Evaluating the Survival Analysis Method for Estimating Residency Rates

Barbara Lepidus Carlson and Daniel Kasprzyk
Mathematica Policy Research, P.O. Box 2393, Princeton, New Jersey 08543-2393
Joint Statistical Meetings, Toronto, August 2004

Key words: RDD, response rate, residency, Community Tracking Study, CASRO, survival analysis

Abstract. To calculate a response rate for a survey, each sample member must first be classified according to its eligibility status. For telephone surveys with random-digit-dial samples, we try to determine for each sampled telephone number whether it is in service and belongs to a residence. The residency status of a certain portion of the sample, however, is never resolved even after making many call attempts. Several methods have been proposed to estimate residency among unresolved telephone numbers. The most conservative of these assumes the same residency rate among the unresolved cases as was found for the resolved cases. This inevitably overestimates the number of eligible telephone numbers and unduly reduces the response rate. A newer approach proposed in the literature tries to estimate a residency rate by modeling the "time to resolution" using survival analysis. The numbers designated as residential, non-residential, and unresolved at each attempt are incorporated into the model. In this paper, we apply a variation of this methodology to the HSC Community Tracking Study Household Survey and report on how it compared to the more conservative method of estimating residency.

Introduction and Background. In recent years, survey directors have struggled to obtain high response rates. This has been true of most sample surveys, but telephone surveys in particular have observed increasing difficulties in obtaining high response rates. This is especially true of random-digit-dial (RDD) surveys where the telephone is used to sample households and persons. Difficulty in obtaining high response is likely the result of a number of factors, including a societal change in the willingness to participate in voluntary surveys and improved telephone technology that permits households to screen their calls. As a result, RDD surveys have seen an increase in refusals and an increase in overall nonresponse because the data collection organization is unable to make contact with anyone at the dialed telephone number. As a consequence, in the latter situation, the survey practitioner does not know whether the telephone number is a working residential number and thereby eligible for the survey.

In theory, the response rate for a survey is easily defined as the number of responding eligible sampling units divided by the total number of eligible sampling units. In practice, the response rate calculation is more difficult, because the total number of eligible sample units must be estimated. The response rate calculation thus depends on estimating the proportion of sample cases of unknown eligibility that are in fact eligible. In an RDD household survey, this involves an estimate of the proportion of telephone numbers called that are residential households. To be more precise, if we let:

A=eligible household completes
B=eligible household noncompletes
C=ineligible households
D=households with undetermined eligibility
E=non-households (business, cell, non-working number)
F=unresolved whether household

The response rate is defined to be:

$$\frac{A}{A + B + (e_1 \cdot D) + (e_1 \cdot e_2 \cdot F)}$$

where e_1=survey eligibility rate among households, and e_2=residency rate among unresolved telephone numbers. (The ineligible categories, C and E, are often used to help determine e_1 and e_2.)

Many view the response rate as an indicator of the quality of the field operations, and therefore the quality of the survey. Thus, of late, researchers have been more cognizant of the estimation of e_2, sparked largely by their desire to obtain better estimates of the response rate. This paper concerns the estimation of the household residency rate by two methods: one a commonly accepted way of estimating residency rates and the other a more recent method, referred to as the "survival analysis" method, proposed by Brick, Montaquila, and Scheuren (2002).

The paper by Brick et al. discusses methods currently used for determining residency rates. One method, labeled the CASRO method (1983), assumes the residency rate for the unresolved numbers is the same as the rate for the resolved numbers. This rate is available from the survey and is easily calculated; it, however, often results in an overestimate of the true residency rate and, therefore, a somewhat conservative response rate, especially when there is a

limited number of call attempts made. Many argue that, although easy to implement, the estimate cannot be accurate because the assumption on which it is based is simply not correct.

A second method, called the "business office method," develops the estimate of e_2 by drawing a subsample of unresolved numbers and contacting telephone business offices to determine the proportion of numbers that are residential. The advantage is that the data collection organization collects additional information to obtain a more accurate estimate of the residency rate or uses data from studies that have gone to business offices. The method, however, takes time and resources, and often the business office does not comply with the request. And using data from other studies can be problematic, mostly because of the time lag (the latest published estimates are from the mid-1990s) and the rapidly changing telephone services market. Furthermore, establishing residential status often occurs after the survey has been conducted and the collection itself has unknown measurement error properties.

A third method is the "survival analysis method," as proposed by Brick et al. Used in clinical research, survival analysis is a biostatistical technique used to measure the time between entry into a study and a subsequent event (for example, time until death) for different groups. Brick et al. adapted the methodology so that it could be used to obtain a better estimate of residency status, taking advantage of the additional information found in the number of call attempts until resolution (as residential or non-residential), rather than relying only on the final outcome of each case. The method models the time until resolution ("death") for each telephone number in the sample. Being resolved as residential or non-residential is comparable to two different "causes of death" in the biostatistical setting. And each call attempt is a "trial." Unresolved telephone numbers are the equivalent of censored observations in a clinical trial.

The methodology produces survival functions (Kaplan-Meier estimates) for each resolution type (residential, non-residential). These are combined as follows to come up with an estimate of the overall residency rate:

$$\hat{R}_\infty = \frac{\hat{S}_{res}(0)}{\hat{S}_{res}(0) + \hat{S}_{nonres}(0)}$$

where $\hat{S}_{res}(0)$ is the survival function for resolving as residential and $\hat{S}_{nonres}(0)$ is the survival function for resolving as nonresidential.

While the CASRO method assumes the same residency rate overall and among the unresolved telephone numbers, the survival method uses the overall residency rate and the resolved cases to come up with a distinct residency rate for the unresolved cases:

$$\hat{R}_{un} = \frac{(\hat{R}_\infty \cdot n_{tot}) - n_{res}}{n_{un}}$$

where n_{tot} is the total sample size and n_{res} is the number of sample cases resolved as residential. The numerator of this equation represents an estimate of the number of unresolved cases that are residential (estimate of total residential minus known residential).

Brick et al. applied the survival analysis method to two large scale RDD surveys, the 1999 National Household Education Survey and the National Survey of America's Families, and described the limitations and difficulties with the method. Despite some issues in its application, the use of survey–specific data is very compelling in the estimation of the survey's residency rate. This paper adds to the knowledge of the method by implementing it on two rounds of the household component of the Community Tracking Study (CTS).

Methods. We used data from the Community Tracking Study Household Survey (CTS) to test the survival method and compare it to the CASRO method. The CTS documents how the U.S. health care system is changing over time, and tracks how those changes affect people. The survey focuses on local health care markets as well as the nation as a whole.[1] There have been four rounds of CTS, roughly every two years since 1996. The CTS has a clustered design where the primary sampling units

---

[1]The CTS is funded by the Robert Wood Johnson Foundation, and sponsored and designed by the Center for Studying Health System Change. The survey samples and weights were designed and implemented by Mathematica Policy Research, who also collected the data. Social and Scientific Systems did the final data processing and file production. Technical publications can be found at the Center's website: www.hschange.org.

were 60 nationally representative sites. It has a list-assisted RDD sample. Starting in round 2, the sample has been a mix of "re-interview," other "overlap," and new sample. The "re-interview" cases are those where the telephone number resulted in a complete in the prior round. (We did not follow households from round to round if they were not at the same telephone number.) The other "overlap" cases were telephone numbers released for interviewing the prior round, but did not result in completed interviews. The new sample contained telephone numbers not part of the prior round sample (both those that could have been sampled in the prior round and those that did not exist at that time).

The round 4 field period was approximately 12 months (February 2003 through February 2004); however, we released sample in random batches over the course of that year. About three-quarters of the cases were actively pursued for 3 months or less, and 90 percent were completed within six months of release. We note that it is possible for some telephone numbers to change status (for example, from a working number to a disconnected number) over the course of several months.

When applying the CASRO and survival methods to the CTS data, the first step was the classification of final status codes as (1) residential, (2) non-residential/non-working, and (3) unknown. To use the survival method, we also had to classify interim status codes (such as "call-back," "circuit problems," "needs Spanish-speaking interviewer"). Codes classified as residential in round 4 included (among others): completes, hang-ups, refusals, language barrier, and call-backs. Codes classified as non-residential were: computer/fax/modem, non-working number, cell phone/pager, and non-residence. Codes classified as unknown were: maximum calls-probable residence, ring/no answer, answering machine or service, busy, circuit problems, temporarily not in service, funny signal/no ring.

Next, the survival method required counting the number of call attempts until a case was resolved; for example, suppose a case had the following outcomes over five call attempts:

1. Busy
2. No answer
3. Firm call-back
4. No answer
5. Complete.

Although there were five call attempts, we established the case as residential at call attempt 3. A case with 20 call attempts resulting in only "no answer" or "busy" outcomes would be classified as unresolved after 20 attempts.

The survival method requires a distribution of call attempts in order to estimate the survival functions. Mathematica's survey operations center has a set of standard maximum number of call attempts for different status codes, set at a very high number, which causes the distribution of call attempts to cluster at that maximum. To ensure a more complete distribution of call attempts in round 3 of CTS, we randomly truncated the number of call attempts. For example, before releasing sample, we randomly selected 70 percent of the cases to have a maximum number of calls set to 12 if they had a status of "ring, no answer"; another 20 percent were randomly selected to have a maximum of 20 calls in that situation, and the remaining 10 percent had a maximum of 30 calls.

In round 4, we did not impose this limitation on the survey operations staff. Instead, to implement the survival method in round 4, we retroactively "pseudo-truncated" calls randomly—and for this evaluation we did it two different ways. In the first variation, we randomly truncated 70 percent of cases at 40 calls, 20 percent at 50 calls, and 10 percent at 60 calls. In the second variation, we randomly truncated 70 percent of cases at 20 calls, 20 percent at 30 calls, and 10 percent at 40 calls. We note that, although truncated, these are still fairly high numbers of call attempts. An example will illustrate how this pseudo-truncation worked.

Suppose a telephone number was attempted 22 times, with the first 21 attempts resulting in outcomes of "no answer," "busy," and "answering machine" (all classified as unresolved) and the 22nd call resulting in a refusal. Because a refusal is classified as residential, the case would be classified as residential. However, if this case was one of the 70 percent of cases randomly chosen to be pseudo-truncated at 20 calls, it would be re-classified as unresolved, because at the time of call attempt 20, there was no evidence of residential status.

Following the algorithm described in the Brick et al. paper, we programmed the survival method calculation of residency rates into a SAS program. Because of the nature of the CTS sample and the likelihood that the rates would be quite different for the "re-interview," "other overlap," and "new" samples, we calculated the residency rates separately for these three groups.

We ran much of our sample through a screening program run by our sampling vendor (Marketing

Systems Group-Genesys) that attempts to remove most businesses and non-working telephone numbers from the sample before releasing to interviewers. For both the CASRO and survival analysis methods, we included in the counts of those "resolved as non-residential" those telephone numbers that were rejected by this screening program, and therefore never released. For the survival method, they were included in the same category as those resolved as non-residential at the first call attempt.

Results. Table 1 shows the results for round 4 of the CTS. Each column represents one of the three sample components described above. The top third of the table shows the results before applying the pseudo-truncation of call attempts. The middle of the table shows the results after truncating randomly at 40, 50, or 60 calls, and the bottom of the table shows the results after truncating randomly at 20, 30, or 50 calls. Within each of these, we present the overall residency rate according to the CASRO method (R_{casro}), the overall residency rate according to the survival method (R), and the residency rate among unresolved cases according to the survival method (R_{un}).

Table 1. CTS Round 4 Estimated Residency Rates

	Re-interview sample	Other overlap sample	Sample new to round 4
No truncation of call attempts			
R_{casro}	.781	.341	.329
R	.776	.330	.331
R_{un}	.451	.257	.372
Random truncation at 40, 50, or 60 calls (70%, 20%, 10%)			
R_{casro}	.781	.340	.328
R	.775	.326	.323
R_{un}	.372	.235	.234
Random truncation at 20, 30, or 50 calls (70%, 20%, 10%)			
R_{casro}	.781	.338	.326
R	.770	.304	.312
R_{un}	.220	.086	.084

While the estimate of R is not used as a residency rate in the survival method (it is merely an interim step), we first compare the two overall residency rates. Across all three sample types, and all three truncation scenarios, the overall residency rate is quite similar between the CASRO and survival methods. For example, among the new sample cases, and the truncation at the higher values, the overall residency rate under CASRO is .328, and under the survival method is .323. Within sample type, these rates remain fairly constant across the three truncation scenarios. As expected, the residency rate

among unresolved cases is, in all but one situation, smaller than the overall residency rate. But, unlike the overall rates, the rates for unresolved cases vary widely across truncation scenarios (within sample type). For the new sample, the residency rates vary from .372 (without truncation) to .084 (with truncation at the lower values).

Moving to the Round 3 results, recall that we actually truncated the call attempts in Round 3 (rather than retrospectively pseudo-truncating the call attempts as we did in round 4). But in round 3, the first round in which we applied the survival method, we struggled with how to classify three ambiguous status codes: (1) maximum calls-residence, (2) maximum calls-probable residence, and (3) answering machine. They all involved some interviewer or supervisor judgment as to the likelihood that the telephone number was residential, so we debated whether they should be classified as "residential" or "unresolved."

Table 2 shows the round 3 results under three different classification scenarios: (1) the two "maximum calls" codes were classified as "residential" while the "answering machine" code was classified as "unresolved," (2) all three codes were classified as "unresolved," and (3) all three codes were classified as "residential."

Table 2. CTS Round 3 Estimated Residency Rates

	Re-interview sample	Other overlap sample	Sample new to round 3
Maximum calls-(probable) residence= "residence" and answering machine = "unresolved"			
R_{casro}	.797	.247	.340
R	.792	.266	.347
R_{un}	.505	.424	.439
Maximum calls-(probable) residence and answering machine = "unresolved"			
R_{casro}	.796	.242	.337
R	.789	.241	.334
R_{un}	.411	.241	.296
Maximum calls-(probable) residence and answering machine = "residence"			
R_{casro}	.797	.248	.341
R	.788	.227	.322
R_{un}	.104	.047	.044

Once again we see that the two methods produce very similar overall residency rates within sample type. Here we also see that the overall residency rates stay fairly consistent across the various classifications of the three ambiguous status codes. Among the new sample, the overall residency rates fall in the range .322 to .347 across the two

methods and the three classification scenarios. But once again we see that the rates for unresolved cases vary widely across classification scenarios (within sample type). For the new sample, the residency rates vary from .439 (where the two maximum calls cases are classified as residential and answering machines are classified as unresolved) to .044 (where all three are classified as residential); in other words, changing the classification of answering machines from unresolved to residential drastically reduced the residency rate among unresolved cases while the overall residency rate (only an interim step in the survival method) remained fairly constant.

Discussion and Conclusions. Using survival methodology to estimate residency rates in RDD surveys is an innovative idea and one that is promising theoretically. As the results showed, the CASRO and survival methods produced virtually identical overall residency rates in both rounds 3 and 4 of the CTS (which had a high number of call attempts). These rates were not sensitive to changes in truncation strategy nor to changes in the classification of a few ambiguous status codes. There was no practical difference between the two methods, and the overall residency rates appear to converge when there are a large number of call attempts, as there were for the CTS. On the other hand, the estimated residency rate among unresolved cases, the primary outcome of the survival method, varied significantly, and was highly sensitive to these changes. It also resulted in some estimated residency rates that were so low (4.4 percent) that they strain credulity.

There is variability associated with these Kaplan-Meier survival functions, and the variability increases as we move further to the right on the distribution (that is, as the number of call attempts increases). Because of the large number of call attempts in CTS, the variability surrounding the overall residency rate is expected to be large, with small changes in the residency rate on the upper tail having a disproportionate effect; however, we observed relatively small variation in the overall residency rate. The variability we saw was in the rate for unresolved cases. Likewise, the Brick et al. paper presented widely varying residency rates for *unresolved* cases across surveys, and associated standard errors that were fairly large.

Furthermore, because the unresolved population changed when we truncated the call attempts, we would expect the residency rate to change accordingly. The changes we observed had more to do with the last step—using the *estimated* overall residency rate and the *actual* number of cases

resolved as residential to estimate the residency rate for unresolved cases.

For example, when comparing the untruncated to the truncated residency rates for new sample in round 4 (truncated at the higher level), only eight telephone numbers changed from being classified as residential to unresolved. And the overall residency rate changed slightly from .331 to .323. But these small changes had a large impact on the unresolved residency rate. When the overall residency rates are applied to the sample size of 33,117 to estimate the total number of residences in the sample, and the resolved-as-residential cases are then subtracted, we are left with very different estimates of the residences among the roughly 1,850 unresolved cases: 685 versus 428 (resulting in residency rates among the unresolved of .372 versus .234).

The initial programming of the survival method is more difficult than the CASRO method, but only has to be done once (within minor modifications each time it is run for a survey). However, there is work involved in using the call history data to obtain call attempts and individual outcomes, and further work involved to ensure a real or simulated distribution of call attempts. Is it worth the effort? We found that we were uncomfortable reporting the survival method residency rate estimates for CTS due to the instability we were seeing, and decided to report the CASRO method results instead.

While not implemented with the CTS survey, incorporating whether the phone number is published or unpublished into the survival method may improve residency rate estimates. More practical questions, such as cost issues related to the number of call attempts, the impact on estimates of these additional attempts, and which cases to continue calling, should also be investigated further.

While RDD surveys appear to be decreasing in popularity in recent years, it is important that others doing such surveys look at how the survival method works for them when trying to estimate residency. One of the reasons for the decline in popularity is the corresponding decline in response rates for this type of sample survey, which may be due to overly conservative estimates of the eligible sample—something that the survival analysis method was trying to address.

When all is said and done, this is an intellectual effort aimed at improving the estimate of the residency rate for unresolved telephone numbers. But in practical terms, it does little to ascertain the quality of the survey and, in fact, one might argue

that intellectual effort to measure response bias may be more fruitful and productive.

References.

Frankel, LR, 1983. "The Report of the CASRO Task Force on Response Rates." In *Improving Data Quality in a Sample Survey*, edited by F. Wiseman. Cambridge, MA: Marketing Science Institute.

Brick MJ, J Montaquila, F Scheuren, 2002. "Estimating the Residency Rates for Undetermined Telephone Numbers." Public Opinion Quarterly, volume 66:18-39. American Association for Public Opinion Research.

Strouse R, B Carlson, J Hall, 2003. "Community Tracking Study Household Survey Methodology Report, 2000-01 (Round Three)." Report submitted to the Center for Studying Health System Change. Princeton, NJ: Mathematica Policy Research, Inc. (HSC Technical Publication #46, www.hschange.org).