Public Opinion Quarterly, Vol. 72, No. 1 2008, pp. 6–27

A COMPARISON OF ADDRESS-BASED SAMPLING (ABS) VERSUS RANDOM-DIGIT DIALING (RDD) FOR GENERAL POPULATION SURVEYS

MICHAEL W. LINK
MICHAEL P. BATTAGLIA
MARTIN R. FRANKEL
LARRY OSBORN
ALI H. MOKDAD

Abstract Valid and reliable public health data are becoming more difficult to obtain through random-digit dial (RDD) telephone surveys. As a result, researchers are evaluating different survey designs (i.e., sampling frame and survey mode combinations) as complements or alternatives to RDD. Traditionally, mail surveys of the general public have been limited due to a lack of a complete sampling frame of households. Recent advances in electronic record keeping, however, have allowed researchers to develop a sample from a frame of addresses (e.g., the U.S. Postal Service Delivery Sequence File, which appears to provide coverage which rivals or possibly exceeds that obtained through RDD sampling methods). To test the use of this frame for surveying adults aged 18 years and older across a wide geographic area, a pilot study was conducted as part of the 2005 Behavioral Risk Factor Surveillance System (BRFSS). The pilot compared use of a traditional, RDD telephone survey methodology to an approach using a mail version of the questionnaire completed by a random sample of households drawn from an address-based frame.

MICHAEL W. LINK is with Nielsen Media Research, 3784 Ardsley Ct., Marietta, GA 30062, USA. MICHAEL P. BATTAGLIA is with Abt Associates, 55 Wheeler St., Cambridge, MA 02474, USA. MARTIN R. FRANKEL is with Baruch College, City University of New York, 14 Patricia Lane, Cos Cob, CT 06807, USA. LARRY OSBORN is with Abt Associates, 55 Wheeler St., Cambridge, MA 02474, USA. ALI H. MOKDAD is with Centers for Disease Control and Prevention, 4770 Buford Highway, NE MS:K66, Atlanta, GA 30341, USA. The authors thank the BRFSS Coordinators, including Bonnie Davis (California), Bruce Steiner (Illinois), Kenneth O'Dowd (New Jersey), Ziya Gizlice (North Carolina), Jimmy Blanton (Texas), and Katrina Wynkoop Simmons (Washington), and the staff at Marketing Systems Group, particularly Meg Ryan, for their timely assistance with this project. The findings and conclusions in this report are those of the authors and do not necessarily reflect the views of the Centers for Disease Control and Prevention, Abt Associates, or Baruch College. Address correspondence to Michael W. Link; e-mail: Michael.Link@Nielsen.com

doi:10.1093/poq/nfn003 Advanced Access publication February 7, 2008
Published by Oxford University Press 2008.

The findings indicate that the mail survey approach can achieve higher response rates in low-response-rate states (<40%) than RDD (particularly when two mailings are sent). Additionally, the address frame with mail survey design provides access to cell phone only households and offers cost savings over the telephone approach. The resulting sample, however, significantly overrepresents non-Hispanic whites and people with higher levels of education.

For more than 30 years, random-digit dial (RDD) telephone surveys have been the workhorse of the survey research industry. During the past decade, however, participation in most RDD telephone surveys has declined due most likely to factors such as the growth of call-screening technologies, heightened privacy concerns in the face of increased telemarketing calls, and the proliferation of nonhousehold telephone numbers which are typically nonvoice and unassigned numbers (Steeh et al. 2001; Curtin, Presser, and Singer 2005). Additionally, coverage provided by RDD sampling frames has increasingly been called into question. RDD frames do not include households that do not have a telephone of any type (approximately 2.2 percent in 2005; Blumberg and Luke 2007). The increased use of cellular telephones has exacerbated this problem with 12.8 percent of households reported to be cell phone only during the last half of 2006 (Blumberg and Luke 2007). Because most RDD samples typically include only landline numbers, cell phone only households end up being excluded. Additionally, most survey organizations have adopted "list-assisted" RDD sampling approaches, which exclude telephone numbers (approximately 3–4 percent of all households) that are included in "zero blocks," that is, banks of 100 telephone numbers with no directory-listed households (Brick et al. 1995). When we consider all sources, undercoverage in RDD frames may be as high as 15–19 percent.

Probability sample design alternatives to RDD that are of comparable speed, efficiency, and cost are, however, scarce. Face-to-face area probability surveys tend to achieve higher response rates, but the costs associated with traditional counting and listing procedures (i.e., those based on in-person methods rather than mail lists) and conducting in-person interviews are often prohibitive. Conversely, mail surveys have tended to provide a less expensive means of collecting information, although rarely, if ever, has an address frame been available that could provide sufficient coverage of the general population. Likewise, internet penetration, while high (as of May–June 2005, an estimated 68 percent of American adults reported using the internet), does not provide sufficient coverage for conducting surveys of the general adult population (Fox 2005).

More recently, however, the growth of database technology has allowed for the development and maintenance of large, computerized address databases, which may provide survey researchers with an inexpensive address-based sampling (ABS) alternative to RDD for drawing household samples. In particular,

the Delivery Sequence File (DSF) used by the U.S. Postal Service (USPS) is a computerized file that contains all delivery-point addresses serviced by the USPS, with the exception of general delivery (USPS 2005). Each delivery point is a separate record that conforms to all USPS-addressing standards. Initial evaluations of the DSF as a means of reducing the costs associated with enumeration of primarily urban households in area probability surveys have proven to be promising (Iannacchione, Staab, and Redden 2003; O'Muircheartaigh, Eckman, and Weiss 2003; Staab and Iannacchione 2004). These initial studies have shown that for a survey of the general population, the DSF offers potential coverage of 97 percent of U.S. households thereby providing a cost-effective and timely sampling frame. The frame's standardized format also facilitates geocoding of addresses and linkage to other external data sources, such as the Census Zip Code Tabulation Areas data. These data can be used to stratify the frame for sampling target populations.

Use of the DSF does have some drawbacks. Researchers cannot obtain DSF information directly from the USPS, but rather must purchase the information through a nonexclusive license agreement with private list vendors. The quality and completeness of the address information obtained from these vendors can vary significantly based on how frequently the company updates the listings, the degree to which the listings are augmented with information from other available databases, and if the company purges records based on requests from householders to not release their information (Link et al. 2005). Moreover, vendors differ in their experience with drawing probability samples from the DSF list. This can be problematic for researchers who do not wish to draw their own samples and tend to rely upon vendor expertise for this task. The DSF coverage in rural areas also tends to be lower than that in urban areas (Link et al. 2005). Additionally, in some rural areas the DSF contains simplified (i.e., city, state, and zip code only) listings, rather than full street addresses. The percentage of these types of addresses in the database is declining, however, as local governments adopt emergency 911 protocols, which require that all households be identified with a street address. The DSF contains post office (PO) boxes and multidrop addresses (i.e., multiple persons associated with the same address), which may be problematic for in-person and telephone surveys where a street address is required to locate the household or an associated telephone number. Such addresses may be less problematic for mail surveys. Finally, households with multiple mailing addresses (e.g., a street address and a residential PO box) induce selection multiplicities in mail surveys. Iannacchione, Staab, and Redden (2003) provide some evidence that a large percentage of households with residential PO boxes in their Dallas County study also have mail delivered to their street address. In a national sample based on the DSF, however, Staab and Iannacchione (2003) were not able to develop a reliable estimate of the percent of households with a PO box that also received home mail delivery. It is likely that in some areas, households with a PO box do not receive home mail delivery. This may be more likely to occur in rural areas where a PO box is provided at

no cost and no home mail delivery is made. Thus, including PO boxes may be necessary to ensure coverage of all households. Despite these limitations, the DSF appears to be a promising source of information for developing sampling frames of residential addresses.

In this study, we extended ABS-based survey assessment by comparing its use with RDD sampling methods for conducting surveys of the general public across a wide geographic area. In particular, we sought to answer the following questions: What design factors impact case resolution and response rates in ABS-based mail surveys? How do RDD telephone surveys and ABS-based mail surveys compare in terms of response rates and resulting respondent demographics? Can ABS-based mail surveys reach households without telephones and cell phone only households, both of which are currently excluded from most RDD sampling frames? And, finally, how do these different approaches compare in terms of cost?

Methods and Design

As one of the world's largest RDD computer-assisted telephone interview health surveys, the Behavioral Risk Factor Surveillance System (BRFSS) collects uniform, state−specific data on preventive health practices and risk behaviors linked to morbidity and mortality among adults (further details on the BRFSS survey design, methodology, and questionnaire are available at http://www.cdc.gov/brfss). Six states participated in the 2005 BRFSS mail survey pilot: California, Illinois, New Jersey, North Carolina, Texas, and Washington. We selected these states because (1) five of the six (North Carolina being the exception) have response rates below 50%, based on calculations using American Association for Public Opinion (AAPOR) response rate formula #4; (2) they represent various geographic regions of the United States; and (3) when combined, they provide a good representation of the racial and ethnic mix of the U.S. population (AAPOR 2006). Data collection for five of the states was conducted from March 15 through May 15, 2005, while the field period in New Jersey was from March 30 through May 30, 2005.

ABS MAIL SURVEY SAMPLE

Households were sampled from the DSF sample frame, which is based on residential housing unit addresses. The frame includes city-style addresses and PO boxes as well as single-unit, multi-unit, and other types of housing structures. To ensure the most complete coverage possible, we also included units identified by the USPS as being seasonal or vacant units, as well as throwback units (i.e., housing units with addresses of residents who do not want mail delivered to their house, but prefer to pick it up at the local PO) and drop-point units (i.e., locations where mail is dropped off, such as a general

store in a rural area or a trailer park office, and the residents of those addresses pick up their mail at that location). For multi-unit structures, the DSF allows for the unique identification of apartments because it includes fields for the house number, street name, and apartment number. Known business addresses were excluded. A national survey sample vendor provided access to the DSF file and conducted the sampling following our specifications. For the pilot survey, the frame was first sorted by county FIPS code within each of the six participating states. Separate samples of 1,680 addresses per state were then drawn as a systematic random sample, for a total of 10,080 addresses across the six states.

SPLIT SAMPLE EXPERIMENTS TO IMPROVE PARTICIPATION

Embedded within the mail survey pilot were several experiments designed to test the effectiveness of various contacting and within-household selection procedures. Randomization of cases within each of these experiments was conducted independently across the four embedded experiments. These included the following:

- Inclusion of surname/family name on the mailing envelope. Two database vendors matched the sampled addresses with any name(s) they could associate with the address. Cases with a surname match were randomized in an equal fashion into one of two groups: (i) addressed to "The <Surname> Household or Current <State> Resident," or (ii) addressed to "<State> Resident." Cases in which a surname could not be matched were addressed to "<State> Resident."
- Postcard reminder. All cases were equally randomized to one of two groups: (i) received a postcard one week after the initial questionnaire mailing or (ii) did not receive a postcard.
- Second questionnaire mailing. All cases were equally randomized to one of two groups: (i) nonrespondents, who received a second mailing after four weeks, including cover letter and questionnaire, or (ii) nonrespondents, who did not receive a second mailing.
- Alternative within-household selection techniques. Sampled addresses were randomized equally to one of three respondent selection methods: (i) any adult in the household, with the household deciding who responds (a non-probability approach hypothesized to have the lowest associated respondent burden and potentially the lowest level of nonresponse), (ii) adult with the next birthday (based on selection procedures used widely in a number of RDD surveys), or (iii) every adult in the household.

MAIL SURVEY WEIGHTING

The ABS mail survey data were weighted to adjust for probability of selection at both the residential address and the within-household respondent

selection levels (depending on the type of within-household selection used), poststratified by sex and age of the respondents, then ratio-adjusted to equalize weighted state sample sizes. First, household base sampling weights (BSW) were calculated by state. For each state, this sum equaled the DSF population count of residential addresses divided by the sample size. Because we did not include questions on the survey about type of mail delivery, the BSW did not include an adjustment for potential multiplicity of addresses, such as where households have mail delivered to both a street address and a PO box.

Next, a design weight (BSW_2) for version 2 (respondent selection = next birthday) completed questionnaires was calculated as BSW × the number of adults in the household, where the maximum value for the number of adults in a household was capped at 5. For version 1 (respondent selection = any adult) and version 3 (respondent selection = all adults) completed questionnaires, BSW_2 = BSW. A version 3 (all adults) nonresponse adjustment was made (BSW_3) and calculated as BSW_2 times the ratio: (number of adults in the household/number of adults in household that completed a questionnaire), where the maximum value for the number of adults in a household was capped at 5. For version 1 and 2 completed questionnaires, BSW_3 = BSW_2.

For all completed questionnaires in a state combined, BSW_3 was poststratified to 2004 population control totals (provided by Claritas) for 13 age-by-gender cells to produce a poststratified weight (BSW_4). Males aged 18–24 years were combined with males aged 25–34 years, because of the small sample size in the younger age group.

Finally, BSW_4 was ratio-adjusted to produce a final weight (FINALWT) such that the sum of the weights in each state equaled the average of the total adult population across the six states. FINALWT was used to produce the estimates presented in the analyses below because it gave each state an "equal" contribution to the combined state estimates (i.e., the estimates were not dominated by California and Texas).

RDD TELEPHONE SURVEY

The ABS mail pilot surveys were conducted in parallel with the ongoing, monthly RDD data collection, thereby facilitating the comparison of results across the two designs. Telephone survey data from the six participating states for the months of March, April, and May 2005 were used in this analysis. These data were weighted to account for sampling designs, poststratified using the same gender and age categories specified for the mail survey data, and ratio-adjusted so that the sum of the FINALWTs in each state equaled the average of the adult population totals across the six states. More details on BRFSS design and methodology are available elsewhere (Mokdad, Stroup, and Giles 2003) and at http://www.cdc.gov/brfss.

RESPONSE RATE CALCULATIONS

To maximize comparability between the mail and telephone surveys, outcome disposition codes and response rate calculations recommended by AAPOR were used (AAPOR 2006). AAPOR provides a set of case outcome codes for RDD telephone surveys and mail surveys of specifically named persons. For the telephone survey, the original BRFSS disposition codes were mapped to the AAPOR-specified codes and response rates were calculated using AAPOR response rate formula #4. Because the AAPOR mail survey disposition codes apply to surveys in which the respondent's name is known upfront, some modifications were required to handle sampled cases that might not be identified with an eligible residence. Survey packets that were returned from the USPS as undeliverable were coded according to the reason given for nondelivery. Cases in which the survey packet could not be delivered due to an address problem, an address was no longer in service, or because the unit was vacant, including packets marked "cannot be delivered" (no reason given), "cannot be delivered as addressed," "insufficient address," "no mail receptacle," "no such number," "PO box closed," and "vacant," were treated as ineligible.

Determining eligibility in a self-administered mail survey in the absence of a completed questionnaire is more difficult since we cannot determine if someone aged 18 or older resides in the household. Given this, one approach would be to consider all returned cases without a completed interview to be "undeliverable cases with unknown eligibility." However, given that the vast majority of households in the United States typically have at least one adult, we decided it would be more accurate for response rate calculations if returned cases where we could reasonably infer that an occupied household had been reached be considered "eligible interviews." This would have the effect of actually lowering response rates, rather than treating these as "unknown eligibility" where only a portion of the cases would be counted toward the response rate denominator. To do so, returns without a completed questionnaire were determined to be "eligible noninterviews" if (a) someone at the address returned a blank questionnaire in the return envelope or (b) if the case was part of the group where a surname was used on the envelope and the reason for return indicated that "addressee not available" or "addressee no longer at this address." The former is considered a refusal and the latter is assumed to be a household and further assumed to have at least one eligible respondent who did not complete the questionnaire.

Finally, all cases in which no return (either from the respondent or from the USPS) was received were considered to have unknown eligibility and a percentage of these cases were included in the response rate denominator. Unfortunately, determining the residency rate to apply to cases with either an "undeliverable – unknown eligibility" or "no questionnaire returned" disposition was not a clear-cut process. AAPOR only provides guidance for such calculations for mail surveys where the sample members' names are known

upfront (that is, it is a very safe assumption that all or nearly all returns are eligible cases). Using a variant of the methodology used in estimating residency rates in RDD surveys, we initially calculated the residency rates based on cases with known eligibility. Using that approach, approximately 75 percent of the nonrespondent cases were determined to be likely eligible households. Other studies (such as Iannacchione, Staab, and Redden 2003), however, estimate the residency rate to be closer to 90 percent, a percentage which has yet to be empirically confirmed. Like RDD surveys, therefore, there are several potential ways in which this percentage could be calculated. Because of the screening conducted when the sample was selected, we were reasonably confident that the 75 percent rate was low. In the absence of any other empirical basis upon which to make this decision, therefore, we calculated all response rates using an assumption that 90 percent of the undeliverable and no-return cases were eligible households. We believe that the 90 percent estimate is likely on the higher side, so the resulting response rates are "conservative" estimates – that is, there is a higher likelihood that the "true" response rates are higher than those reported here rather than lower.

COST CALCULATIONS

Cost is an important component in the evaluation of any survey design. The data collection costs per 1,000 completed interviews were calculated for both the telephone and mail surveys (assuming a design involving an initial questionnaire mailing, a follow-up postcard, and a second questionnaire mailing) using (1) actual unit costs for materials and supplies based on the pilot study experience, (2) production statistics from the pilot effort, and (3) estimates of industry averages for direct hourly rates and indirect cost rates (i.e., fringe benefits, general and administrative expenses, indirect technical costs, and materials support expenses). Other costs assumed to be nearly equivalent regardless of the survey design were not included, such as overall project management, survey design development, and post-data collection weighting and analysis.

Results

DESCRIPTION OF MAIL SURVEY SAMPLE UNITS AND MATCH RATES

In total, 10,080 addresses were sampled for the ABS mail survey (1,680 per state). The states varied considerably in the ways in which their residents received mail. For instance, nearly two-thirds (62 percent) of North Carolina residents received their mail curbside, compared to less than one-in-five (17 percent) of those living in California. Conversely, a higher percentage (38 percent) of Californians received their mail via either a residential cluster box or a delivery point within a building (i.e., residential central) than did residents of the five other states surveyed. Likewise California residents

(8 percent) along with Texas residents (8 percent) were more likely than residents of the four other states to have their mail delivered to some type of PO box.

The surname match rates were relatively high, ranging from 78 percent in New Jersey to 66 percent in Texas. Surname-matching rates varied considerably, however, depending on how a household received its mail. Among households where mail was delivered curbside, via door-to-door (walking route), or door-slot delivery, the surname match rates were approximately 84 percent. These rates were lower among residential cluster (71 percent) and residential central (48 percent) mail recipients. Among PO box holders, the percentage of surnames identified was significantly lower at 14 percent.

Although telephone numbers for ABS sample units were not a part of the analysis presented here, they were identified for a related follow-up to verify within-household selection (Battaglia et al. in press). The percentage of telephones matched to addresses was slightly lower than the surname match rates, ranging from 66 percent in North Carolina to 52 percent in California. Again, the match rates varied considerably by delivery type: residential curb service (74 percent), residential walking/slot mail route (72 percent), residential cluster (62 percent), residential central (40 percent), and PO box (8 percent). One additional finding of interest was that we were able to match telephone numbers to 62 percent of the addresses classified as nonresponding, potentially eligible households. Additional follow-up by telephone, therefore, while not used in this pilot for completing interviews due to resource constraints, appears quite feasible.

MAIL SAMPLE EFFICIENCY

A total of 3,010 completed mail surveys were obtained across the six states, representing 2,550 unique households (since some households were asked to complete interviews with multiple adults). At the household level, the final disposition of cases was as follows: 2,550 completed questionnaires; 50 eligible noninterviews; 29 undeliverable cases with unknown eligibility; 6,593 cases with no returns resulting in unknown eligibility; 857 undeliverable cases considered ineligible; and 1 case deemed ineligible due to age (respondent reported being younger than 18 years of age).

Not surprisingly, there was variation across both type of address and state in terms of the percentage of addresses determined to be ineligible and those with one or more completed interviews (see table 1). Among traditional city-style addresses, 26.7 percent resulted in a completed interview, while 6.2 percent were determined to be ineligible across the six states. For PO boxes, the percentage of completed interviews was lower (16.6 percent) and the percentage of ineligible cases was higher (12.7 percent) than for city-style addresses. There was also a greater variation in rates across states, with completion rates varying from 4.5 percent (New Jersey) to 20.0 percent (California). The overall rates for

Table 1. Percentage of Completed Interviews and Confirmed Ineligibles by Type of Address by State

Type of address	Total	State					
		CA	IL	NJ	NC	TX	WA
City-style							
Sampled addresses (*n*)	8,968	1,521	1,476	1,454	1,504	1,473	1,540
Confirmed ineligible (%)[a]	6.2	6.2	5.2	3.7	6.9	8.6	6.5
Completed interview (%)[b]	26.7	24.7	31.2	18.4	27.5	24.9	32.7
Post office box							
Sampled addresses (*n*)	561	125	63	89	90	109	85
Confirmed ineligible (%)	12.7	16.8	9.5	10.1	13.3	13.8	9.4
Completed interview (%)	16.6	20.0	12.7	4.5	18.9	18.3	22.4
Throwback or drop unit							
Sampled addresses (*n*)	215	7	81	79	11	22	15
Confirmed ineligible (%)	13.5	28.6	3.7	3.8	36.4	36.4	20.0
Completed interview (%)	16.3	14.3	19.8	13.9	9.1	18.2	13.3
Vacant unit							
Sampled addresses (*n*)	307	22	60	42	75	68	40
Confirmed ineligible (%)	63.2	68.2	50.0	73.8	69.3	67.6	50.0
Completed interview (%)	8.5	0.0	11.7	4.8	5.3	7.3	20.0
Seasonal unit							
Sampled addresses (*n*)	29	5	0	16	0	8	0
Confirmed ineligible (%)	20.7	40.0	NA	12.5	NA	25.0	NA
Completed interview (%)	17.2	20.0	NA	12.5	NA	25.0	NA

[a]Calculated as number of confirmed ineligible addresses/total number of sampled addresses.
[b]Calculated as number of households with at least one completed interview/total number of sampled addresses.

throwback and drop units were similar to those of PO boxes (16.3 percent completion rate; 13.5 percent ineligible rate); however, there was a greater variation across states both in terms of the percentage of total sampled addresses which were either throwback or drop units and in terms of the percentage of cases determined to be ineligible. Among vacant units, 63.2 percent of the addresses were determined to be ineligible, while 8.5 percent resulted in a completed interview. By far, seasonal units were the smallest address-type category which accounted for just 29 of the 10,080 total sampled addresses.

EFFECT OF DESIGN FACTORS ON MAIL SURVEY RESPONSE

Examining the effect of various survey design experiments embedded in the mail survey after obtaining a completed interview from at least one respondent in the addresses sampled, table 2 provides the results of a logistic regression model predicting the effects of the design components on the odds of obtaining

Table 2. Logistic Regression Model: Odds of Receiving a Completed Survey by Survey Design Feature

Address type	Completed interview from total addresses mailed	
	AOR	(95% CI)
Other type	1.00	
City-style	2.27***	(1.74, 2.95)
PO box	1.83***	(1.30, 2.58)
Postcard		
Not sent	1.00	
Sent	1.12*	(1.02, 1.22)
Second questionnaire		
Not sent	1.00	
Sent	1.58***	(1.44, 1.73)
Surname on mailing		
No name available	1.00	
Name not used	2.01***	(1.77, 2.29)
Name used	1.83***	(1.62, 2.09)
Respondent selection		
Any adult	1.00	
Next birthday	0.91	(0.81, 1.02)
All adults[a]	0.90	(0.81, 1.01)
(*n*)	(10,080)	

NOTE.—AOR = adjusted odds ratio; CI = confidence interval.
[a]At least one completed interview received from the household.
Significance: $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

a completed survey from all of the addresses to which a questionnaire was mailed. The odds of receiving a completed interview were 127 percent higher than all other types of addresses (i.e., seasonal, drop-point, throwback, and vacant units) if a city-style address was available and 83 percent higher if a PO box was used. The odds of receiving a completed questionnaire using a family name or surname on the mailing label were 83 percent higher than addresses for which no surname could be identified. However, not using a surname when one was available also had a significant positive effect, doubling the odds of a completed survey (101 percent higher). Sending a second questionnaire improved the odds of a completed survey by 58 percent and sending a postcard reminder one week after the original mailing improved the odds by 12 percent. The within-household respondent selection method used (i.e., any adult, next birthday, or all adults) did not have a significant effect on the odds of receiving a completed survey (see Battaglia et al. in press) for a more detailed analysis of the effects of within-household selection techniques).

Next, we calculated the response rates for the various treatment groups (e.g., combinations of surname use, postcard reminder, and second mailing). As

Table 3. Response Rate by Survey Design Group

Treatment group	Estimated eligible households	Response rate
Name not used, postcard, second questionnaire	782	40.4
Name not used, no postcard, second questionnaire	790	39.8
Name used, postcard, second questionnaire	810	38.0
Name used, no postcard, second questionnaire	803	33.9
No name match, postcard, second questionnaire	500	29.8
Name not used, postcard, no second questionnaire	815	29.0
Name used, postcard, no second questionnaire	807	27.4
No name match, no postcard, second questionnaire	490	25.9
Name not used, no postcard, no second questionnaire	814	25.3
Name used, no postcard, no second questionnaire	810	24.8
No name match, no postcard, no second questionnaire	574	18.3
No name match, postcard, no second questionnaire	576	16.8

NOTE.—Response rate calculated using American Association for Public Opinion Research Response.

Response rate formula #4 (AAPOR 2006). The percentage of mail survey cases with unknown eligibility included in the response rate denominator was set at 90% for all states.

shown in table 3, we obtained the highest response rates for the groups where a name was available but not used and a second questionnaire was mailed. The addition of a postcard reminder to these two factors had little effect on response rates (40.4 percent versus 39.8 percent). The lowest response rates were for the groups where no surname was identified and no second mailing was sent.

COMPARISON OF RESPONSE RATES

Considering all cases in the ABS mail survey, we found that in Washington the mail survey resulted in a substantially higher household-level response rate (i.e., where at least one mail survey was returned from the sampled address) than did the telephone survey (see table 4). California, Illinois, New Jersey, and Texas had rates that were statistically equivalent across the two modes. In North Carolina, the state with the highest RDD response rate, the mail survey's response rate was nearly 15 percentage points lower than the RDD rate.

However, when examining only those cases in the treatment group that received a second mailing, we found that the difference in rates was much starker, with the mail survey performing significantly better in five of the six states: Washington (+6.7 percent), Texas (+5.4 percent), California (+4.5 percent), Illinois (+4.1 percent), and New Jersey (+3.7 percent). In North Carolina, the second mailing markedly improved the response rates for

Table 4. Comparison of DSF Mail Survey and RDD Telephone Survey Response Rates by State and Experiment Condition

State	Response rates		
	RDD telephone survey % (*n*)	DSF mail survey: All cases % (*n*)	DSF mail survey: Cases in 2nd mailing group[a] % (*n*)
Ca	29.4 (5,771)	28.1 (1,432)	33.9*** (691)
Il	35.8 (3,323)	33.7 (1,456)	39.9*** (720)
NJ	22.5 (14,965)	20.0 (1,450)	26.2*** (713)
NC	45.8 (9,782)	31.1*** (1,402)	37.0*** (691)
Tx	31.1 (6.902)	29.0*** (1,375)	36.5*** (661)
Wa	34.1 (17,304)	36.9*** (1,443)	40.3*** (698)

NOTE.—Response rate calculated using American Association for Public Opinion Research.

Response rate formula #4 (AAPOR 2004). The percentage of mail survey cases with unknown eligibility included in the response rate denominator was set at 90% for all states.

RDD = random-digit dialed; DSF = Delivery Sequence File; *n* = estimated number of households.

[a]Includes all cases randomly assigned to this treatment group, including those which complete the survey on the first mailing and did not require a second mailing.

Significance: $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

the mail survey, yet the rate was still significantly lower than that obtained by the telephone survey.

COMPARISON OF DEMOGRAPHIC CHARACTERISTICS

We also looked at the demographic characteristics obtained using the telephone and mail surveys and compared these with results from the 2003 Current Population Survey (CPS). The CPS totals presented here were only from the six states included in the pilot and the CPS weights, like the pilot data weights, were adjusted so that the sum of the weights in each state equaled the average of the total adult population across the six states. The six-state CPS totals were then used as a "gold standard" against which the BRFSS telephone and mail results were compared. Estimates for the telephone and mail surveys were poststratified to adjust for sex and age differences using 2000 Census estimates updated for 2004 by Claritas. Both the telephone and mail surveys differed significantly from the CPS estimates in a number of characteristics (see table 5). Most striking were the differences in the respondents' education levels. In the telephone survey, 59.7 percent of the respondents reported having at least some college education, as did 71.8 percent of those who responded to the mail survey. Both of these results were higher than the 53.8 percent estimated by the CPS.

Table 5. Comparison of Weighted Demographic Characteristics, DSF Mail Survey, RDD Telephone Survey, and Current Population Survey

Demographic characteristics	CPS population estimates[a] (%)	RDD telephone survey (%)	DSF mail survey (%)	Significant levels: RDD versus CPS	DSF versus CPS	RDD versus DSF
Sex				n.s.	n.s.	n.s.
Male	48.5	48.7	48.3			
Female	51.5	51.3	51.7			
Age				n.s.	n.s.	n.s.
18–34	32.6	32.2	32.0			
34–54	29.4	30.6	30.5			
55–64	23.2	21.5	22.1			
65+	14.8	15.6	15.4			
Race				.001	.001	.001
White, non-Hispanic	64.9	68.5	76.1			
Other	35.1	31.5	23.9			
Education				.001	.001	.001
Less than high school	16.9	13.7	7.8			
High-school diploma/GED	29.3	26.5	20.4			
Some college or more	53.8	59.7	71.8			
Income				n.s.	.05	.01
<$50,000	53.6	54.5	51.4			
$50,000+	46.4	45.5	48.6			
Marital status				.001	.01	n.s.
Married/couple	56.6	60.2	59.1			
Not married/single	43.4	39.8	40.9			
Number of children in household				.001	n.s.	.001
None	59.8	56.8	61.0			
One or more	40.2	43.2	39.0			

Continued.

In terms of other demographic groups, the telephone survey overestimated the percentages of white, non-Hispanics; persons in households with incomes of $50,000 or more; and married people, and underestimated the percentage of persons in households with three or more adults. The mail survey also differed significantly from the telephone survey with regards to household education

Table 5. Continued

Demographic characteristics	CPS population estimates[a] (%)	RDD telephone survey (%)	DSF mail survey (%)	Significant levels: RDD versus CPS	DSF versus CPS	RDD versus DSF
Number of adults in household				.01	.001	.001
One	16.2	16.7	19.3			
Two	54.9	56.2	59.5			
Three	28.9	27.1	21.2			
Metropolitan statistical area (MSA)				n.s.	.001	.001
In MSA	86.2	86.8	89.7			
Not in MSA	13.8	13.2	10.3			
[*n*]	[32,963]	[18,780]	[3,010]			

NOTE.—Data are weighted to adjust for sample design, poststratified by sex and age, and ratio-adjusted so that state sample sizes are equivalent.

CPS = Current Population Survey; RDD = random-digit dialed, DSF = Delivery Sequence File.

[a]CPS data include only the six pilot study states: California, Illinois, New Jersey, North Carolina, Texas, and Washington.

level and income as well as number of children and adults in the household. The mail survey also differed significantly from both the CPS and RDD estimates with respect to metropolitan statistical area (MSA) status. Of the mail survey respondents, 89.7 percent lived within an MSA and 10.3 percent lived outside of an MSA (i.e., in a less urbanized area). This latter percentage compares to 13.8 percent from the CPS and 13.2 percent from the RDD survey.

We next examined the success of the mail survey in reaching cell phone only households and households with no telephone, both of which are missed by RDD surveys. We compared findings with estimates from interviews conducted during January through June 2005 as part of the National Health Interview Survey (NHIS), a face-to-face survey with a relatively high response rate. As shown in table 6, 6.5 percent of the adults who responded to the DSF-based mail survey indicated that their household could only be reached by cell phone. This finding for the combined six states in the pilot was similar to the national figure of 6.7 percent reported for the NHIS (Blumberg et al. 2006). Approximately 1 percent of the mail survey respondents stated that they had no telephone access in their household, compared with 1.7 percent of respondents interviewed in the NHIS.

Table 6. Percentage of Adults by Type of Household Telephone Access

Household telephone access	National Health Interview Survey % (95% CI)	BRFSS DSF mail survey % (95% CI)
Land-line	91.6 (91.1, 92.1)	92.6 (90.0, 94.0)
Land-line only	–	*14.9 (13.5, 16.4)*
Land-line and cellular phone	–	*77.7 (75.7, 79.6)*
Cellular phone only	6.7 (6.2, 7.2)	6.5 (5.1, 8.2)
No telephone	1.7 (1.5, 1.9)	1.0 (0.6, 1.4)
[*n*]	[33,614]	[2,947]

NOTE.—Based on interviews NHIS conducted from January to June 2005.
SOURCE.—Blumberg et al. (2006).
CI = confidence interval; *n* = estimated number of households.

COMPARISON OF COSTS

The operational costs of conducting the telephone survey ($79,578 per 1,000 completed interviews) were 12 percent greater than the costs associated with the mail survey ($70,969 per 1,000 completed interviews), assuming a design which included two questionnaire mailings with a follow-up postcard in between (table 7). Although the cost of materials was higher for the mail survey (rates including indirect/overhead charges: $3,938 for telephone survey, $49,600 for mail survey), the telephone survey was much more labor intensive for the same number of completed interviews (rates including indirect/overhead charges: $75,640 for telephone survey, $21,369 for mail survey). The higher indirect rates for labor (estimated to average 150 percent) compared with the indirect costs of materials and supplies (estimated to average 25 percent) further exacerbated these differences.

Discussion

Mail surveys conducted with respondents selected using address-based sampling methods show some promise as an alternative or complementary approach to RDD surveys. In this study, the ABS mail survey produced significantly higher response rates than those obtained in the RDD surveys in five of the six states studied when a second questionnaire mailing was used. The benefit of a second questionnaire mailing was consistent with the findings of other mail surveys (Dillman 2000). The use of a reminder postcard sent one week after the initial mailing provided a modest, but not statistically significant boost to response rates.

Additionally, differences were found in participation rates between respondents with addresses with a surname match and those where a surname could

Table 7. Cost Comparisons per 1,000 Completed Interviews for Conducting an RDD Telephone Survey versus a DSF-Based Mail Survey (With Two Questionnaire Mailings and a Postcard Follow-Up)

	RDD telephone survey	DSF mail survey[a]
Assumptions		
Number of sampled telephone numbers/addresses (per 1,000 completed interviews)[b]	5,000	2,722
Cost calculations for materials/supplies		
Telephone sample ($0.08 per number)/mail sample ($0.11 per address)[c]	$400	$299
Telephone connect charges ($0.55 per number)[c]	$2,750	NA
Printing and postage for mail survey package (@ $6.80 per address)[c]	NA	$37,019
Return postage (@ $1.00 per completed interview)[c]	NA	$1,000
Printing and postage for postcard reminder (@ $0.50 per address)[c]	NA	$1,361
Subtotal for direct cost of materials/supplies	*$3,150*	*$39,680*
Subtotal for indirect costs of materials/supplies (@ 25%)[d]	*$788*	*$9,920*
Total cost for materials/supplies (direct and indirect)	$3,938	$49,600
Cost calculations for labor		
Hours of interviewer time required (@ 2.75 hours/completed interview)[b]	2,750 hours	NA
Hours to print and assemble mailing packages (@ 100 packages/hour)[b]	NA	54 hours
Hours to handle postcard mailing (@ 750 per hour)	NA	4 hours
Hours of receipt/control time required (per 1,000 completed interviews)[b]	NA	500 hours
Hours of data entry time required (per 1,000 completed interviews)[b]	NA	219 hours
Hours of supervisor/monitor/quality control time required (@ 25% of interviewer/receipt-control/data entry time)[d]	688 hours	194 hours

Continued.

Table 7. Continued

	RDD telephone survey	DSF mail survey[a]
Interviewer/receipt-control/data entry time ($8.00 per hour)[c]	$22,000	$6,217
Supervisor or monitor time ($12.00 per hour)[d]	$8,256	$2,331
Subtotal for direct labor costs	*$30,256*	*$8,548*
Subtotal for indirect labor costs, including fringe benefits (@ 150%)[d]	*$45,384*	*$12,822*
Total cost for labor (direct and indirect costs)	$75,640	$21,369
Total cost of materials/supplies and labor (absent fee/profit) per 1,000 completed interviews	$79,578	$70,969

NOTE.—Comparison includes only costs of actual data collection and does not include other costs such as survey design, statistical support, analysis, or project management (other than direct supervision costs noted in the table).

[a]Costs based on mail survey design yielding the highest response rate, which includes two questionnaire mailings to all sampled addresses with a postcard follow-up between questionnaire mailings.

[b]Based on production statistics from pilot study.

[c]Based on cost data from pilot study.

[d]Based on estimates of average rates across survey research industry.

not be identified, with the former being more likely to respond regardless of whether the name was actually used on the mailing envelope. This finding is similar to the differences found in RDD surveys between telephone numbers with an identifiable address and those without an address match (Link and Mokdad 2006). It appears that persons who are more readily identifiable in public databases, such as those used for surname- or address-matching, tend to be more willing to participate in surveys than people who are more difficult to identify. Although the differences in participation between the two surname-identified groups were not statistically significant, there are potential issues that might make not using the surname preferable, even when available. If the surname match is incorrect, household members may be more likely to discard the mailing without opening it. Although the differences in participation rates seen here were not statistically significant, the group in which surnames were available but not used in the mailing had the highest overall response rates. Use of a surname may also influence respondent selection, particularly in households where adults may not share the same last name. Finally, use of a surname may raise concerns about confidentiality among some respondents leading them to alter their responses, particularly to sensitive questions (Link et al. 2006).

The ABS mail survey also provided access to households with only cell phones, and to a smaller degree, to households with no telephone coverage. The former group is increasingly becoming a focus of concern among researchers, whereas the latter group has always been unreachable by telephone surveys. The percentage of cell phone only households across the six states examined here was similar to the percentage reported at the national level (Blumberg et al. 2006). Unfortunately, there are currently no state-level data on the percentage of cell phone only households with which to compare the pilot study findings.

The mail survey was also advantageous in that it cost less to conduct. To obtain the same number of completed interviews, the telephone survey cost was 12 percent higher than the amount required for the mail survey.

Nonetheless, the ABS mail survey approach had some drawbacks. First, improvement in response rates were obtained only in those states where the RDD response rates were low (i.e., below 50 percent). In North Carolina, where the RDD response rate was above 45 percent, the mail survey response rate was significantly lower than the telephone response rate, even when two questionnaire mailings were sent. Second, the mail survey obtained responses from a much higher percentage of non-Hispanic whites and people who had at least some college education and from a significantly lower percentage of persons who do not live in an MSA than did either the RDD survey or the CPS. This skewed distribution across these key demographic groups raises some concerns about potential bias in the estimates (see Link et al. 2006 for more detailed analysis of this issue). Third, use of the mail survey approach would likely force some fundamental changes in the way in which a surveillance system, such as BRFSS,

currently operates. Mail surveys require a longer fielding period (typically eight weeks or more) than the current monthly schedule for the BRFSS telephone survey. Use of a mail survey would also reduce the length and flexibility of the BRFSS questionnaire. The telephone version of the BRFSS contains a core survey of 70–75 questions (asked in all states), optional modules of 1–20 questions (standardized topic modules that can be adopted by the states), and state-added question modules of 1–50 questions (typically unique to each state, focusing on state-specific health issues). The 2005 mail survey pilot tested only the core questionnaire. Lengthening the mail questionnaire could increase respondents' reluctance to complete the survey, and customizing each state survey to include the optional or state-added modules would significantly increase the operational complexity of administering the survey.

Another issue that is not addressed completely in this study is the potential of multiplicity of mail addresses by including non-city-style addresses in the sampling frame (i.e., PO boxes, drop-point units, etc.). Because this was a pilot study, we wanted to maximize coverage of housing units and so we included all residential address types in the sampling frame: city style, vacant, seasonal, throwbacks, and drop-point units. The throwback addresses (0.3 percent of total addresses) and drop-point unit addresses (2.0 percent of total addresses) do not necessarily represent duplication of units in the sampling frame, but the inclusion of drop-point units may lead to some subjectivity as to which housing unit associated with a drop point receives the survey mailing. This could be handled in the sample design by creating a separate stratum for drop points and including all drop-point units associated with the sample drop points in the sample (i.e., a one-stage cluster sample design). Given the small percentage of drop-point units in the sampling frame, this approach may not be warranted in relation to the bias that may be incurred.

Potential duplication in the frame caused by PO boxes is a more important issue because residential PO boxes account for 7 percent of the addresses in the sampling frame. As noted earlier, however, it is unclear under what conditions this overlap is most likely to occur. Because this was a pilot survey and we needed to balance respondent burden issues with the desire to obtain as much information on mail delivery as possible, we decided not to add a question to the sample addresses that are PO boxes asking them whether they also have home mail delivery and vice versa for those with city-style addresses. Inclusion of such questions would allow for refinement of the weight adjustment to account for multiplicity and should be a focus of future research with ABS-based samples.

The study also highlights areas where data collection efficiencies may be gained. In addition to city-style addresses, it appears that inclusion of PO boxes, throwback and drop-point units is important for both coverage and the nonignorable number of completed interviews obtained from these types of addresses. The same does not appear to apply, however, to addresses identified as vacant or seasonal by the UPSP. More than 60 percent of the addresses

identified as vacant units were confirmed to be ineligible addresses. Although the percentage of vacant units resulting in a completed interview was relatively high in some states (20 percent in Washington), the number of completed interviews from this address type as a percentage of the total number of completed interviews was quite small (8 of 532 in Washington). Exclusion of vacant units from an ABS sample design is unlikely, therefore, to have a significant negative effect on coverage bias. Likewise the extremely small number of addresses designated as seasonal units argues for their exclusion as well.

While the DSF appears to be an effective frame for conducting address-based sampling of the general population, its true potential may be in facilitating mixed-mode surveys. Crossreferencing addresses with other public databases yielded telephone numbers for half to two-thirds of the addresses depending on the state. Moreover, among the subset of nonrespondents cases with unknown eligibility that received two mail questionnaires, over 60 percent had identifiable telephone numbers. Although additional research will be required to determine how accurately these telephone numbers match with the addresses, early indications are that conduct of a telephone survey follow-up to the mail survey is quite feasible. Moreover, address-based sampling may facilitate the more cost-effective use of other interviewing modes, such as web surveys or interactive voice response (IVR). Households could be sampled through an address-based frame, such as the DSF, then provided a link to a web site, given the telephone number for an IVR survey, mailed a hardcopy questionnaire, or any combination of these approaches. Recent studies have shown that combining telephone surveys with either web or mail survey options can produce higher response rates in general population surveys than use of telephone alone (Link and Mokdad 2006). Resources permitting face-to-face surveys could also be added to this mix, particularly since use of the DSF was initially tested as a means of identifying households for such surveys (Iannacchione, Staab, and Redden 2003; O'Muircheartaigh, Eckman, and Weiss 2003; Staab and Iannacchione 2004). The DSF, therefore, has the potential to serve as a sampling base for a wide variety of single or multimode survey designs.

A great deal of study is needed before use of address-based sampling can be recommended as a standard sampling approach for ongoing surveys such as the BRFSS. The findings do, however, offer encouragement, particularly for states and areas with low RDD response rates, urban areas where address coverage is higher, and surveys where all households are eligible. Future research efforts should continue to evaluate the expansion of address-based coverage as more rural areas adopt city-style addresses that conform to 911 emergency number rules. Use of the DSF in particular as a sampling frame for the conduct of surveys via other modes (telephone, web, IVR, face-to-face, etc.) as well as combinations of modes needs to be explored more fully as complements to RDD designs. Given the continued decline in RDD response rates and the increased use of cell phones it seems clear that an alternative design is needed to fill a growing gap as the new mainstay for survey research.

References

American Association for Public Opinion Research (AAPOR). 2006. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. 4th ed. Lenexa, KS: AAPOR.

Battaglia, Michael P., Michael W. Link, Martin Frankel, Larry Osborn, and Ali Mokdad. "An Evaluation of Respondent Selection Methods for Household Mail Surveys." *Public Opinion Quarterly* (forthcoming).

Blumberg, Stephen, and Julian Luke. 2007. "Wireless Substitution: Early Release of Estimates Based on Data from the National Health Interview Survey, July–December 2006." Available at http://www.cdc.gov/nchs/data/nhis/earlyrelease/wireless200705.pdf.

Blumberg, Stephen J., Julian V. Luke, and Marcie L. Cynamon. 2006. "Telephone Coverage and Health Survey Estimates: Evaluating the Need for Concern about Wireless Substitution." *American Journal of Public Health* 96:926–31.

Brick, J. Michael, Joseph Waksberg, Dale Kulp, and Amy Starer. 1995. "Bias in List-Assisted Telephone Samples." *Public Opinion Quarterly* 59:218–35.

Curtin, Richard, Stanley Presser, and Eleanor Singer. 2005. "Changes in Telephone Survey Nonresponse over the Past Quarter Century." *Public Opinion Quarterly* 69:87–98.

Dillman, Donald. 2000. *Mail and Internet Surveys: The Tailored Design Method*. New York: Wiley.

Fox, Suzanne. 2005. Digital Divisions, Pew Internet and American Life Project Report. Available at http://www.pewinternet.org/pdfs/PIP_Digital_Divisions_Oct_5_2005.pdf, accessed February 13, 2006.

Iannacchione, Vince G., Jennifer M. Staab, and David T. Redden. 2003. "Evaluating the Use of Residential Mailing Addresses in a Metropolitan Household Survey." *Public Opinion Quarterly* 76:202–10.

Link, Michael W., Michael Battaglia, Pamela Giambo, Martin Frankel, and Ali Mokdad. 2005. "Assessment of Address Frame Replacements for RDD Sampling Frames." Paper presented at the Annual Meeting of the American Association for Public Opinion Research, Miami Beach, FL.

Link, Michael W., Michael Battaglia, Martin Frankel, Larry Osborn, and Ali Mokdad. 2006. "Address-Based Versus Random-Digit Dialed Surveys: Comparison of Key Health and Risk Indicators." *American Journal of Epidemiology* 164:1019–25.

Link, Michael W., and Ali H. Mokdad. 2006. "Can Web and Mail Survey Modes Improve Participation in an RDD-Based National Health Surveillance?" *Journal of Official Statistics* 22:293–312.

Mokdad, Ali H., Donna Stroup, and H. Wayne Giles. 2003. "Public Health Surveillance for Behavioral Risk Factors in a Changing Environment: Recommendations from the Behavioral Risk Factor Surveillance Team." *Morbidity and Mortality Recommendations and Reports* 52(RR09):1–12.

O'Muircheartaigh, Colm, Stephanie Eckman, and Charlene Weiss. 2003. "Traditional and Enhanced Field Listing for Probability Sampling." *Proceedings of the American Statistical Association, Survey Methodology Section (CD-ROM)*, 2563–67. Alexandria, VA: American Statistical Association.

Staab, Jennifer M., and Vince G. Iannacchione. 2004. "Evaluating the Use of Residential Mailing Addresses in a National Household Survey." *Proceedings of the American Statistical Association, Survey Methodology Section (CD-ROM)*, 4028–33. Alexandria, VA: American Statistical Association.

Steeh, Charlotte, Nicole Kirgis, Brian Cannon, and Jeff DeWitt. 2001. "Are They Really as Bad as They Seem? Nonresponse Rates at the End of the Twentieth Century. *Journal of Official Statistics* 17:227–47.

United States Postal Service. 2005. Delivery Sequence File. Available at http://www.usps.com/ncsc/addressservices/addressqualityservices/deliverysequence.htm, accessed April 12, 2005.