

STUDY SERIES
(*Survey Methodology* #2010-06)

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

Elizabeth Nichols
Jennifer Hunter Childs
Kyra Linse¹

¹Decennial Statistical Studies Division

Statistical Research Division
U.S. Census Bureau
Washington, D.C. 20233

Report Issued: April 19, 2010

Disclaimer: This report is released to inform interested parties of research and to encourage discussion. The views expressed are those of the authors and not necessarily those of the U.S. Census Bureau.

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. In July, staff field tested a computer-assisted personal interview (CAPI) version of the survey instrument 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 basically the same set of questions to gather a household roster and to capture other addresses where people in that household were reported to have stayed during that 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. When we examine the trends in the demographic distributions of the two studies, we highlight how the findings follow the patterns expected for RDD and site test data based on the literature. To our knowledge, comparative literature between RDD and site test data for the main variables of interest in our studies (i.e., household mobility and the identification of people with multiple addresses) does not exist. This paper provides a starting place to look at the patterns of mobility and multiple addresses for these two field pretest methods. The results suggest that using a RDD method to study mobility and multiple residences might be comparable to using a site test.

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

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

I. 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 U.S. Census Bureau usually uses face-to-face site tests to pretest decennial census operations. We seek to compare what data can be gathered about living situations from pretests conducted using either of these testing methods.

In 2006, the 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 the site test was fielded as a part of the 2006 Census Test, which included 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 purpose of the 2006 site test was to test the questionnaire and operational aspects of the CCM in preparation for 2010. From July to September, 2006, the CCM face-to-face survey was field-tested. 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.

The survey instrument for the RDD pretest was a specially designed research survey. In November of 2006, many of the same questions that were asked in the CCM test were also asked in a small-scale nationally-representative (excluding Alaska and Hawaii) RDD survey. The RDD sample was 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 essentially the same set of questions (and the same question order) to gather a household roster and associated demographics of the people in the household, to capture reports of 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. We compare these two field pretest surveys in this paper.

It is important to understand the differences between the two field pretesting methods.

- **Location:** At the most general level, the RDD population varied from the site test population by where they were geographically located in the U.S. The RDD population resided across the contiguous U.S., and the site-test population resided in one of the two site-test locations.
- **Telephone access:** Cell-phone-only households were in the site test population, but not in the RDD population. We did not measure the percent of people who have land-line phone access in either study.
- **Interviewers:** The interviewers used in each study differed. The RDD test used permanent, experienced staff and the site test used temporary staff.

- Mode: The questionnaire mode also differed. Both questionnaires used automated instruments; but the RDD was conducted over the telephone (CATI), and the site test was conducted via personal visit (CAPI).

We sought to investigate whether these differences would translate into differences in the substantive survey data items (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, looking for patterns. We discuss how the trends in the demographic data in each field study follow the expected patterns based on our knowledge of the frame and previous research. Additionally, we discuss the patterns in each field study associated with mobility of people and the number of addresses where a person could have been counted. These data items had never before been collected in an RDD study at the U.S. Census Bureau.

Before we go into details on the studies 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.

II. Background

Pretesting

In the pretesting cycle, after initial cognitive testing comes field 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 a pilot study is to identify and correct problems prior to fielding the full survey or census. This is one of the methods that the Census Bureau uses to test 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) mentions 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.

In this paper, we compare data from a site test to data collected in an RDD split panel design.

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 conducted 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. Several issues have arisen as a result of Groves' 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.

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) more recently conducted a mode experiment and did not find overall higher item nonresponse rates in the telephone mode. In fact, Jäckle and colleagues only found one question, again concerning household income, which produced large mode effects between the telephone and face-to-face samples.

Jäckle and colleagues (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.

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 necessarily 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. To our knowledge, the literature does not address whether or not there is a correlation between our information of interest (i.e., whether people have more than one place to live) and willingness to answer survey questions. Consequently, this

study seeks to compare test data from an RDD study (where response rate is typically much lower) with face-to-face site test data (for which response is typically very high).

Sampling Method: Address-based vs. Telephone frames

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 land-line 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 land-line numbers is small.

Link and colleagues (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 significantly higher education than the CPS estimates. The telephone sample had higher proportions of white non-Hispanics and married people compared to CPS. It also had lower rates of persons in households with three or more adults and persons in households with no children compared to CPS.

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., mobility and alternate places to live), our data may be affected.

III. 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 to conduct 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 are reported in Childs, Nichols, Dajani, and Rothgeb (2007). 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 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 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 are compared to data resulting from the census operations. 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;¹
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).²

The final sections of QDERS 2006 and the 2006 CCM PI differed. Data from the portions of the surveys that differed are not included in this analysis.

1. 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. (This was the focus of the main split-panel experiment.)
2. 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.

¹ 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 roster questions. People identified by one of these additional questions were excluded from this analysis.

² The site test used a shorter reference period for several of the alternate address questions than the RDD study, but since the "catch-all" question had a similar reference period, the results should be similar.

While the total sample size was between 5,000 and 6,000 for both operations, Table 1 shows that the response rate and number of cases eligible for analysis differed.

Table 1: Sample Size and Response Rates

	<u>2006 QDERS</u>	<u>2006 CCM PI</u>
Total Sample Size	5,992	5,468
Response Rate	58.4%	97.1%
Cases Eligible for analysis	1,870	4,590

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.³ These response rates were significantly different from one another ($p < .01$). The overall response rate was 58.4 percent as shown in Table 1. 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 South Dakota 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” (the QDERS study) and “site” (the CCM PI field test).

IV. Limitations

Although we present demographic population distributions, our primary focus is on the distribution patterns of the mobility data elicited by the questionnaire in the two field tests. The discussion of differences within this paper is made only by a visual inspection of the response distribution patterns because of the following issues:

- 1) We assume ignorable nonresponse bias in the RDD data. Montaquila et al. (2006) concluded that increased efforts to raise RDD response rates do not affect nonresponse bias. In our study, the RDD’s response rate was 58 percent and the response rate for the site test was 97 percent. We do not have information to determine whether there is bias associated with the data we would have liked to collect from the RDD nonresponders as compared to those who responded in that test.
- 2) The test frames and sampling methods were different. The site test was conducted in two relatively small geographical areas. The RDD test used a nationwide simple-random sample. There were too few RDD cases in the test sites to make a geographic subset analysis worthwhile. Additionally, we did not have access to the weights for

³ Rates reflect the AAPOR RR6 definition (AAPOR, 2006).

the site test data. Large blocks in the test sites were sampled at a higher rate than were the small and medium blocks.

- 3) In addition to the difference in mode and universe, the interviewers differed by field 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 mistakenly classified as duplicates or unique. We do not believe this limitation is severe enough to hamper the conclusions presented below.

V. Results

We examine three aspects of the data: 1) demographics (both person and household) to compare the samples, 2) the mobility of the populations (the substantive questions of the survey), and 3) item nonresponse.

Person Demographics

First we compare the differences in demographic data between the test samples and examine 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 2000.⁴

In Table 2, the sex distribution for all people rostered in both the RDD and site test data collections were within 5 percent of the population distribution in Census 2000.

Table 2: Sex Distribution

	Number of People				Census 2000
	RDD		Site		
Male	2339	47.8%	5982	51.6%	49%
Female	2549	52.1%	5583	48.1%	51%
Item nonresponse	7	0.1%	40	0.3%	
Total # People	4,895		11,605		281,421,906

In Table 3, we see very different race/Hispanic origin distributions between the two sample populations (RDD and site) and the census. We did not expect the site test to reflect the general race/Hispanic origin demographics of the country. As mentioned earlier, the 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 were almost the reverse of what was reported in Census 2000. The RDD test had an

⁴ Census 2000 data are from American Fact Finder, available on the Internet at: <http://factfinder.census.gov/>.

eight percent higher white/non-Hispanic population than Census 2000. Link and colleagues (2008) also found his RDD study to have a higher proportion of White/non-Hispanic than the population, estimated by the CPS. Thus, our finding is consistent with previous research.

Table 3: Race/Hispanic Origin Distribution

	Number of People				
	RDD		Site		Census 2000
White Non-Hispanic	3774	77.1%	4056	35.0%	69%
Other ⁵	1121	22.9%	7549	65.1%	31%
Total # People	4,895		11,605		281,421,906

Finally, in Table 4, we see the disadvantages of both the RDD and site tests in the representation of age groups proportional to the population. The RDD had a lower percentage of people under 34 compared to Census 2000. This is consistent with Blumberg and Luke’s (2007) finding that over a quarter of young adults were in wireless households (and thus not available by RDD). The site test had a higher sample proportion of young adults (18-34) than 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.

Table 4: Age Distributions

	Number of People				
	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

Household Demographics

Next we will examine household level variables, again showing the sample distributions across the three data collections.

Table 5 shows that there was a smaller percentage of one-person households in the RDD survey and a higher percentage in the site test compared to Census 2000. These differences might be attributable to the less reachable wireless-only population for the RDD test, which tends to be made up of people living alone or with unrelated people (Blumberg and Luke, 2007). The fact that one of the sites was near the University of Texas also could provide an explanation for the increase in the number of single resident households in the site test.

⁵ Don’t know and Refused are included in the Other category.

Table 5: Household (HH) Size Distributions

No. of people	Number of Households				
	RDD		Site		Census 2000
1 person HH	407	21.8%	1408	30.7%	26%
2	670	35.8%	1375	30.0%	33%
3	323	17.3%	732	16.0%	17%
4	277	14.8%	560	12.2%	14%
5	122	6.5%	286	6.2%	7%
6	51	2.7%	123	2.7%	2%
7 or more	20	1.1%	106	2.3%	2%
Total # HH	1,870		4,590		281,421,906

Table 6 shows that the RDD study had a higher percentage of owners compared to the general population, which might again be attributed to the exclusion of wireless-only households, since past research suggests that they tend to be renters (Blumberg and Luke, 2007). Over half of the households in the site test were renters which, again, might be due to the university and characteristics of the college-age population.

Table 6: Household Tenure Distribution

	Number of Households				
	RDD		Site		Census 2000
Owner	1510	80.8%	1915	41.7%	66%
Renter	322	17.2%	2554	55.4%	34%
Occupied rent free	26	1.4%	65	1.4%	Not calculated
Item nonresponse	12	0.6%	66	1.4%	Not calculated
Total # HH	1,870		4,590		281,421,906

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 was no population benchmark. These data pertain to the mobility of the current occupants of the sampled unit. In this paper, we define mobility to mean the percentage of people living or staying at multiple addresses.

The Census Bureau needs to (1) count each person 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.⁶ 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

⁶ See the report from the National Research Council (2006) on the complexities of applying this rule.

where they stay the most. Members of these latter two groups of people 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. The appendix contains the questions used to elicit these alternate addresses for both data collections.⁷ 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 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 beach house address 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.

⁷Panel 2 of the RDD data collection used the same initial question about where a person lived on April 1st as the Site test. Panel 1 of the RDD had a slightly different initial question. We will note when we excluded Panel 1 responses from the analysis.

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.⁸

Table 7 shows that the distribution of people identified in each study using each of the three roster questions looked nearly identical.

Table 7: Distribution of people by the roster question where they were first mentioned

	Number of People			
	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, 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 those identified with the first question. This pattern in the data 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 8 shows the distributions of alternate addresses collected for each person identified during the data collection.

⁸ The site test uses “here” instead of “there” for these questions when the interview is at the sample address.

Table 8: Number of alternate addresses collected per person

	Number of People			
	RDD		Site	
No alternate addresses	3982	81.4%	9796	84.4%
One alternate address	807	16.5%	1648	14.2%
Two alternate addresses	99	2.0%	148	1.3%
Three alternate addresses	6	0.1%	12	0.1%
Four alternate addresses	1	<0.1%	1	<0.1%
Total # People	4,895		11,605	

Table 8 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 distribution pattern of collected addresses appears similar.

Table 9 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 and populations.

Table 9: Distribution of Unique Alternate Addresses in an Interview

Number of Alternate Addresses per HH	Number of Households			
	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 9 shows that about 30 percent of RDD and site test households had at least one person with mobility, but that did not translate into 30 percent of people having mobility. In Table 8, we see that less than 20 percent of people had mobility. This could mean either the non-mobile households contained more people than the mobile households, or there were some non-mobile people living in households with mobile people. We did not investigate either of these possibilities further for this paper.

Table 10 examines the point in the interview when alternate addresses were mentioned for each person. For this analysis, we will only examine one panel of the RDD study, since it had the same initial question as the site-test.⁹ There were 888 households in the panel of interest in the RDD study, and there are 4,590 households in the site-test. So, the second panel of the RDD is about one-fifth the size of the site-test.

The appendix contains the questions in the two studies that were used to elicit alternate addresses. The initial alternate address question was identical in the site test and Panel 2 of the RDD. That question asked for the address where the person was living or staying on April 1st. The “College,” “Seasonal,” and “Often” alternate address questions were virtually identical between studies. The reference period for the RDD and site test differed in the question text for the “Military,” “Job,” and “Relative” alternate address questions. The reference period for the RDD was the year, while the reference date for the site test was March or April. The RDD also had two additional questions that the site test did not have and vice versa.

Table 10: Question which elicited the alternate addresses
Number of Households with an Alternate Address
(Percentage of all Alternative Addresses First Reported Here)

Alternate address Question Type	RDD (Panel 2)		Site	
Initial Question	195	37.7%	1011	51.0%
Other alternate address questions with the same reference period*	221	42.7%	523	26.4%
Other alternate address questions with different reference period**	81	15.7%	164	8.3%
RDD: unique questions	15	2.9%	N/A	
Site: unique questions	N/A		247	12.4%
Group quarters	5	0.9%	39	2.0%
Total alternate addresses	517		1,984	

* includes College, Seasonal, and Often alternate address questions

** includes Military, Job, and Relative alternate address questions

Whereas the numbers of alternate addresses looked very similar across the two data collections in Tables 8 and 9, the question at which the address was first identified for the person looks different as shown in Table 10. The initial question, though identical between the panels compared here, elicited 51 percent of the alternate addresses in the site test compared to 38 percent in the RDD study. We speculate this could be for several reasons: first, the site test had a higher proportion of renters, which could impact the number of people that had moved since April 1st, and second, the site test was in a college town, which could increase the proportion of people who moved between April and the summer (due to the change in semester).

While small numbers were seen in both cases, in the site test the Group Quarters question elicited over twice the percent of reports as the RDD study (2% to 0.9%). A flashcard was used

⁹ There were 1027 alternate addresses mentioned for all people across both panels and about half of them (517) were collected in the panel of interest, Panel 2.

with this question in the site test and might be part of the reason for the higher percentage due to a better understanding of the types of places included in this question.

Item Nonresponse

Our final analysis looks at the item nonresponse or completeness of the data collected during the two tests. In a meta-analysis of telephone versus face-to-face studies, de Leeuw and van der Zouwen (1988) reported a slight mode effect of more “do not know” and “no answer responses” in a telephone survey than in a face-to-face interview. We investigated whether those findings would be replicated in our data collection.

First, we examine the residence status assigned by the automated processing system. For this paper, we collapsed the codes into three categories that would be meaningful for a general audience: 1) The person is a resident of the sampled unit on Census Day (and thus should have been counted there), 2) The person is not a resident of the sampled unit on Census Day (and should not have been counted there), or 3) The analyst could not determine whether the person should or should not have been counted there. The same codes were used in both studies.

Table 11 contains the distribution of residence status codes by study. The most surprising finding is the percent of unresolved residence status codes in the two tests. The site test had about 22 percent of people with an unresolved residence code compared to about 4 percent in the RDD study. We speculate this happened because the site test had a much higher response rate than did the RDD study (97% to 58%). In the site test, if no one answered the door after a number of attempts, the interviewer was allowed to conduct a proxy interview with a neighbor. The ability to conduct proxy interviews in the site test and the fact that the face-to-face interviewers were persistent in converting refusals might have led to more inconsistent data (causing an unresolved status) or higher proportions of missing data than in the RDD study, where the respondents who answered the phone and participated were less reluctant to report on their households.

Table 11: Distribution of People by Residence Status

	RDD		Site	
Resident of the Sample Address on Census Day, (April 1, 2006)	4282	87.5%	8774	75.6%
Not a Resident of the Sample Address on Census Day, (April 1, 2006)	435	8.9%	319	2.8%
<u>Undetermined Residence Status</u>	178	3.6%	2512	21.7%
Total # People	4,895		11,605	

In Table 12, we examine the completeness of the alternate addresses given by the respondent during the interview. Collecting exact addresses is critical to this survey because we not only need to determine where the person should be counted, but must also verify that the person was not counted more than once. Staff at the Census Bureau must be able to clearly identify any

other address where the person might have been counted. Over 60 percent of the alternate addresses were complete, or were just missing a ZIP Code in the site test compared to 31 percent in the RDD study.

Table 12: Item nonresponse: Alternate Addresses

	RDD		Site	
	Count	Percentage	Count	Percentage
All address fields have values	181	25.8%	930	55.7%
All except zip	37	5.3%	79	4.7%
Street, city and state, but no house number	45	6.4%	69	4.1%
City and state only	225	32.1%	261	15.6%
State only	39	5.7%	25	1.5%
Other combination of address info	158	22.5%	246	14.7%
<u>No address info: Item nonresponse</u>	<u>16</u>	<u>2.3%</u>	<u>60</u>	<u>3.6%</u>
Total # of unique alternate addresses	701		1,670	

In Table 13, we examine the completeness of people’s names given during the interview. In the site test, 98 percent of names provided were full names, while only 77 percent of names in the RDD study were full names. Collecting complete names is also critical to this survey to ensure that everyone is counted once and only once in the census. The item nonresponse patterns presented in Tables 12 and 13 are similar, with more item nonresponse in the RDD study than in the site test for these data items.

Table 13: Item nonresponse: Names of Household Members

	RDD		Site	
	Count	Percentage	Count	Percentage
First, Middle and Last name or First and Last	3773	77.1%	11377	98.0%
First name only	186	3.8%	130	1.1%
Last name only	N/A		13	0.1%
<u>No name given</u>	<u>936</u>	<u>19.1%</u>	<u>85</u>	<u>0.7%</u>
Total # People	4,895		11,605	

We speculate that there might be more privacy concerns with the RDD study than in a face-to-face interview as well. In the site test, the interviewer can show a badge, and has a Census Bureau bag and other authenticating materials. It is more difficult to verify the authenticity of the caller in an RDD environment.

Our item nonresponse results for names are somewhat confounded with differences in training and how names were collected in the two instruments, which could also have produced the differences that we observed. Although the training manuals were the same, the training script was not verbatim. We suspected an anecdotal discussion in one of the training sessions for one of the RDD panels may have accounted for the greater percentage of missing names in that panel (i.e., interviewers were more inclined to offer to respondents the option of not providing names in order to secure the interview). This hypothesis was supported by examining the completeness of names between Panel 1 and 2 in the RDD study (not shown in this paper). We found many more incomplete names in Panel 2, where the aforementioned anecdotal discussion occurred.

Additionally, the importance of gathering names was not emphasized in the training for the RDD study like it was in the site test training. In the RDD study, if a respondent did not give a first name, the instrument immediately allowed the interviewer to collect a description of the person to be used during the interview. In the site test, if the respondent did not give a first name, the instrument still asked for last name, before skipping to the write-in field for a description. In the production survey, without a name, it is very difficult to assess the accuracy of the census.

VI. Discussion

In this section we summarize what we learned by conducting the two pretest studies. As a reminder, the two pretest studies were conducted in a similar time frame (site: Summer 2006 and RDD: November 2006) and used very similar questions for a portion of the surveys. However, the sampling frames, interviewers, and mode were different between the two studies. The resulting response rate also differed, with the RDD study having only a 58 percent response rate compared to the 97 percent response rate for the site test. We have assumed any nonresponse bias in the resulting data was ignorable.

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 and the high minority populations in the site test. If we were interested in using the demographics gathered in these pretests for estimation or some other decision making purpose, neither method would have been ideal because neither method matched the general population very well.

We did not find any literature on whether we should expect differences between the studies in the mobility of the populations. Because the populations differed, 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 patterns for the numbers of alternate addresses elicited and the percentages of households with a person with an alternate address looked remarkably similar. Different questions might have been more productive in eliciting the alternate addresses in the two studies, but based on these data, either pretest method seemed to be a reasonable venue for testing questions to capture mobility and for making decisions regarding the mobility found. This is consistent with Montaquila et al.'s (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.

We found that the patterns of item nonresponse, specifically for name, addresses, and residence coding in the RDD study did not reflect what the Census Bureau would expect in a production setting like the site test. We suspected more difficulty collecting personally identifiable data (such as names and addresses) over the telephone than in a face-to-face survey. Our data supported that hypothesis. Surprisingly, the RDD collected enough data to assign a residence status code in a higher proportion of interviewed cases than the site test. However, because of the higher unit nonresponse, the RDD study left a higher proportion cases in the total sample unresolved (as non-interviews). For this reason, we believe the site test results show a more accurate picture of how many people would be left unresolved in a production face-to-face

survey. Had we only conducted an RDD pretest, we would not have had an accurate picture of the expected quality for these data items. Consequently, we may have underestimated the need for clerical coding or follow-up in the production survey.

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. The findings suggest that using a RDD study is a reasonable method for pretesting questions regarding mobility within the U.S. population.

Acknowledgements

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

Some of the analyses in this paper were presented by Jennifer Hunter Childs at the 2008 AAPOR meetings and can be found in the JSM proceedings (Nichols et al., 2008).

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. & Luke, J. V. (2007). Wireless Substitution: Early Release of Estimates from the National Health Interview Survey, January – June 2007.
- Brick, M. J. & Tucker, C. (2007). Mitofsky-Waksberg: Learning from the Past. *Public Opinion Quarterly*, 71, 5, 703-716.
- Childs, J.H., Nichols, E., Dajani, A., and Rothgeb, J. (2007.) A new approach to measuring residence status, *2007 Proceedings of the American Statistical Association, Survey Research Methods Section [CD-ROM]*, Alexandria, VA: American Statistical Association.
- De Leeuw, E. D. (1992). *Data Quality in Mail, Telephone, and Face-to-Face Surveys*. Amsterdam: TT-Publicaties.
- De Leeuw, E. D. & 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., & 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., & 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., & 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., & 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., & Linse, K. (2008). RDD versus site test: Mode effects on gathering a household roster and alternate addresses. In *JSM Proceedings, Survey Research Methods Section*. Alexandria, VA: American Statistical Association. Retrieved July 1, 2009 from <http://www.amstat.org/sections/SRMS/Proceedings/y2008/Files/nichols.pdf>
- Schwarz, N., Strack, F., Hippler, H. J., & Bishop, G. (1991). The Impact of Administration Mode on Response Effects in Survey Measurement. *Applied Cognitive Psychology*, 5, 193-212

Appendix: Questions collecting Alternate Addresses, in the actual order

Initial Question

RDD Panel 1: **Did you/NAME live at ADDRESS all of 2006?**

(If not) **What is the other address where you/NAME lived in 2006?**

RDD (Panel 2) and Site: **Were you/Was NAME living at ADDRESS on April 1st or somewhere else?**

(If somewhere else) **What was your/NAME address on April 1st?**

College:

RDD: **During the spring of 2006, < fill all names from roster in appropriate range, example were you or was NAME > attending college?**

Site: **During March or April, <f fill all names from roster in appropriate range, example were you or was NAME > attending college?**

Relative:

RDD: **During 2006, did < fill all names on roster, example you or NAME> live or stay part of the time with another relative?**

Site: **During March or April of this year, did <fill all names on roster, example you or NAME> live or stay part of the time with another relative?**

Military:

RDD: **During 2006, < fill all names from roster in appropriate range, example were you or was NAME> away because of military service?**

Site: **During March or April of this year, <fill all names from roster in appropriate range, example were you or was NAME> away because of military service?**

Job:

RDD: **During 2006, did < fill all names from roster in appropriate range, example you or NAME> have a job that involved living or staying someplace else <other than the military service you just mentioned>?**

Site: **During March or April of this year, did < fill all names from roster in appropriate range, example you or NAME> have a job that involved living or staying someplace else <other than the military service you just mentioned>?**

Seasonal:

Both: <fill with all names on roster **Do you or Does NAME> have a seasonal or second home?**

Often:

RDD: **During 2006, was there any other place <fill with all names on roster you or NAME> stayed often?**

Site: **In the past year, was there any other place <fill with all names on roster you or NAME> stayed often?**

GQ:

Both: **The Census Bureau does a special count of people in places that house groups of people, such as nursing homes, jails and emergency shelters. Even if <fill you weren't or NAME wasn't> living there, did <you/he/she> spend even one night in any of these types of places around April 1st of 2006?**

(Site only) **This card lists the kinds of places I'm interested in:**

The following two questions were asked only for the RDD study (both panels):

Other:

RDD: For people who only stayed at the sample address (after all the alternate addresses were asked.) **I want to make sure we have this right so far.**

<fill You, name, name and name > **lived or stayed at only one address during 2006.**

Is that correct?

If not, **What was the other address where <fill you/NAME> stayed?**

Multiple addresses:

RDD: For people who stayed at multiple addresses (after all the alternate addresses were asked.)

I have collected these addresses for <you/NAME>.

<ADDRESSES>

If respondent reports a problem: **What was the other address where <fill you/NAME> stayed?**

The following two questions were asked only for the site test:

Census Day

Site: **I have collected these addresses for <you/NAME>.**

<ADDRESSES>

Around April 1st, where were you/was NAME living and sleeping most of the time?

Interview Day

Site: **Currently where are you/is NAME living and sleeping most of the time?**