

RDD versus Site Test: Mode Effects on Gathering a Household Roster and Alternate Addresses

Elizabeth Nichols, Jennifer Hunter Childs, and Kyra Linse¹
U.S. Census Bureau 4600 Silver Hill Rd. Washington DC 20233

Abstract

In 2006, U.S. Census Bureau staff conducted two iterations of field testing of a survey instrument that, in part, examines household mobility and identifies people with multiple residences. The purpose of this work is to improve estimates of the accuracy of the decennial census count. In July, field staff tested a computer-assisted personal interview (CAPI) version of the survey in selected areas around Austin, Texas and the Cheyenne River Sioux reservation in South Dakota. In November, many of the same questions were asked in a small-scale nationally-representative (excluding Alaska and Hawaii) computer-assisted telephone interview (CATI) random-digit-dial (RDD) survey. Both surveys collected data using the same set of questions to gather a household roster and to capture other addresses where people in that household stayed during the previous year.

This paper compares the demographic and address data collected in the two studies, comparing results from the RDD study to results from the site test and discusses the biases of each type of field pretest method. We examine the differences in distributions of number and types of people and addresses listed for each household between the two studies. We discuss what we would have learned had we only fielded the site test or the national RDD survey and how fielding both types of studies allowed us to draw broader conclusions than would have been possible with either study alone.

Key words: RDD, CATI, CAPI, field test, mode effects, residence status, within household coverage

1. Introduction

Our research question addresses whether a small-scale nationwide random-digit-dial (RDD) telephone test can be a suitable method for pretesting a large-scale survey on living situations. The Census Bureau usually uses face-to-face site tests to pretest decennial census operations. We seek to examine what data can be gathered concerning living situations from pretests conducted using either testing method.

In 2006, the U.S. Census Bureau conducted two separate field pretests of a survey instrument that, in part, examines household mobility and residency for people with multiple residences. The survey instrument for one of the pretests was part of the 2006 Test for the Census Coverage Measurement (CCM) Program. The purpose of the CCM Program is to provide estimates of the accuracy of the decennial census count. The survey instrument for the other pretest was conducted for research purposes. We compare these two field pretest surveys in this paper.

From July to September, 2006, the CCM face-to-face survey was field-tested in the 2006 Census Test. Temporarily-hired field staff interviewed respondents using a computer-assisted personal interview (CAPI) version of the survey in selected areas around Austin, Texas and in the Cheyenne River Sioux reservation in South Dakota. In November of that year, many of the same questions were asked in a small-scale nationally-representative (excluding Alaska and Hawaii) RDD survey with a sample drawn from banks consisting only of land-line telephone numbers. Permanent telephone staff from one of the Census Bureau's telephone centers used a computer-assisted telephone interview (CATI) instrument to conduct this survey. Both surveys collected data using the same set of questions to gather a household roster and associated demographics of the people in the household, to capture any other places where the people stayed (i.e., if there were any other place where each person might have been counted had this been a real census), and to determine where each person should have been counted according to the census residence rules.

It is important to understand the differences between the two field pretesting methods. At the most general level, the RDD population will vary from the site test population by their residence location in the U.S. Though we did not measure it, theoretically, the RDD study population will also vary from the site test in terms of the percent of people who have land-line phone access. The interviewers used in each study differed. The RDD test used permanent, experienced staff and the site test

¹ This report is released to inform interested parties of ongoing research and to encourage discussion. Any views expressed on statistical or methodological issues are those of the authors and not necessarily those of the U.S. Census Bureau.

used temporary staff. The questionnaire mode also differed. Both used automated instruments; but the RDD was conducted over the telephone, and the site test was conducted via personal visit. We seek to investigate whether these differences translate into differences in the substantive survey data items we are interested in measuring (e.g., the number of places where household members could have been counted in the census).

This paper compares the household roster and address data collected in the two studies, comparing results from the RDD study to results from the CAPI interview in the two sites. We examine the differences in distributions of number of people and addresses listed for each household between the two studies. We discuss what we would have learned had we conducted only the site test or the national RDD survey and how fielding both types of studies allowed us to draw broader conclusions than would have been possible with either study alone.

Before we go into details on the study at hand, we will examine pretesting methods generally, the literature on mode effects between telephone and face-to-face survey administration, and research that investigates the use of address- and telephone-based sampling frames. Then, we will present our study examining field test data gathered by an RDD study as compared to that gathered by a small-scale site test.

2. Background

2.1 Pretesting

After initial cognitive testing comes formal testing, either in the form of a pilot study or a split panel field test (DeMaio 1983). A pilot study employs “a design which duplicates the final proposed survey design on a small scale from beginning to end” (DeMaio 1983, p.57). The goal of the pilot study is to identify and correct problems prior to fielding the study. This is one of the methods that the Census Bureau uses to test and dry run decennial census operations prior to the decennial census. Sites are selected, usually to test specific aspects of the operation, and the entire operation is conducted within those sites. This was the method used to field test the CCM survey in 2006.

The other type of formal testing that DeMaio (1983) references is a split sample test. DeMaio notes that “split sample tests are suitable for any data collection mode” (p.75) and suggests that an RDD CATI study is a viable, relatively inexpensive option for a split panel design. This was the method used in the RDD study described here.

Our study compares data from a site test to data collected in an RDD split panel design.

2.2 Mode Effects: Telephone vs. Face-to-face

Most of the literature on mode effects seems to be focused on the effects a switch in mode has on survey estimates. The current study is particularly interested in choosing a mode for field testing a survey that will be fielded nationally as a face-to-face survey.

Differences between telephone and face-to-face surveys are well documented in the literature. Groves (1989) presents a thorough comparison of difference between telephone and face-to-face surveys, and several issues have arisen as a result of this and other comparisons, the most notable issues being response rates and data quality (see Biemer, 2001; de Leeuw, 1992; de Leeuw & van der Zouwen, 1988; Groves, 1989; and Schwarz, Strack, Hippler, & Bishop, 1991). De Leeuw and van der Zouwen (1988) conducted a meta-analysis of telephone and face-to-face interview comparisons. They found higher response rates from face-to-face interviews than from telephone interviews (75 percent versus 69 percent across studies). They also found several indicators of data quality that favored face-to-face interviews. Interestingly, de Leeuw and van der Zouwen provide evidence that the differences between face-to-face and telephone interviews were becoming less exaggerated over the years, indicating that more experience with telephone interviews was increasing their accuracy. Supporting that claim, Biemer (2001) investigated measurement and nonresponse bias in telephone and face-to-face interviews and found no consistent quality difference in favor of either mode. Jäckle, Roberts and Lynn (2006) more recently conducted a mode experiment and did not find overall higher item nonresponse rates in the telephone mode. In fact, Jäckle et al. only found one question, concerning household income, which produced large mode effects between the telephone and face-to-face samples.

Though we know that nonresponse is higher over the phone, nonresponse bias is not necessarily increased. Groves (2006) states that nonresponse bias is the correlation between a survey variable and the willingness to answer the survey question and that nonresponse rates are not proportional to nonresponse bias. Similarly, Montaquila and colleagues (2006) conclude that increased efforts to raise RDD response rates do not affect nonresponse bias.

Thus, past research suggests that the data quality for a telephone survey could be expected to be as good as that of a face-to-face interview.

2.3 Sampling Method: Address-based vs. Telephone frames

While we are interested primarily in the question of how to perform a field test for a large national survey, Jordan, Marcus and Reeder (1980) conducted a similar study that focused on a survey of Los Angeles residents. They found much

higher item nonresponse for a question on income in the telephone sample than they did in the face-to-face sample. Otherwise, their demographic variables produced essentially the same characteristics.

Jäckle, Roberts and Lynn (2006) used a single frame that contained both telephone and address information. Interestingly, they found differences between the people who were willing to be interviewed over the phone and those willing to be interviewed in person: fewer men, manual workers and respondents with low education were willing to be interviewed over the phone than were willing to be interviewed in person.

Telephone prevalence rates become a key consideration when deciding between an RDD and an address-based sample. Prior to recent years, the proportion of households with a landline telephone was on the increase (Brick & Tucker, 2007). However, with the boom in popularity of wireless phones, the trend is changing such that households with landline coverage are now decreasing. Blumberg and Luke (2007) reported that in the first half of 2007, 13.7 percent of American homes did not have a landline telephone, but did have at least one working wireless phone. About two percent of households had no telephone service during that time period. They also reported that over 25 percent of young adults (age 18-24) lived in households that had only a wireless phone (no landline), and more than half of all adults living with unrelated roommates were in wireless-only households. Wireless-only households were also more likely than those with landlines to be low income, to be renting, to consist of a person living alone or with unrelated roommates, and to be located in metropolitan areas. People living in a wireless-only household were more likely to be male, young, Hispanic, and living in the Midwest. These findings suggest there might be differences between people who are available to be sampled by RDD in banks of landline numbers versus an address-based frame. However, Keeter et al. (2007) posit that including wireless-only households in a study does not change estimates very much; indicating the coverage bias in estimates from RDD studies using banks of landline numbers is small.

Link et al. (2008) recently conducted a study involving a mailout survey that used an address-based frame and compared that to a traditional RDD sample. They compared the demographic characteristics of the mail and telephone samples to those of the Current Population Survey (CPS) and found that both samples reported higher education than the CPS estimates, but that the mail respondents' education exceeded that of telephone respondents. The telephone sample had higher proportions of white non-Hispanics, high income households, and married people. It also had lower rates of persons in households with three or more adults.

From this research, we expect that our pretesting methods (RDD and site test) will tap into different demographic groups, and in as much as demographics are related to our key area of interest (i.e., census coverage), our data may be affected.

3. Methodology

The data collected in the RDD CATI survey came from the 2006 Questionnaire Design and Experimental Research Survey (QDERS). QDERS is a split-panel controlled experiment developed by the Census Bureau's Statistical Research Division for conducting methodological experiments offline from the agency's ongoing production surveys. The purpose of the 2006 QDERS was to investigate a new way of measuring residence status according to census residence rules. Results of the larger residence study embedded in QDERS are reported in Childs and Nichols (forthcoming).

The data collected in the site-test CAPI survey came from the 2006 CCM Person Interview (PI), which was part of the 2006 Census Test. The purpose of the 2006 Census Test was to test the census and coverage measurement operations to be used in the 2010 Census. The specific purpose of the 2006 CCM PI was to collect an independent roster of people living in selected housing units and to obtain enough information to determine residence status for each of those people. In order to measure the within-household coverage of the census, data from the CCM PI is compared to data resulting from the census operations.

QDERS 2006 was conducted between November 3 and November 21, 2006, from one of the Census Bureau's centralized telephone centers. The sample was from a landline telephone number bank, selected to be nationally representative (excluding Alaska and Hawaii), with independent samples for each of the two treatments. The 2006 CCM PI was conducted using an area-based sampling frame between July 3 and September 2, 2006, in two sites: selected census tracts in Travis County, Texas (near Austin), and the Cheyenne River Sioux American Indian Reservation and Tribal Trust Lands in South Dakota.

Both QDERS 2006 and the 2006 CCM PI used the same basic survey questions and were developed in the Blaise survey design software. Both instruments first collected a roster of current occupants, then sought information about the demographics of those occupants, and finally asked questions about residence status. The two surveys had very similar questions for these three sections of the survey:

1. The first three roster questions collecting names of people who currently live at the housing unit in sample;²
2. Most demographic questions; and
3. Questions collecting alternate addresses (i.e., other addresses where a person lived or stayed during the year, and where they might be counted in the census).³

The final sections of QDERS 2006 and the 2006 CCM PI differed. One panel of QDERS 2006 used the same residence status questions as the 2006 CCM PI; but the other QDERS panel used a different set of residence status questions (and was the focus of the main split-panel experiment). In addition, the 2006 CCM PI also collected information on people who had moved out of the housing unit since Census Day, April 1st, whereas the QDERS 2006 did not collect this information. Data from the portions of the surveys that differed are not included in this analysis.

The total sample size for each QDERS panel was 2,996. Using the response rate calculation standards established by the American Association for Public Opinion Research (AAPOR, 2006), excluding cases of ineligibility and unknown eligibility, the response rates for Panels A and B were 60.77 percent and 55.92 percent, respectively.⁴ These response rates were significantly different from one another ($p < .01$). The overall response rate was 62.4 percent. A total of 1,870 interviews were completed: 982 in Panel A and 888 in Panel B.

The total sample size in the 2006 CCM PI was 5,468: 495 in SD and 4,973 in Texas. The overall response rate was 97.13 percent. A total of 4,609 interviews were either fully or partially completed. Because some of these were vacant or at places which were out of scope for CCM (e.g., group homes), for this paper we examine 4,590 interviews at occupied housing units.

Within the results section, we refer to the two tests as RDD (referring to the QDERS study) and site (referring to the CCM PI field test).

4. Limitations

Our analysis focuses on the performance of the questionnaire as opposed to making inferences about a population. One must also keep in mind the difference in how the respondents were selected. The site test was conducted in two relatively small geographical areas. The RDD test used a nationwide sample. There were too few RDD cases in the test sites to make a subset analysis worthwhile. Much of the discussion of differences within this paper is made only by a visual inspection of the response distributions.

We are performing an unweighted analysis even though the large blocks in the test sites were sampled at a higher rate than were the small and medium blocks. Since our focus is the questionnaire, the additional complication of using the sampling weights will not affect our conclusions. The RDD survey used simple random sampling. In addition, we calculate confidence intervals under the assumption of simple random sampling only when the statistics of interest from the two surveys appear close. Otherwise, the differences are large enough that inspection is adequate to draw conclusions about the questionnaire.

In addition to the difference in mode and universe, the interviewer's experience also differed by test. The RDD test used experienced telephone center staff, some with more than 10 years' experience, while the site test used recently hired, temporary field staff, many of whom had no previous data collection experience.

A final limitation concerns the site test data. Address identifiers used to define uniqueness were assigned in post processing. Due to limitations in the software, a small percentage of addresses may have been classified as duplicates or unique when that was not true. We do not believe this limitation is severe enough to hamper the conclusions presented below.

5. Results

We examine person demographics and the mobility of the populations (the substantive questions of the survey). For more analysis see Nichols, Childs and Linse (forthcoming).

² Results reported within this paper include only the people identified with one of the first three roster questions. There were only three roster questions in the RDD study and these three questions matched the first three questions in the site test. The site test included three additional questions. People identified by one of these additional questions were excluded from this analysis.

³ The site test used a shorter reference period for several of the alternate address questions than the RDD study, but the "catch-all" question had a similar reference period, thus the results should be similar.

⁴ Rates reflect the AAPOR RR6 definition (AAPOR, 2006).

5.1 Demographics

First we compare the demographic data between studies to see if there are differences between the samples studied and to see how those differences relate to national statistics. For each table, data from the nationally representative RDD study are presented first, followed by data from the personal-visit site test and relevant percentages from Census 2000.⁵

In Table 1, we see very different race/Hispanic origin distributions between the two sample populations (RDD and Site) and the census. Our site test was held in and near Austin, Texas and on an Indian Reservation in South Dakota – one area was highly Hispanic and the other area was highly American Indian. These sites were chosen because they contained hard-to-enumerate populations. For those reasons the race/Hispanic distributions from the site test are almost the reverse of what was reported from Census 2000. This is an example of how the particular sites selected for a pretest can impact the sample characteristics. The RDD test seems to have a higher proportion of white/non-Hispanic population than the general census, consistent with previous research findings with RDD studies (Link et al., 2008)

Table 1: Race/Hispanic Origin Distribution

	RDD		Site		Census 2000
White Non-Hispanic	3774	77.1%	4056	35.0%	69%
Other ⁶	1121	22.9%	7549	65.1%	31%
Total # People	4,895		11,605		281,421,906

In Table 2, we see the disadvantages of both the RDD and site test in the representation of age groups proportional to the population. The RDD appears to have a lower percentage of people under 34 compared to the census. This is supported by Blumberg and Luke's (2007) finding that over a quarter of young adults were in wireless households (and thus were not available by RDD). The site test seems to have a higher proportion of young adults than does the national average. The main campus of the University of Texas (one of the largest universities in the U.S.) is located in Austin, the location of one of the test sites. We speculate that this factor drove up the college-age population in the site test. This is another example of how the particular site selected may influence the sample.

Table 2: Age Distributions

	RDD		Site		Census 2000
<18	1131	23.1%	2759	23.8%	26%
18-34	738	15.1%	4258	36.7%	24%
35-54	1447	29.6%	2734	23.6%	29%
55-64	729	14.9%	698	6.0%	9%
65+	776	15.9%	714	6.2%	12%
Item nonresponse	74	1.5%	442	3.8%	
Total # People	4,895		11,605		281,421,906

5.2 Mobility of Populations

Next, we will examine the data of substantive interest to this survey. These data were only available in a comparable way through the pretests using the RDD and the site samples, so there will be no census benchmark. These data pertain to the mobility of the current occupants of the sampled unit. By mobility, we mean that some people live or stay at more than one place and thus are mobile.

The Census Bureau needs to (1) count everyone once and only once (i.e., no one should be counted twice, or duplicated, and no one should be missed, or omitted) and (2) count each person at the correct place according to the census residence rule.⁷ In order to fulfill the mission of getting each person counted once and only once and in the correct place, the site test and RDD data collections attempted to elicit a complete roster of everyone living at the sampled address, as well as those people who stay at the unit more than any other place or who have no one place where they stay the most. These latter two groups of people are the types of people who could be missed in the census.

Both studies collected the address where a person lived on Census Day, April 1, 2006 (if it was different from the sampled address) and any other place the person stayed often (e.g., college, a relative's home, a location often stayed at because of travel for work or military service, or a second home). These are the places where a person could be duplicated in the census. We refer to all the different addresses (other than the sampled address) as alternate addresses. See Nichols et al.

⁵ These data are gathered from American Fact Finder available on the web at: <http://factfinder.census.gov/>.

⁶ Don't know and Refused are included in the Other category.

⁷ See the report from the National Research Council (2006) on the complexities of applying this rule.

(forthcoming) for the questions used to elicit these alternate addresses for both data collections.⁸ One difference between the studies was that several of the alternate address questions used a shorter reference period in the site test than did the RDD test. In the site test, the questions about college, relative, job and military addresses asked about alternate addresses in March and April. In the RDD test, these questions asked about the year up until the study date in November. However, the question on seasonal residences and the final “catch-all” question used the same reference period. Because the catch-all questions used the same time frame (the year up until that point), we hypothesize that the distributions of addresses collected should be similar.

Using the automated instrument assisted in the efficient collection of alternate addresses. As might be suspected, multiple people within the household sometimes have the same alternate address. For example, John and Jane Doe living at 101 Main Street, Anytown U.S. have a second home at the beach where they spend the summer. They share that second address. Instead of having to retype the address for each person in the household, both instruments (RDD and site) allowed the interviewer to pick from a list of previously recorded addresses for that interview. For example, once the beach house was reported for John, when the interviewer asked about Jane’s second home, the interviewer could select the address for the beach house that was already entered. Of course, the interviewer could also add a new address, if needed. Within this paper, when we evaluate only the unique addresses within a household we are talking about John and Jane’s one beach house (we will not count it twice) but when we examine the number of alternate addresses per person, both John and Jane each have one alternate address.

The first mobility indicator in this study involves the questions used to identify the current occupants of the sampled address. The questions identify core members of the household currently living at the sampled address and anyone else who might be considered a resident according to the census residence rule. The introduction and first three questions (which we refer to as roster questions) were identical in both data collections. Figure 1 provides the question text.

Introduction: Now I will ask some questions about where you live. We’ll start by making a list of people. We want to include people who live there all the time and people who stay there often, even if you don’t think of them as members of your household.

Roster Question 1: First, please tell me the names of everyone who lives there now. Let’s start with you. What is your name?
Anyone else?

Roster Question 2: Is there anyone else who has another place to live but who stays there often?
What is that person’s name?
Anyone else?

Roster Question 3: Is there anyone else who is staying there until they find a place to live?
What is that person’s name?
Anyone else?

Figure 1: Roster question introduction and text for both RDD and site test, except that the site test uses “here” instead of “there” for these questions when interview is at the sample address.

Table 3 shows that the distribution of people identified in each study using each of the three roster questions looked nearly identical.

Table 3: Distribution of people by the roster question where they were first mentioned

	RDD		Site	
Roster Question 1	4745	96.9%	11252	97.0%
Roster Question 2	134	2.7%	307	2.7%
Roster Question 3	16	0.3%	46	0.4%
Total # People	4,895		11,605	

The majority of people were identified with the first question that asked who lives at the address now. About three percent were identified with the remaining two probes. The last two probes tried to elicit non-core members of the household,

⁸Panel 2 of the RDD data collection and the site test used the same initial question about where a person lived on April 1st. Panel 1 of the RDD had a slightly different initial question. We will note when we excluded Panel 1 responses from the analysis.

who either had another place to live but stayed at the address often, or had no place to live and were staying at the address temporarily. If we assume that the respondent answered the latter two questions accurately, this three percent of people are more mobile than are the people identified with the first question. This suggests our two responding populations (RDD and site) were not very different from each other in terms of mobility as measured by these questions.

The second mobility indicator addressed in this study is how many other unique addresses each person had. People living or staying often at the sampled unit might have a second (or third, etc.) address where they could have been counted in the census. Table 4 shows the distributions of alternate addresses collected for each person identified during the data collection and the 95 percent confidence interval (CI) surrounding that percent as calculated by the “exact binomial” option in SAS’ Proc Freq.

Table 4: Number of alternate addresses collected per person

	RDD		RDD CI	Site		Site CI
No alternate addresses	3982	81.4%	(.8023 - .8243)	9796	84.4%	(.8374 - .8507)
One alternate address	807	16.5%	(.1546 - .1756)	1648	14.2%	(.1357 - .1485)
Two alternate addresses	99	2.0%	(.0165 - .0246)	148	1.3%	(.0108 - .0150)
Three alternate addresses	6	0.1%	(.0004 - .0027)	12	0.1%	(.0005 - .0018)
Four alternate addresses	1	<0.1%	(.0000 - .0011)	1	<0.1%	(.0000 - .0005)
Total # People	4,895			11,605		

Table 4 shows that for 81 percent of the people identified in the RDD study, no other address where that person might have lived or stayed during 2006 was mentioned during the interview, and that 84 percent of the people in the site test did not have any other address where they lived or stayed during 2006. No one in either data collection had more than four alternate addresses. Despite the fact that the two tests used slightly different reference periods for some of the alternate address questions, the resulting number of addresses collected were very similar. Of course, due to the large sample sizes, even these seemingly small differences become significantly different as shown by the confidence intervals.

Table 5 shows that in 71 percent of households in the RDD study and in 73 percent of the households in the site test, no other addresses were reported for any household member. Again, these percentages appear very similar despite the slight differences in reference periods.

Table 5: Distribution of Unique Alternate Addresses in an Interview

Number of Alternate Addresses per HH	RDD		Site	
0	1330	71.1%	3333	72.6%
1	412	22.0%	960	20.9%
2	105	5.6%	216	4.7%
3	16	0.9%	53	1.2%
4	6	0.3%	24	0.5%
5	0		2	<0.1%
6	0		1	<0.1%
7	0		1	<0.1%
8	1	0.1%	0	
Total # HH	1,870		4,590	

Table 5 tell us that about 30 percent of RDD and site test households have at least one person with mobility, but that does not translate into 30 percent of people having mobility. In Table 4, we see that less than 20 percent of people have mobility. This means either the non-mobile households contain more people than the mobile households, or not everyone in the households with mobile people had another address. We did not investigate either of these possibilities further for this paper.

6. Conclusions

Finally, we focus on what we learned by conducting the two pretest studies and what we would not have learned if we had only conducted one of them. As a reminder, the pretest studies were conducted in a very similar time frame (site: summer 2006 and RDD: November 2006) and used very similar questions for a portion of the surveys. The RDD study, however, had only a 62 percent response rate compared to the 97 percent response rate for the site test.

We suspected there would be different demographics in the two studies, and there were. Differences generally matched up to the literature on the different modes. If we were interested in using these demographics for estimation or some

other decision making purpose, neither method would have been ideal. However, for the purpose of a pretest, the studies showed remarkably similar results on the substantive topics.

We suspected that we would find more differences in the mobility of the two studied populations than we found. In fact, even though there were major demographic differences, the mobility of the populations in the two studies seemed relatively stable. Based on these data, either pretest method seemed to be a reasonable venue for testing questions to capture mobility and to make decisions regarding the mobility found. This is consistent with Montaquila and colleagues' (2006) finding that increased efforts to raise RDD response rates do not affect nonresponse bias. We suggest further analysis on the RDD and site test data to see if an earlier closeout date (and more nonresponse) would yield similar distributions regarding mobility as those we found with the full dataset.

These findings provide a starting place to examine the advantages and disadvantages of pretesting a national face-to-face survey using two different sampling methods. They also provide a starting point from which to make decisions regarding how to study mobility within the U.S. population.

Acknowledgements

The authors would like to thank Mary Mulry, Betty Murphy, Terry DeMaio, Gia Donnalley, Magda Ramos, and Stephanie Sheffield for their reviews of the paper and Rolando Rodriguez and Aref Dajani for statistical and programming support.

References

- American Association for Public Opinion Research (2006). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. 4th edition. Lenexa, Kansas: AAPOR.
- Biemer, P. (2001). Nonresponse Bias and Measurement Bias in a Comparison of Face to Face and Telephone Interviewing. *Journal of Official Statistics*, 17, 295-320.
- Blumberg, S. J. and Luke, J. V. (2007). Wireless Substitution: Early Release of Estimates from the National Health Interview Survey, January – June 2007.
- Brick, M. J. and Tucker, C. (2007). Mitofsky-Waksberg: Learning from the Past. *Public Opinion Quarterly*, 71, 5, 703-716.
- Childs, J.H. and Nichols, E. (forthcoming). "QDERS 2006: An Experimental Comparison on Measuring Residence Status." RSM2008/x, <http://www.census.gov/srd/www/byname.html>
- De Leeuw, E. D. (1992). Data Quality in Mail, Telephone, and Face-to-Face Surveys. Amsterdam: TT-Publicaties.
- De Leeuw, E. D. and van der Zouwen, J. (1988). Data Quality in Telephone and Face to Face Surveys: A Comparative Meta-Analysis. *Telephone Survey Methodology*, 283-299.
- DeMaio, T. J. (ed.) (1983). Approaches to Developing Questionnaires. Statistical Policy Working Paper 10. Washington, DC: Office of Management and Budget.
- Groves, R. M. (1989). Survey Errors and Survey Costs. New York: Wiley.
- Groves, R. M. (2006). Nonresponse Rates and Nonresponse Bias in Household Surveys. *Public Opinion Quarterly*, 70, 5, 646-675.
- Jäckle, A., Roberts, C., and Lynn, P. (2006). *Telephone Versus Face-to-Face Interviewing: Mode Effects on Data Quality and Likely Causes. Report on Phase II of the ESS-Gallup Mixed Mode Methodology Project*. ISER Working Paper.
- Jordan, L. A., Marcus, A.C., and Reeder, L.G. (1980). Response Styles in Telephone and Household Interviewing: A Field Experiment. *Public Opinion Quarterly*, 44, 2, 210-222.
- Keeter, S., Kennedy, C., Clark, A., Tompson, T., and Mokrzycki, M. (2007). What's Missing from National Landline RDD Surveys?: The Impact of the Growing Cell-Only Population. *Public Opinion Quarterly*, 71, 5, 772-792.
- Link, M. W., Battaglia, M. P., Frankel, M. R., Osborn, L., and Mokdad, A. H. (2008). A Comparison of Address-Based Sampling (ABS) Versus Random-Digit Dialing (RDD) for General Population Surveys. *Public Opinion Quarterly*, 72, 1, 6-27.
- Montaquila, J. M., Brick, J. M., Hagedorn, M. C., Kennedy, C., and Keeter, S. (2006). Aspects of Nonresponse Bias in RDD Telephone Surveys. *Second International Conference on Telephone Survey Methodology, January, Miami, FL*.
- National Research Council (2006). *Once, Only Once, and in the Right Place: Residence Rules in the Decennial Census*. Panel on Residence Rules in the Decennial Census. Daniel L. Cork and Paul R. Voss (Eds.) Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.
- Nichols, E., Childs, J. H., and Linse, K. (forthcoming) RDD versus Site Test: Mode Effects on Gathering a Household Roster and Alternate Addresses RSM2008/x, <http://www.census.gov/srd/www/byname.html>
- Schwarz, N., Strack, F., Hippler, H. J., and Bishop, G. (1991). The Impact of Administration Mode on Response Effects in Survey Measurement. *Applied Cognitive Psychology*, 5, 193-212