Sampling Refusals: Why, When, and How Much?

Timothy Triplett, Adam Safir, Kevin Wang, and Natalie Abi-Habib

1. Introduction

The 2002 National Survey of America's Families (NSAF) is a dual frame survey that relies primarily on a large RDD sample design using over 500,000 telephone numbers. The survey consists of a short three-minute screener survey used to determine eligibility followed by a 45-minute extended interview. Although interviews are completed with close to 40 percent of all initial refusals, the per-interview cost of converted refusals far exceeds that of initial cooperators. In addition, the data collection period is lengthened by refusal conversion. Almost half of all potential respondents initially refused to participate, creating cost and scheduling problems. These problems contributed to the decision not to attempt to convert screener refusals in the final 20 percent of the NSAF 2002 sample release. This strategy for dealing with nonresponse is sometimes referred to as "double sampling" or "two-phase sampling." The purpose of using a double sampling approach is that it offers a method of balancing costs and errors in deciding what efforts should be made to measure sample persons who refuse initial interview attempts (Groves 1989). While we will mention some of the literature that has tried to quantify the balancing of costs and errors, this paper will focus on possible auxiliary effects to the cooperation, response, and refusal conversion rates as a result of the double sampling approach.

2. Double Sampling

The double sampling theory has been around for a long time (Neyman 1938; Deming 1953; Hansen and Hurwitz 1946). A double sampling strategy is appealing when the total cost of a survey must be within the amount appropriated, and when some clearly defined portion of your population of interest is more difficult to interview. The double sampling strategy as it pertains to this research involves dividing the sample into two groups, those who initially refuse to do the survey and those who do not refuse. The cost and time associated with completing interviews in households that initially refuse is assumed to be greater. You would attempt to complete an interview in 100 percent of the households that never refused, while attempting to complete an interview in some fraction (80 percent in the NSAF) of households that initially refused. Resources saved from having to do less refusal conversion will effectively increase the overall size of the sample that can be worked given a fixed budget. This increased sample size will improve the precision of the survey estimate. However, for the final sample to be representative, the initial refusal sample should be weighted by the inverse of the sampling fraction (1.25 in the 80 percent example). This additional weighting factor will increase the variance of the survey estimate. Choosing the optimal fraction of refusals to call back depends on the trade-off between increasing final sample size and increasing the variance associated with the weighting adjustment from sampling refusals. The difficulty with using a double sampling approach for dealing with refusals is that choosing an optimal sampling fraction depends on cost and error information that is not known before data collection Further complicating matters is that some other factors could be associated with the double sampling process itself that could alter cost and error estimates. Deciding not to call refusals may affect interviewer morale, interviewer's perception of the

importance of not getting a refusal, and could also affect interviewer work assignments. There are probably other factors that may affect the cost and error estimates from deciding to sample households that initially refused. The main purpose of this paper is to see how much of an impact other factors may have had on the double sampling process carried out in the 2002 NSAF.

3. NSAF Sample Design

The purpose of the NSAF is to assess the impact of recent changes in the administration of a number of assistance programs for children and the poor. The NSAF sample is designed to generalize to 13 specific states, as well as to the nation as a whole. The design also includes an oversampling of households that were estimated to be under 200 percent of the federal poverty level as well as households with children. All three rounds of NSAF data collection (1997, 1999, and 2002) were done for the Urban Institute by Westat.[8]

The NSAF consists of both a screening and an extended interview. The screener consists of about 3 minutes of questions designed to assess household eligibility and select a respondent for the extended interview. The sampling of initial refusals applied only to the screening interview. The RDD telephone sample for the NSAF was randomly divided into 101 replicates. The first 81 replicates were designated as the refusal conversion sample while the final 20 replicates were designated as the nonrefusal conversion sample. Therefore, the fraction of nonrespondents who received standard refusal conversion was actually slightly higher than 80 percent (81/101 replicates). The no-refusal conversion sample (last 20 replicates) was released by August 2002 and data collection ended on November 3, 2002, therefore all telephone numbers in the 101 replicates were able to receive their full complement of call attempts.

4. Possible Auxiliary Effects

If sampling refusals did affect the survey process, you would expect to see differences in the initial cooperation rate between the first 81 replicates where standard refusal conversion efforts were carried out versus the final 20 replicates where refusal conversion was not done. Additionally, you might expect a change to occur in the refusal conversion rate, since fewer refusals will be called back, which would affect the interviewers work assignments.

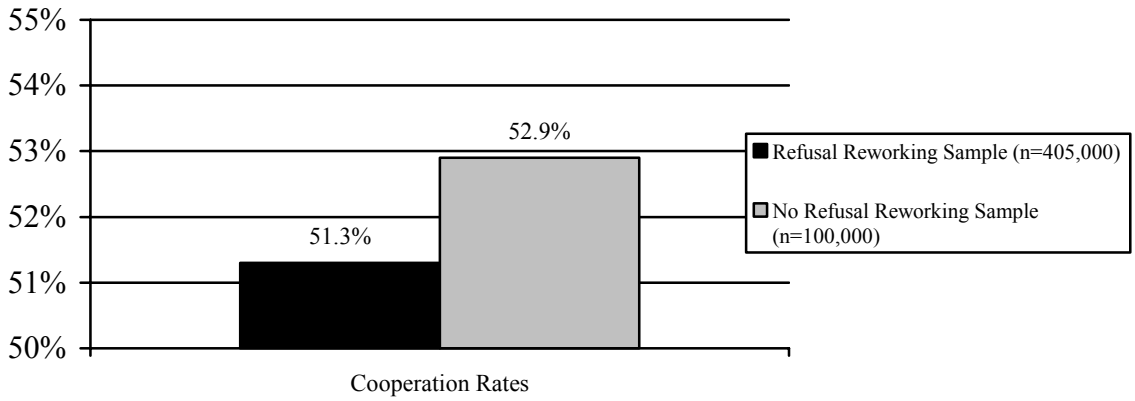
We used the NSAF data to test whether sampling nonrespondents affects the initial cooperation or refusal conversion rates. There are two cooperation rate comparisons of interest. First, there is the simple comparison of initial cooperation rates from the refusal conversion sample versus the nonrefusal conversion sample. Second, we tested for an effect on the initial cooperation rate of the refusal conversion sample in correlation with the release of the no refusal conversion sample. Testing for an effect on refusal conversion rates involves comparing refusal conversion rates before and after the release of the nonrefusal conversion sample.

8 The screener response rate has declined from 77% in 1997 to 66% in 2002.

5. Results of the Tests

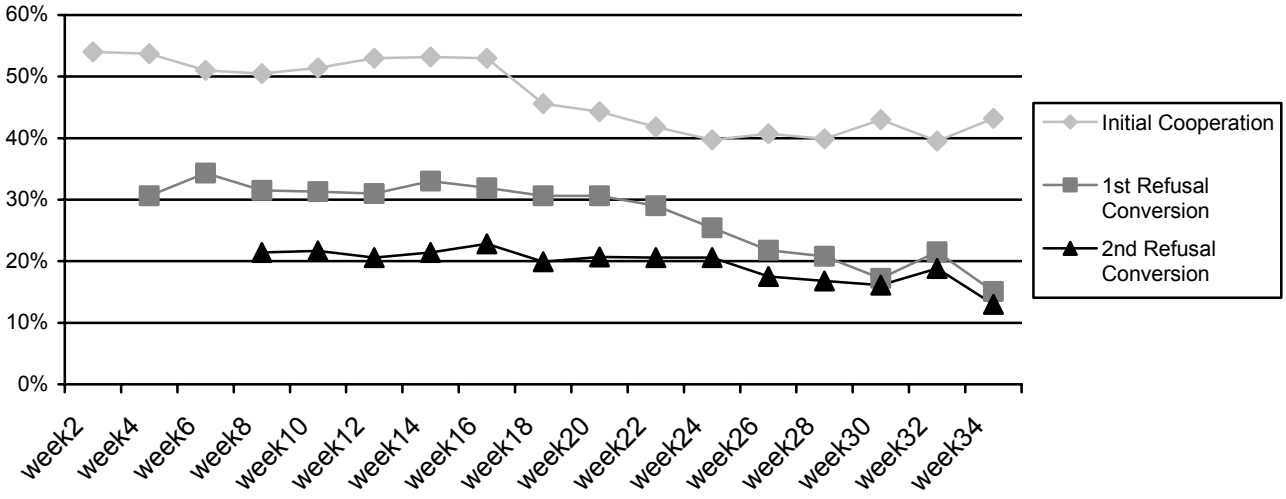
The first test compared initial cooperation rates in both samples, first the 81 replicate releases (405,000 telephone numbers) in which refusals were reworked versus the 20 replicates (100,000 telephone numbers) in which refusals were not called As seen in figure 1, the initial cooperation rate is higher for the nonrefusal conversion sample. The increase from 51.3 percent to 52.9 percent is statistically significant, although not large. This increase would support the hypothesis that sampling nonrespondents could improve morale or interviewer staffing. However, given how small the increase is overall, it could also be explained by factors unrelated to the double sample process, such as changes in interviewers skills and morale that evolve over time.

Figure 1. Initial Cooperation Rates



Determining whether the initial cooperation rate in the convert refusal sample was affected by the introduction of the sample where no refusals were reworked is more difficult. This is because the initial cooperation rate of RDD sample usually begins to decline as the number of call attempts needed to reach a household increases (Lynn et al. 2002; Triplett 2002). However, there is some evidence that the initial cooperation rate in the convert refusal sample dropped more dramatically than expected. This can be seen by the sudden drop in the initial cooperation rate indicated by the arrow in figure 2. At week 18 of NSAF data collection, the no-refusal conversion sample began being called by the interviewers. Therefore, any effect from the change in how the sample is worked would be expected to occur shortly thereafter. It is generally expected that there would be a decline in the initial cooperation rate of older sample any time an additional sample is released. However, the decline in week 18 was larger than occurred during any other week of the study, and there were several other times during the study when a new sample was released. Thus, it appears that giving interviews a fresh sample that was not going to require refusal conversion had unintended negative effects on the initial cooperation rate of the older sample. Was this decline a result of reduced enthusiasm for making call attempts on the older sample that requires refusal reworking? While this could be an explanation, it is not something we are able to test for.

Figure 2. Refusal Reworking Time Line



Testing if the refusal conversion rate benefited from the reduction in total refusals needing rework is also quite difficult to measure. The effects from having fewer refusals to convert would likely occur a few weeks after the no-refusal conversion sample was released, since conversion attempts require at least a two-week waiting time on the NSAF. We would expect softer refusals would be converted sooner, thus over time, the refusal conversion rate would fall. The refusal conversion rate for both first and second refusals gradually decreases over time (figure 2). Since the no-refusal conversion sample was released in week 18, the benefit from having fewer refusals to rework would have begun by week 20. It appears that having fewer refusals to convert may have helped refusal conversion. In week 20, the refusal conversion rate on first refusals stopped declining, while the conversion rate for second refusals increased slightly. These changes are not large and could have other explanations, but they support the hypothesis that by reducing the number of refusals that need converting, it becomes easier to assign the better interviewers to rework refusals.

6. Summary Discussion

We have explored some of the subtler effects of sampling nonrespondents. A more important issue in deciding whether to sample refusals and how many is how well refusal conversions improve estimates. Recent studies have found little or no reduction in nonresponse bias from efforts at increasing response rate (Lynn et al. 2002; Teitler et al. 2003; Curtin et al. 2000; Keeter et al. 2000). These findings suggest that the fraction of refusals that should be called back could be set much higher than the 20 percent chosen for the NSAF without affecting survey estimates.

One thing we know from the NSAF data collection is that we were able to finish the study earlier than we would have had we not sampled refusals. We also were able to increase the amount of overall sample released. It is less clear whether we benefited from higher initial cooperation rates due to sampling refusals, but we certainly did not do any worse. Since we achieved a higher cooperation rates from the no-refusal conversion sample, the weight adjustment (total initial

refusals/initial refusals from the sample receiving conversion attempts) was slightly smaller than expected. A smaller weighting adjustment reduces the variance associated with this weight.

While our findings on the effects of sampling refusals on cooperation and refusal conversion rates were not very strong, this could be in part due to this not having been a controlled experiment. A controlled experiment with careful interviewer assignments would have improved our ability to test for effects on cooperation rates from sampling refusals. While a controlled experiment might provide additional evidence of auxiliary effects from sampling refusals, it is still likely that these effects will be small and even less likely that they would wind up negatively affecting the survey. With respondents increasingly refusing surveys, it is likely that there will be more double sampling as way of handling nonresponse in the future. Therefore, more research is needed to assess the effects of sampling nonrespondents on the overall survey error.

References

Brooks, Camilla A., and William D. Kalsbeek. 1982. "The Double Sampling Scheme and Its E(MSE)." In *Proceedings of the Section on Survey Research Methods* (235–39). Alexandria, VA: American Statistical Association.

Curtin, Richard, Stanley Presser, and Eleanor Singer. 2000. "The Effect of Response Rate on the Index of Consumer Sentiment." *Public Opinion Quarterly* 64(4): 413–28.

Deming, Edwards W. 1953. "On a Probability Mechanism to Attain an Economic Balance between the Resultant Error of Response and the Bias of Nonresponse." *Journal of the American Statistical Association* 48(264): 743–72.

Elliott, M.R, and Roderick J.A. Little. 2000. "Subsampling Callbacks to Improve Survey Efficiency." *Journal of the American Statistical Association* 95:730–38.

Groves, Robert M. 1989. *Survey Costs and Survey Errors*. New York: Wiley.

Hansen, Morris H., Hurwitz. 1946. "The Problem of Non-Response in Sample Surveys." *Journal of the American Statistical Association* 41(236): 517–29.

Keeter, Scott, Carolyn Miller, Andrew Kohut, Robert M. Groves, and Stanley Presser. 2000. "Consequences of Reducing Nonresponse in a National Telephone Survey." *Public Opinion Quarterly* 67(2): 125–48.

Lynn, Peter, Paul Clarke, Martin, and Paul Sturgis. 2002. "The Effects of Extended Interviewer Efforts on Nonresponse Bias." In *Survey Nonresponse*, edited by Robert M. Groves, Don Dillman, John L. Eltinge, and Roderick J.A. Little (135–47). New York: Wiley.

Neyman, J. 1938. "Contribution to the Theory of Sampling Human Populations." *Journal of the American Statistical Association* 33(201): 101–16.

Teitler, Julien O., Nancy E. Reichman, and Susan Sprachman. 2003. "Costs and Benefits of Improving Response Rates." *Public Opinion Quarterly* 67(1): 126–38.

Triplett, Timothy. September 2001. "What Is Gained from Additional Call Attempts and Refusal Conversion, and What Are the Cost Implications?" Unpublished ongoing research paper. http://mywebpages. comcast.net/ttriplett13/tncpap.pdf.