g., grocery stores, theatres, physicians, and mail-order catalogs) where households purchase various types of goods and services. The sample for each panel is selected via RDD.<sup>1</sup> Each quarter about 45% of the sample is new RDD sample and the rest is returning sample ( $n \approx 32,000$  cases per quarter). Once a household is selected and interviewed, it may be interviewed up to three more times over the next nine to 12 month period. The number of targeted completed cases each quarter is approximately 14,000 urban households. The sample is a stratified national RDD sample of urban units drawn in the second quarter of 2001 and retired in the first quarter of 2002.<sup>2</sup> Response rates for the four quarters were 57%, 55%, 52%, and 51% and are based on AAPOR's (2000) standard definitions and calculation rate (RR4), using an estimate of .27 for the unknowns.

### Final call distribution

*Interviews* include completed interviews and partially completed interviews. *Refusals* include verbal refusals and break-offs (hang-ups) at the household and respondent level. *Noncontacts* include cases where the telephone number is confirmed as eligible but a respondent is never contacted or never available (e.g., broken appointments (callbacks or CBs), answering machine message that indicates that it is an eligible case) and other eligible cases that cannot be completed (e.g.,

hard of hearing, language barrier, absence, and hospitalized). *Unknowns* are situations where the number is always busy, a ring-no-answer, or call-blocking and call-screening systems are encountered. For this study the *ring-no-answers* will be removed from the unknown category and analyzed separately. For TPOPS, the ring-no-answers are sub-sampled in the first quarter and half are returned to sample for the subsequent quarters. This is done to improve efficiency; the cases that remain are weighted in the actual data used for TPOPS. The data used for this study are unweighted. One of the reasons for conducting this study was to evaluate the decision to sub-sample these cases.

*Not eligible* are cases so designated by the requirements of the study (e.g., telephone numbers outside the sampled geographic area, military and farm households, and non-residential numbers). *Nonworking numbers* are also ineligible but are kept separate for the purpose of this study. Nonworking numbers are removed in the first quarter, but those numbers that become nonworking after the first quarter are not removed from the sample because they are considered eligible sample units. Other cases that become ineligible in subsequent quarters also remain in sample.

### Calling Rules

Calling rules are important because they limit the number of call attempt based on predefined rules. These rules can directly impact the distribution of the final outcomes. For example, stopping after 2 refusals truncates the case to be a refusal, whereas making a third attempt at refusal conversion may have resulted in a completed interview.<sup>3</sup>

TPOPS's first calling rule deals with the ring-no-answer problem (i.e., the cases that are never contacted). After 12 consecutive ring-no-answers, one half of these cases are removed from sample and the other half of the cases receives no more calls until the next quarter. Also included in the 12 call rule are answering machines (AM). After 12 consecutive calls to AMs, calling stops until the next quarter. AMs can be coded out as a noncontact if the message indicates it is a household (e.g., *hi this is Bill and Sue please leave a message*) or as an unknown if the message is vague (e.g., *you have reached*

<sup>1</sup> Genesys® sample.  
<sup>2</sup> "Urban" includes A, B, and C size PSUs.

<sup>3</sup> The final call outcome code is the last outcome code associated with the case. Unfortunately, this does not tell us much about what happened with the case during the calling period.

202-999-9999, please leave a message). After 30 call attempts, the second rule requires that calling stops on all other noncontact cases (mostly composed of CBs and AMs). Appendix A shows the entire list and definitions of variables used within the call counter.

There are also calling rules for refusals. Calls stop after two refusals within one quarter. Refusals are removed from sample after two refusals in two consecutive quarters; hard refusals are removed immediately. It is important to note here that the tables presented will not include the refusals deleted from the sample based on the calling rules.<sup>4</sup> This is because the interest in this study is the viable sample that is available to call for each quarter.

The intent of this study is to assess the impact of nonworking numbers, noncontacts, and ring-no-answer numbers on the number of completed interviews across time and also the amount of effort involved to contact the telephone numbers based on the number of call attempts. It will also examine the viability of subsampling ring-no-answer numbers to determine if this provides a way to reduce coverage bias due to the loss of these cases.

## Results

The first set of tables examines the final call distributions across time in sample. Of interest in Table 1 is the high loss of sample due to nonworking numbers. The loss of nearly 30% of the sample is costly when you also consider that another 17.6% of the sample is lost due to ineligibility. Only 26.2% of the sample ends up being a completed interview, illustrating the challenges of dealing with an RDD generated sample frame.

Noncontact is relatively low at 3.8% of the sample (mainly CBs and AMs, the rest consist of things such as absence, ill, hard of hearing, and language barriers). The *ring-no-answer* are about 10% of the sample (n=1461). The ring-no-answer were sub-sampled in quarter one and returned to sample for subsequent calling (n=708).

**Table 1: Quarter One Final Call Distribution**

Final Outcome	Percent	Frequency
Interview	26.2	3789
Refusal	13.3	1930
Noncontact	3.8	543
Unknown	0.1	15
Ring-no-answer	10.1	1461
Not Eligible	17.6	2544
Nonworking	28.9	4179
Total	100%	14461

<sup>4</sup> In Quarter one, 333 refusals are removed; in Quarter two, 846, and in Quarter three, 218.

Table 2 illustrates how efficiency is improved by the removal of the ineligible cases (n=6723), one-half of the ring-no-answers (n=753), and the hard refusals from the sample (n=333). About 54% percent of the RDD sample is removed after the first quarter of calling (n=7809). Completed interviews are now about 50% of the sample and continue around 50% for the next two quarters. Nonworking numbers (7.3%) and not eligible cases (2.8%) now represent a small part of the sample. The *ring-no-answer* cases again represent about 10% of the sample; whereas noncontacts increase to about 10%.

**Table 2: Quarter Two Final Call Distributions**

Final Outcome	Percent	Frequency
Interview	49.9	3319
Refusal	19.5	1298
Noncontact	10.2	676
Unknown	0.6	40
Ring-no-answer	9.7	643
Not Eligible	2.8	189
Nonworking	7.3	487
Total	100%	6652

Table 3 shows where the calling rule for refusals reduces the sample size by 845 cases in quarter three. Refusals now represent about 12% of the sample in Quarter 2 and Quarter 4. The number of nonworking numbers increases to about 12% of the sample and remains about the same in quarter four (See Table 4). Noncontacts stay about the same as well. However, it is not safe to assume that the case outcomes are stable.

**Table 3: Quarter Three Final Call Distribution**

Final Outcome	Percent	Frequency
Interview	51.7	3004
Refusal	12.5	724
Noncontact	9.7	564
Unknown	2.0	114
Ring-no-answer	9.1	526
Not Eligible	3.2	187
Nonworking	11.8	687
Total	100%	5806

**Table 4: Quarter Four Final Call Distribution**

Final Outcome	Percent	Frequency
Interview	51.4	2870
Refusal	11.5	641
Noncontact	9.3	519
Unknown	0.1	4
Ring-no-answer	10.7	597
Not Eligible	4.2	236
Nonworking	12.9	721
Total	100%	5588

### Quarter 2 Nonworking Numbers and Noncontacts

Since nonworking numbers are removed in the first quarter, the comparison uses quarter two data to compare nonworking numbers to noncontacts across time. Of the 487 cases of nonworking numbers in quarter two, more than half of this number will remain nonworking for the subsequent two quarters (69.2% and 53%). For the noncontacts in quarter two, 32.8% and 31.7% remain noncontacts in quarters three and four respectively (See Table 5 and Table 6). About 13% of the nonworking numbers in quarter two are completed interviews in quarter three and 18.7% in quarter four. Completed interviews for quarter two noncontacts are 24.6% and 22.3%, respectively.

**Table 5: Quarter Two Nonworking Numbers by Quarter 3 and Quarter 4 Final Call Distribution**

Quarter 2 Final Outcome	Quarter 3		Quarter 4	
	Percent	Freq.	Percent	Freq.
Interview	12.9	63	18.7	91
Refusal	6.2	30	10.3	50
Noncontact	4.7	23	6.6	32
Unknown	0.8	4	0.0	0
Ring no ans.	3.3	16	5.0	25
Not Eligible	2.9	14	6.4	31
Nonworking	69.2	337	53.0	258
Total	100%	487	100%	487

**Table 6: Quarter Two Noncontact by Quarter 3 and Quarter 4 Final Call Distribution**

Quarter 2 Final Outcome	Quarter 3		Quarter 4	
	Percent	Freq.	Percent	Freq.
Interview	24.6	166	22.3	151
Refusal	23.5	159	24.1	163
Noncontact	32.8	222	31.7	214
Unknown	0.7	5	0.0	0
Ring no ans.	4.1	28	4.3	29
Not Eligible	5.7	35	5.8	39
Nonworking	9.0	61	11.8	80
Total	100%	676	100%	676

### Callbacks and Answering Machines

To further understand the noncontacts the data were analyzed for quarter two CBs and AMs across quarters three and four (n=676). These two outcomes constitute 80% of the noncontacts (40% are CBs and 40% are AMs). The rest of the noncontacts include 12% who are absent or ill for the duration of the interviewing period, 5% are due to language barriers, and 3% for other reasons.

### Callbacks

There are 272 CB cases for quarter two (See Table 7). Of these cases in Quarters three and four, 31.3% and 28.3%, respectively, become completed interviews, 20.6% and 22.4% are refusals, 19.5% and 16.0% remain CBs, and 8.8% and 23.8% become AMs as the final call distribution. Answering machines appear to be used as a screening device by quarter four. The majority of the cases remain eligible, 87.5% in quarter three, and 83.4% in quarter four.

**Table 7: Quarter Two Callbacks by Quarter 3 and Quarter 4 Final Call Distribution**

Quarter 2 Final Outcome	Quarter 3		Quarter 4	
	Percent	Freq.	Percent	Freq.
Interview	31.3	85	28.3	77
Refusal	20.6	56	22.4	61
Noncontact	2.9	8	2.9	8
Callback	19.5	53	16.5	45
Ans. Mach	8.8	24	23.8	64
Unknown	0.3	1	0.0	0
Ring no ans.	4.0	11	5.2	14
Not Eligible	5.9	16	5.2	14
Nonworking	6.6	18	11.4	31
Total	100%	272	100%	272

### Answering Machines

Fewer AMs cases result in completed interviews than for CBs. Table 8 shows that of the 269 cases in quarter two that were AMs, 19.3% and 16.7% resulted in completed interviews for quarters three and four respectively. Refusals are somewhat more likely with 22.7% for quarter three and 23.1% for quarter four. About one quarter remain as AMs, and 6.7% become CBs. Again, the majority of the cases remain eligible, 82.5% in quarter three and 77.7% in quarter four.

**Table 8: Quarter Two Answering Machines by Quarter 3 and Quarter 4 Final Call Distribution**

Quarter 2 Final Outcome	Quarter 3		Quarter 4	
	Percent	Freq.	Percent	Freq.
Interview	19.3	52	16.7	45
Refusal	22.7	61	23.1	62
Noncontact	2.2	6	1.9	5
Callback	6.7	18	7.4	20
Ans. Mach	24.5	66	23.8	64
Unknown	1.5	4	0.0	0
Ring no ans.	5.6	15	4.9	13
Not Eligible	5.2	14	8.2	22
Nonworking	12.3	33	14.1	38
Total	100%	269	100%	269

### Ring-no-answer Sub-sample

Finally, what about the decision to sub-sample the ring-no-answer numbers (RNA) during the first quarter (n=708)? Did this decision help to reduce potential bias? As shown in the next three tables, the majority of the RNA cases results in a subsequent code of RNA (78.2%, 62.3%, and 69.3% respectively). The lower percent in quarter three is largely accounted for by the large increase in the number of unknowns (12.1%). The rest of the cases are represented by nonworking numbers and other not eligible cases.

**Table 9: Quarter One Sub-sampled *Ring-no-answer* By Quarter Two Final Call Distribution**

Final Outcome	Percent	Frequency
Interview	1.6	11
Refusal	2.1	15
Noncontact	2.5	18
Unknown	0.0	0
Ring no ans.	78.2	554
Not Eligible	7.8	55
Nonworking	7.8	55
Total	100%	708

**Table 10: Quarter One Sub-sampled *Ring-no-answer* By Quarter Three Final Call Distribution**

Final Outcome	Percent	Frequency
Interview	1.6	11
Refusal	1.7	12
Noncontact	2.4	17
Unknown	12.1	86
Ring no ans.	62.3	441
Not Eligible	7.2	51
Nonworking	12.7	90
Total	100%	708

**Table 11: Quarter One Sub-sampled *Ring-no-answer* By Quarter Four Final Call Distribution**

Final Outcome	Percent	Frequency
Interview	1.8	13
Refusal	1.5	15
Noncontact	2.6	18
Unknown	0.28	2
Ring no ans.	69.3	488
Not Eligible	8.95	63
Nonworking	15.5	109
Total	100%	708

Fifty percent of the sub-sampled RNA cases remain RNA across all four quarters. Five percent become nonworking across the next three quarters and four percent are not eligible for the next three quarters. Of the completed interviews, only two cases are completed interviews all three quarters, the majority of the rest of cases tend to bounce between the nonworking, unknown, and not eligible outcomes.

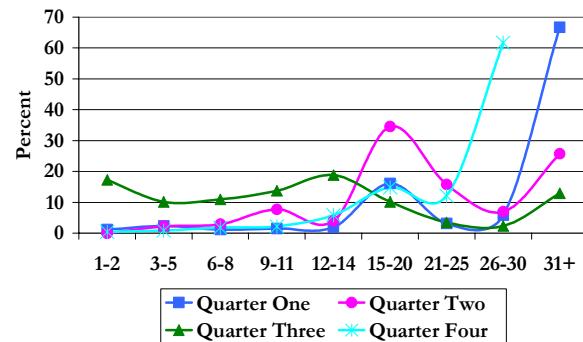
If we combine Quarters 2-4, the combined sample is 18,046 telephone numbers. Of these 18,046 cases, interviews (n=35), refusals (n=42), and noncontacts (n=53) result in a total net of 130 instances where an eligible household might have been contacted across three quarters of interviewing (130/18,046 or 0.0072). In contrast, the cases that remain or become RNA (n=1483) unknown (n=88) or ineligible (n=169) represent 1.5% of the combined sample. It is questionable whether continuing to call RNAs reduces nonresponse bias.

### Calling Effort

So what is the effort involved with implementing the calling rules? The next set of figures shows the number of call attempts broken into groupings that account for the calling rules (12-14 calls represent the 12 call attempt rule; +31 or more represents the 30 call rule).

Callbacks and AMs constitute the majority of the noncontacts. How long is it worthwhile to continue to call AMs and make CBs? Examining Figure 1, it appears that in quarter one and quarter four the thirty call rule pushes the vast majority of CBs into the thirty call limit. However, in the second and third quarters many cases end in a CB before the calling limit is reached. It is likely the refusal calling rule of two refusals in two consecutive quarters (refusal conversion) limits the number of CBs that can be completed.

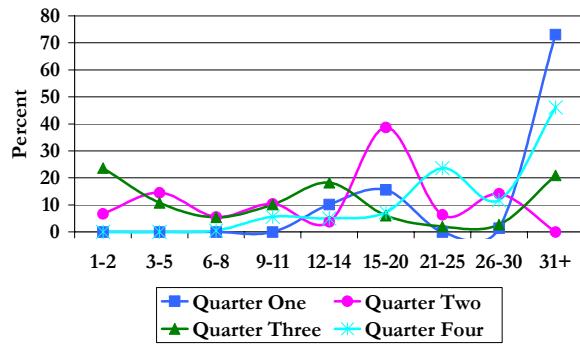
Figure 1. Callbacks by Call Attempts



However, the 12 consecutive AMs calling rule does not appear to work very well. Over 70% of the cases in Quarter 1 and over 45% of the cases in Quarter 4 end with the final outcome as an AMs (See Figure 2). Probably busy signals and RNAs are affecting the 12

attempt calling rule since it must be 12 consecutive AMs. A better AM calling rule might be some combination of these outcomes (RNA, CB, and AM), such as a total of five AMs among 10 attempts with no contact with the household. Again, it appears that making attempts at refusal conversion affect the number of call attempts made to other cases in quarters 2 and 3.

**Figure 2. Answering Machines by Call Attempts**



Since the cases that become ineligible after the first quarter stay in sample, it is worth noting that most of the ineligible cases are resolved with 3 to 5 call attempts, similar to interview or refusal distributions. There is also shifting between eligible to ineligible status that is probably due to “soft refusals” or misunderstanding the screening question(s) since the respondent for the household can change for TPOPS. Further investigation of the question wording is probably warranted.

## Conclusions

Overall, the findings indicate that it is probably not worth the cost to continue to sub-sample the RNA numbers in quarter one. It is more worthwhile to consider using an estimate of the eligible units within the RNAs outcomes and not continue calling them after the first quarter. This would be more cost effective and have little effect on bias.

In contrast, it might be worthwhile trying a sub-sampling of the nonworking numbers in quarter one as a test. Since all of the cases were removed in the first quarter it is difficult to draw any firm conclusions, but based on the small sample in quarter two ( $n=487$  nonworking numbers) it indicates that maybe 13 to 15% could result in completed interviews. Nonworking numbers also represent nearly 30% of the sample loss in the first quarter, in comparison to only 10% RNA. Since the majority of nonworking numbers take one call to determine if the number is still nonworking, compared to 12 calls for the RNA group, taking a sub-sample of this

group might be more economical and have a greater likelihood of reducing bias.<sup>5</sup>

Making 30 or more call attempts adds a good deal of cost to the study with only a small portion of the effort adding to the usable data. For example, if the second calling rule was limited to 20 call attempts for this panel it would mean 57,267 fewer call attempts (Quarter 1  $n=19,357$ , Quarter 2  $n=9,220$ , Quarter 3  $n=12,377$ , Quarter 4  $n=16,313$  call attempts). For completed interviews, this reduction in calls would result in a loss of 3.11%, 1.99%, 2.03%, and 2.51% across the four quarters. Thus, making fewer call attempts on cases less likely to yield a completed interview might be a better solution. It would be worthwhile to consider other ways to optimize the calling rules. For example, for CBs and AMs would it be worthwhile to create calling protocols based on contact probability scores or use a priority score approach similar to what are commonly used for cold call cases (See Weeks 1988 for a discussion). It would be worthwhile to examine call history data more closely to come up with a better calling protocol for these cases.

This study did not examine the impact of making refusal conversion calls directly, but indirectly they appear to impact what happens to the rest of the calling effort for the sample. Brick et al. (2003:4) examined refusal conversion between two rounds of an RDD panel survey (Round 1 in 1997, Round 3 in 2002) and discovered that it took 2.6 more call attempts to complete a refusal in four years. In 2002 it averaged 3.8 call attempts to complete an interview, a successful refusal conversions averaged 10.0 call attempts.

Finally, we need to start thinking about the impact we are having on future respondents. For example, for TPOPS, we make it very hard to refuse. The survey procedures require two refusals in two consecutive quarters before removal or at least 30 call attempts. Recently a respondent actually recorded a message on his AMs indicating that he did not want to receive any more calls regarding the survey. When does repeated calling become harassment?

<sup>5</sup>Kathy Frankovic reported at a recent conference that CBS recalls nonworking numbers and this yields about 1% of their completed interviews. CBS calling periods (2 days to 5 days) tend to be very short compared to TPOPS (8 weeks of calling).

## Appendix A: Call Counter

Call Counter	Maximum Limit on Call Attempts	Rule which adds "1" to counter
Refusal Counter (RC)	2	outcome = 30, 31, 33, or 34
Ring-no-answer (NCC) Counter	12	outcome = 80, 81, 82, 83, 84, 86 and NO previous outcome = 30, 31, 33, 34, 40-45, 47, 48, 49, 79, or 89
Total Contact Counter (TCC)	30	outcome = 30 - 89 and a previous outcome = 30, 31, 33, 34, 40-45, 47, 48, 49, 79, or 89

### 30-34 Hostile Breakoffs and Refusals

- 30 Hostile breakoff by respondent following interview progress\* on this call.
- 31 Refusal by respondent at or after >Intro\_1st< without interview progress\* on this call
- 33 Refusal prior to >Intro\_1st<
- 34 Immediate hang-up

\* Interview progress is defined as getting past the front and into the middle.

### 40-41 Callback Needed and Acceptable

- 40 Partial interview obtained: callback needed and acceptable after interview progress on this call.
- 41 Callback needed and acceptable without interview progress on this call.

### 42-44 Delays in Reaching Household Respondent; No Interview Progress

- 42 All respondents temporarily absent or away
- 43 All respondents temporarily absent or in hospital

### 45-49 Special Situations

- 45 Residential/special place undetermined, contact information obtained
- 47 Language barrier or problem: refer to supervisor or language specialist
- 48 Respondent is deaf
- 49 Reached answering service or AMs identified for the telephone number

### 70-79 Unresolved and Interim RDD Contacts

- 75 Unconfirmed non-working number (unconfirmed because information obtained from a recording)
- 79 Confirmed residential from other source (i.e., AMs leads you to believe it's residential)

### 80-89 Calls Without Contacts to Sample Telephone Number

- 80 Ring-no-answer
- 81 Normal busy or circuits busy/FAX
- 82 Fast or WATS busy
- 83 Number could not be completed as dialed
- 84 No signal, funny signal
- 85 Bad connection
- 86 Number temporarily not in service
- 87 Wrong number dialed or reached
- 88 Possible wrong number. Person answering would not confirm sample number
- 89 Answering machine--unknown if reached sample number