Using a Large-Scale Field Study to Estimate Nonresponse and Noncoverage Biases in an RDD Survey

Wendy Van de Kerckhove[1], Jill Montaquila[1], J. Michael Brick[1]
[1]Westat, 1650 Research Boulevard, Rockville, MD 20850

Abstract
Concerns about declining response and landline telephone coverage rates in Random Digit Dial (RDD) surveys led to the development and administration of an independent Bias Study that was conducted in conjunction with the 2007 National Household Education Surveys Program (NHES:2007). For the Bias Study, a clustered sample of addresses was selected, telephone numbers were matched to the addresses to the extent possible, and the matched sample used the regular RDD protocol. Addresses that could not be matched or completed by telephone were sent into the field for in-person attempts. In this paper, we describe the methods used to estimate nonresponse and noncoverage bias, present estimates of the two types of bias, and interpret the findings in the broader context of other nonresponse and noncoverage bias studies.

Key Words: CATI, response rate, coverage

1. Motivation and Background

Throughout the past decade, both response rates and coverage rates of landline random digit dial (RDD) samples have been declining precipitously. Curtin, Presser, and Singer (2005) found an average annual decline in response rates of 1.5 percentage points for the Survey of Consumer Attitudes between 1997 and 2003. Groves et al. (2002), Battaglia et al. (2008), and Holbrook, Krosnick, and Pfent (2008) provide further discussion of declining response rates. In addition, the limitation of RDD surveys to landline telephone numbers has been recently detrimental to coverage rates because of the increasing prevalence of households having only cell telephone service. In 2004, 4 percent of adults were in wireless-only households compared to 15 percent in late 2007 (Blumberg and Luke, 2008).

The National Household Education Surveys Program (NHES), sponsored by the National Center for Education Statistics, is one set of ongoing, periodic surveys that is facing these declining response and coverage rates. Conducted nine times from 1991 to 2007, NHES Screener response rates have declined from above 80 percent in the early 1990s to 53 percent in 2007. Declines have been observed in NHES extended interviews as well. For the adult extended interview, the response rate was 85 percent in 1991 and fell to 62 percent in 2007. In addition, the NHES survey is restricted to households with landline telephone numbers, and thus is subject to noncoverage of cell phone only households. Further background on NHES is provided in section 1.1

As a result of these trends in response and coverage rates, there is increasing concern among survey methodologists and practitioners about the potential for bias in RDD survey estimates. A general discussion of the effects of nonresponse and noncoverage on bias is given in section 1.2. To better evaluate the effects on the NHES estimates, an extensive Bias Study was conducted in 2007 involving an in-field follow-up of cases for which an interview could not be completed through the usual RDD procedures. The design and implementation, estimation methods, and results of the Bias Study are described in sections 2, 3, and 4, respectively.

1.1 National Household Education Surveys (NHES) Program
The NHES surveys are designed to collect information on a range of educational issues through computer-assisted telephone interview (CATI) surveys of households. Each NHES administration includes a brief screening interview (Screener) and two or three topical interviews. Efforts have been made to maximize response rates through the use of extended calling protocols, increased refusal conversion efforts, and monetary incentives.

The NHES:2007 RDD study was conducted from January to May 2007. The RDD sample for NHES:2007 included 278,490 telephone numbers. Address matches were available for 55 percent of the sampled telephone numbers[1], and an advance letter with a $2 incentive was mailed to these addresses. The NHES data collection protocol included up to at least 14 call attempts to complete a Screener. A random subsample of the refusal households with matched addresses received a letter, with a $2 incentive, encouraging response, and refusal conversion attempts were made. At the screening stage, household members were enumerated as needed to identify eligible household members and apply the within-household sampling algorithm. The NHES interviews were conducted in English and Spanish. Screeners were completed with 54,034 households.

NHES:2007 included three topical surveys. The School Readiness (SR) Survey contains measures of preschoolers' (ages 3 through 6) literacy and numeracy, as well as items on early childhood care and programs. The Parent and Family Involvement in Education (PFI) Survey contains items on school characteristics, student experiences, communication with other parents, and family involvement in school for children enrolled or home schooled in kindergarten through twelfth grade. The Adult Education for Work-related Reasons (AEWR) Survey was administered to sampled persons 16 years or older who were not enrolled in twelfth grade or below, and contains items pertaining to participation in both formal and informal work-related adult education. Numerous call attempts, refusal conversion attempts, and refusal letters were also used at the interview stage to encourage response. The overall response rates for the NHES:2007 SR, PFI, and AEWR surveys in the RDD study were 41 percent, 39 percent, and 33 percent.

1.2 Nonresponse and Noncoverage

The theory of sampling that is the basis for the majority of surveys conducted for the federal government assumes that accurate responses are obtained for all the sampled units and that the sample covers the target population. To the extent that those who respond to surveys and those who do not are different in important ways, there is a potential for biases in estimates from survey data.

As survey response rates decline, understanding the relationship between response rates and nonresponse bias has become increasingly important. One approach to understanding the relationship is to conduct nonresponse bias studies. The literature is replete with examples of such studies (see, for example, Boyle et al., (2002); Cohen and Duffy, (2002); Garretsen et al., (2002); and Keeter et al., (2000)). Even when there is the potential for nonresponse bias in survey estimates, statistical adjustments that account for differences in response propensities can effectively reduce nonresponse bias. Groves and Petcheyeva (2008) presents the results of a meta-analysis involving 59 studies that revealed virtually no relationship between response rates and nonresponse bias.

Differences in telephone noncoverage rates, especially differential rates among population subgroups, such as those defined by region, age, race/ethnicity, and household composition, are of concern to telephone survey methodologists because they too can introduce bias in the estimates. Tucker et al. (2004) and Blumberg and Luke (2008) examined differences in characteristics among persons and households having no telephone service, cellular service only, and landline service (including both landline only, and landline and cellular). Although there are differences in landline noncoverage (e.g., young adults, adults in one-person households, renters, and Blacks and Hispanics are less likely to have landlines), as with adjustments for nonresponse, statistical adjustments that account for these subgroup differences can reduce noncoverage bias.

Although statistical adjustments may effectively reduce or remove bias due to nonresponse or noncoverage, these adjustments rely on knowledge of the characteristics associated with response propensities and noncoverage, and the ability to measure these characteristics and use them in the adjustments. Thus, even when the estimation methodology includes such adjustments, studies to evaluate bias are important. The most common methods use existing data to evaluate bias and can be effective at detecting and helping to correct for it. Such methods have been used in the evaluation of bias in earlier NHES surveys. (See Brick, (1996); Brick et al., (1997); Brick, Burke, and West, (1992); Montaquila, Brick, and Brock, (1997); Nolin et al., (2000); Nolin et al., (2004); and Roth, Montaquila, and Chapman, (2007)). The study described here differs from these past evaluations in that it involves a full-scale data collection effort directed entirely at estimating bias in key NHES survey outcomes. As a result, this study complements those earlier investigations.

[1] Using two-phase selection, telephone numbers with matched addresses were subsampled at higher rates than those without matched addresses.

2. Design and Implementation

The purpose of the Bias Study for the NHES:2007 was to obtain a more direct assessment of bias in key survey outcomes through the use of an in-field follow-up. The Bias Study involved a sample of addresses, including both households with a landline telephone and those without. The sample design is described further in section 2.1. Cases were first attempted by telephone, when possible, and cases not completed by telephone were then attempted in the field. The implementation of the Bias Study and success of the data collection effort are described in section 2.2. The Bias Study was primarily designed to evaluate overall bias in the NHES:2007 estimates, but it also allowed for a separate analysis of the unit nonresponse and noncoverage bias components. Following up with nonrespondents to the telephone effort allowed an investigation of nonresponse bias, and following up with households with no landline telephone allowed an investigation of noncoverage bias.

2.1 Sample Design

The Bias Study involved of a clustered sample of 7,500 addresses, selected independently of the main RDD survey. The addresses were sampled in two phases. First, a large sample of addresses was selected, and an attempt was made to match the addresses to telephone numbers. Then, the final sample of addresses was selected from the first phase sample, where addresses with a matching telephone number were sampled at a higher rate than those without a match. In the final sample, 72 percent of the addresses had a matching telephone number. The oversampling was done to ensure that enough cases were available to attempt in the initial telephone effort for the evaluation of nonresponse bias.

To sample persons for the topical surveys, the same within-household sampling algorithms were used as with the main RDD survey. Sample sizes were set to allow for a detection of a 5 percentage point difference between the Bias Study estimates and RDD survey estimates of key statistics from each of the topical surveys. Differences of 5 percentage points or more are generally considered to be of substantive importance for NHES.

2.2 Data Collection

In order to assess bias in the NHES:2007 estimates, the Bias Study sample cases were first attempted by telephone, using the same protocol as the main RDD sample. After the initial telephone data collection effort, cases with no matching telephone number or an incorrectly matched telephone number, as well as nonrespondents to the initial telephone effort, were sent to the field for in-person follow-up. The goal of the field staff was to gain cooperation from a household member to complete the Screener. Upon gaining cooperation, the field staff would then connect the respondent with the telephone center so that the respondent could complete the interview by telephone. Several methods were used to maximize response rates in the field, including an advance letter to the cases with no matching phone number (since they did not receive one prior to the telephone effort), an additional incentive offered at the door, and numerous attempts to complete the Screener. The interviewer was also responsible for completing an Interviewer Observation Form (IOF) before approaching the household. The IOF contained questions about the household and neighborhood characteristics for all cases sent to the field, and thus could be used in analyzing the nonresponse bias from cases that did not respond to the Screener in either the telephone or field effort.

Of the 7,500 Bias Study sample cases, 5,433 were attempted in the initial telephone effort. Of these 5,433 cases, 2,396 (44 %) were completed through this effort. An additional 49 cases were finalized after the initial telephone effort as hostile refusals plus 9 as other nonrespondents (e.g. the only member of households was suffering from dementia). The remaining 5,104 cases were sent to the field for in-person follow-up.

A total of 3,576 cases (approximately 70 % of cases sent to the field) were those for which the sampled address had not been previously attempted in the telephone effort, i.e. the cases with no matching telephone number or an incorrectly matched telephone number. A telephone number was determined to be incorrectly matched if the number obtained for the sampled address was found to be nonworking or nonresidential, or if the address provided by the respondent to the telephone interview did not match the sampled address. The rest of the cases sent to the field were the nonrespondents to the telephone effort, consisting of nonhostile refusals (693 cases), maximum calls (477 cases), noncontacts (310 cases), and language problems (48 cases). Overall, 50 percent of cases sent to the field were completed. The highest response rates in the field (52 to 54 %) were for cases that had not been previously attempted by telephone. However,

with the exception of language problems (with a 23 % response rate), the in-field effort proved successful for telephone nonrespondents as well, with in-field response rates of 42 to 46 percent.

After the initial telephone effort and the following field effort, Screeners were completed with a total of 4,894 households. The Screener response rate was 68 percent, 15 percent higher than that for the RDD survey. The overall response rates for the SR, PFI, and AEWR surveys were 54 percent, 51 percent, and 41 percent. While the Bias Study response rates are an improvement over the RDD survey, there are still 32 percent of cases for which no Screener or topical survey data was obtained.

3. Estimation of Bias

Because the Bias Study covers nontelephone households and includes in-field followup of telephone interview nonrespondents, it can be used to assess bias in the RDD estimates. The analysis includes an assessment of overall bias (bias attributable to both noncoverage and nonresponse), as well as the separate noncoverage and nonresponse bias components.

Exhibit 1 contains a classification of all cases with completed Screeners in the NHES:2007 RDD and Bias Study samples. This exhibit provides notation that is useful for describing which sets of cases were used to estimate components of bias. For example, S_{Bmf} is the set of Bias Study cases with completed Screeners that are landline telephone households with matching telephone numbers, and were completed as a result of the field effort.

Exhibit 1: Classification of cases with completed screeners in the NHES:2007 RDD and Bias Study samples

		Completed by phone	*Completed in field*
RDD sample	*Telephone household with matching telephone number*	S_{Rmp}	*
	Telephone household without matching telephone number	*	*
	Nontelephone household	*	*
Bias Study sample	*Telephone household with matching telephone number*	S_{Bmp}	S_{Bmf}
	Telephone household without matching telephone number	*	S_{Bzf}
	Nontelephone household	*	S_{Bnf}

** Indicates combinations that, by design, were not possible.*

Using the notation given in Exhibit 1, we define the *reduced effort* as $RE = S_{Bmp} \cup S_{Bzf} \cup S_{Bnf}$, and the *full effort* as $FE = RE \cup S_{Bmf}$. Each of these subsets, *RE* and *FE*, was fully weighted separately, using the same methodology.

Overall bias was estimated by examining differences between estimates from the RDD sample (S_{Rmp}) and the full effort Bias Study sample ($S_{Bmp} \cup S_{Bmf} \cup S_{Bzf} \cup S_{Bnf}$). Because the Bias Study sample includes both telephone and nontelephone households, and the data collection protocol involved attempts to contact both subsets while minimizing the possibility of mode effects, estimates of noncoverage bias in the RDD estimates can be obtained as the differences between Bias Study estimates for telephone households ($S_{Bmp} \cup S_{Bmf}$) and Bias Study estimates for all households ($S_{Bmp} \cup S_{Bmf} \cup S_{Bzf} \cup S_{Bnf}$).

Nonresponse bias can be assessed by examining differences in estimates from the reduced effort and the full effort. Nonmatched cases (cases in sets S_{Bzf} and S_{Bnf}) could not be attempted in the initial telephone effort, but were attempted in the field. Because these cases comprise part of the target population, the reduced effort estimates were

calculated using data from both the TRC respondents and the nonmatched cases, in order to eliminate the effect nonmatched cases would have on estimates of nonresponse bias.

To better understand the rationale for including the nonmatched cases in both sets, consider the difference between the reduced and full effort estimates (and here we illustrate this using estimates of totals) as

$$\hat{Y}_{RE} - \hat{Y}_{FE} = \left[\sum_{i \in S_{Bmp}} w_i y_i + \sum_{i \in S_{Bzf}} w_i y_i + \sum_{i \in S_{Bnf}} w_i y_i \right] - \left[\sum_{i \in S_{Bmp}} w_i' y_i + \sum_{i \in S_{Bmf}} w_i' y_i + \sum_{i \in S_{Bzf}} w_i' y_i + \sum_{i \in S_{Bnf}} w_i' y_i \right].$$

If the weights (w_i and w_i' for the reduced effort and the full effort, respectively) were identical for the two sets, then it is clear that including the nonmatched cases would have no effect on the estimated difference. In practice, the weights do differ, mainly due to differences in the adjustments to population control totals, so this simple result does not hold exactly but the inclusion of the nonmatched cases in both sets has little effect.

An alternative treatment of the nonmatched cases would be to use propensity modeling to assign each a response status for the initial telephone effort. A logistic regression model was fit with zip-code level Census 2000 variables as the predictors. The model proved ineffective, since the mean response propensity for telephone respondents was 0.78, compared to 0.77 for the telephone nonrespondents who responded in the field. The predicted response propensities for the nonmatch cases then ranged from 0.63 to 0.88.

For the analysis of overall bias, as well as the components, the comparisons involved estimates of important demographic and socioeconomic indicators as well as key survey outcomes. For the School Readiness (SR) Survey, over 35 characteristics were evaluated for bias, such as the percentage of preschoolers who participated in center-based care and the percentage of preschoolers with household incomes under $15,000. For the Parent and Family Involvement in Education (PFI) Survey, over 30 characteristics were evaluated, such as the percentage of student's whose parents participate in 5 or more activities in the student's school and the percentage of student's whose parents' highest level of education is a high school diploma or below. For the Adult Education for Work-Related Reasons (AEWR) Survey, over 20 characteristics were evaluated, such as the percentage of adults who participate in adult education for work-related reasons and the percentage of adults who are employed.

4. Results

Despite declining response rates and the restriction of the sample to landline telephone households, the results of the Bias Study indicate no systematic pattern of substantial bias in the NHES:2007 key survey outcome estimates. Specific results from the analysis of overall bias and the bias components are presented in sections 4.1 and 4.2, respectively.

4.1 Overall Bias

In comparing estimates from the RDD study and the Bias Study, no substantive, statistically significant differences were identified for the PFI Survey of school-aged children. Differences were found for the six SR Survey and AEWR Survey estimates shown in Table 1. The differences are expressed in terms of relative bias, which is the bias (i.e. the difference between the RDD estimate and the Bias Study estimate) divided by the Bias Study estimate, expressed as a percentage. These differences might suggest potential bias in RDD study estimates for these characteristics. However, it is also possible the differences are a result of measurement error, nonresponse or noncoverage bias in the Bias Study estimates, or other sampling errors.

For preschoolers, the statistical significance of the lower estimates for the percentage that could count to 20 or higher, whose speech is often understandable to a stranger, and who watch 2 or more hours of TV in a typical weekday for the RDD study compared to the Bias Study could be a result of the large number of comparisons performed. There was no systematic relationship between the RDD and Bias Study estimates of key SR Survey characteristics that would indicate bias. In addition, the differences in the sex distribution of preschoolers is evidence of an anomaly in the Bias Study sample, rather than of bias in the RDD study estimates, since the RDD estimates closely match those of outside

sources[2]. Finally, the higher RDD estimate for the percentage of preschoolers whose mothers are not in the labor force might be evidence of potential bias in the NHES:2007 RDD study estimate of this characteristic. The difference may be an indication of accessibility, with mothers who are not in the labor force being more available and more willing to complete the interview by telephone than mothers with other employment status.

For the AEWR survey of adults, the one estimate found to have potential bias was the percentage of adults who are currently married. Since the estimate from the Bias Study closely matches the estimate from the March 2006 Current Population Survey (CPS), this does suggest bias in the NHES:2007 RDD estimate of this characteristic. This may be indicative of the relative inaccessibility and higher prevalence of cell phone-only households for unmarried adults. However, there was no evidence of bias in the percentage of adults who participate in adult education for work-related reasons, which is the main estimate of interest for the AEWR survey.

Table 1: NHES:2007 characteristics with differences in RDD Study and Bias Study estimates
(SR: school readiness, AEWR: adult education for work-related reasons)

Characteristic	*Relative bias (%)*
SR survey	
Child counts to 20 or higher	–8
Child's speech is often understandable to a stranger	–7
Child watches 2 or more hours of TV in a typical weekday	–13
Child's sex: male	–19
Mothers' employment status: not in labor force	37
AEWR Survey	
Marital status: currently married	10

4.2 Components of Bias

The results of the nonresponse bias analysis indicated no differences between estimates for the reduced effort and full effort. To estimate the bias from cases that failed to respond to either the initial telephone effort or the field effort, IOF and Census characteristics for the field respondents and nonrespondents were compared. A higher proportion of field respondents were found to live in zip codes with lower median home values and lower median income deciles. Also, interviewers classified a higher proportion of field respondents as living in working class or poor households, having evidence of children, and being on blocks where no households had signs for private security, compared to field respondents. For the majority of characteristics examined, however, the field respondents were found to be similar to field nonrespondents.

For the analysis of noncoverage bias, the estimated percentage of preschoolers whose parents' highest level of education was beyond a high school diploma was 14 percent higher for telephone households compared to the full Bias Study estimate. No other indication of noncoverage bias in the NHES:2007 estimates was observed.

5. Summary

The results from the Bias Study showed little evidence of bias in the NHES:2007 estimates. These results were consistent with past bias analyses for NHES surveys. No evidence of nonresponse bias in estimates of the key survey outcomes was found, but a limitation of this study is that it does not examine the potential for nonresponse bias due to the 32 percent of households that did not complete the Screener. Some general characteristics of this group were examined, but there is no information about their key survey outcomes. Some potential for noncoverage bias was found in the estimates of preschoolers whose parents' highest level of education was beyond a high school diploma. Although estimates of noncoverage bias in other final estimates examined in this study do not appear to be sufficiently large to be of substantive importance, noncoverage bias may become more of an issue in the future as more households drop their landline telephone service.

[2] To evaluate the effect of the skewed sex distribution on the analysis of overall bias, the Bias Study weights were re-calibrated to Current Population Survey (CPS) population totals by sex, and the analysis was rerun. The general results were unchanged.

Acknowledgements

The authors acknowledge the support of the National Center for Education Statistics, Institute of Education Sciences for its sponsorship of the Bias Study. The authors are grateful for the collaboration of Marilyn Seastrom, Chris Chapman, and Gail Mulligan throughout this effort.

References

Battaglia, M.P., Khare, M., Frankel, M.R., Murray, M.C., Buckley, P., and Peritz, S. (2008). Response rates: How have they changed and where are they headed? In Lepkowski, J.M., et al. (Eds.).*Advances in Telephone Survey Methodology*. Hoboken, NJ: John Wiley & Sons.

Blumberg, S.J., Luke J.V. *Wireless substitution: Early release of estimates from the National Health Interview Survey, July – December 2007*. National Center for Health Statistics. Retrieved August 22, 2008 from http://www.cdc.gov/nchs/nhis.htm.

Boyle, F.M., Cook, M.D., Dunne, M.P., Najman, J.M., and Purdie, D.M. (2002). Health and demographic characteristics of respondents in an Australian national sexuality survey: Comparison with population norms. *Journal of Epidemiology and Community Health*, 56, 748-753.

Brick, J.M. (1996). *Undercoverage bias in estimates of characteristics of adults and 0- to 2-year-olds in the 1995 National Household Education* Survey *(NHES:95)* (NCES Publication No. 96-29). Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Brick, J.M., Burke, J., and West, J. (1992). *Telephone undercoverage bias of 14- to 21-year-olds and 3- to 5-year-olds* (NCES Publication No. 92-101). Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Brick, J.M., Collins, M., and Chandler, K. (1997). An *experiment in random-digit-dial screening* (NCES Publication No. 98-255). Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Cohen, G., and Duffy, J.C. (2002). Are nonrespondents to health surveys less healthy than respondents? *Journal of Official Statistics*, 18, 13-23.

Curtin, R., Presser, S., and Singer, E. (2005). Changes in telephone survey nonresponse over the past quarter century. *Public Opinion Quarterly*, 69, 87–98.

Garretsen, H.F., Jansen, H.A., Lahaut, V.M., and van de Mheen, D. (2002). Non-response bias in a sample survey on alcohol consumption. *Alcohol & Alcoholism*, 37, 256-260.

Groves, R.M., Dillman, D.A., Eltinge, J.L., and Little, R.J.A. (2002). *Survey Nonresponse*. New York: John Wiley and Sons.

Groves, R.M. and Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias: A meta-analysis. *Public Opinion Quarterly*, 72(2), 167–189.

Holbrook, A. L., Krosnick, J. A., and Pfent, A. (2008). The causes and consequences of response rates in surveys by the new media and government contractor survey research firms. In J. M. Lepkowski et al. (Eds.), *Advances in telephone survey methodology*. New York: John Wiley & Sons.

Keeter, S., Miller, C., Kohut, A., Groves, R., and Presser, S. (2000). Consequences of reducing nonresponse in a large national telephone survey. *Public Opinion Quarterly*, 64, 125-148.

Montaquila, J.M., Brick, M.J., and Brock, S.P. (1997). *Working paper: Undercoverage bias in estimates of characteristics of households and adults in the 1996 National Household Education Survey* (NCES Publication No. 97-39). Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Nolin, M.J., Montaquila, J., Nicchitta, P., Hagedorn, M.C., and Chapman, C. (2004). *National Household Education Surveys Program: 2001: Methodology report* (NCES Publication No. 2005-071). Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Nolin, M.J., Montaquila, J., Nicchitta, P., Kim, K., Kleiner, B., Lennon, J., Chapman, C., Creighton, S., and Bielick, S. (2000). *National Household Education Survey of 1999: Methodology report* (NCES Publication No. 2000-078). Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Roth, S.B., Montaquila, J.M., and Chapman, C. (2007). *Nonresponse bias in the 2005 National Household Education Surveys Program* (NCES Publication No. 2007-016). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Tucker, C., Brick, J.M., Meekins, B., and Morganstein, D. (2004). Household telephone service and usage patterns in the U.S. in 2004. *Proceedings of the Survey Research Methods Section, American Statistical Association* (CD-ROM).