Mode and Address Frame Alternatives to Random-Digit Dialing

Michael W. Link*[1], Michael P. Battaglia[2], Martin R. Frankel[3], Larry Osborn[2], and Ali H. Mokdad[1]

[1] Centers for Disease Control and Prevention, [2] Abt Associates, [3] Abt Associates and Baruch College, CUNY,
* MLink@cdc.gov, Centers for Disease Control and Prevention, 4770 Buford Highway NE, K-66, Atlanta, GA 30341

Abstract

Traditionally, mail surveys of the general public have been limited by lack of a complete sampling frame of households. Advances in electronic record keeping, however, have allowed researchers to develop and sample from a frame of addresses (the U.S. Postal Service Delivery Sequence File), which appears to provide coverage which rivals that obtained through RDD sampling methods. Testing the use of this frame for surveying the general adult population, a pilot study was conducted as part of the 2005 Behavioral Risk Factor Surveillance System (BRFSS), comparing use of RDD methodology to an approach using a mail version of the questionnaire completed by a random sample of households drawn from an address-based frame. The findings indicate that higher response rates can be achieved in low response rate states (< 40%) using the mail survey approach (particularly when two mailings are sent). Additionally, the address frame / mail survey design provided access to cell phone-only households and offered considerable cost savings over the telephone approach. The resulting sample, however, significantly over-represented those with higher levels of education and non-Hispanic whites and under-represented people in less urban areas.

1.0. Introduction

For more than 30 years, random-digit dialed (RDD) telephone surveys have been the workhorse of the survey research industry. Over the past decade, however, participation in most RDD telephone surveys has declined due to factors such as the growth of call screening technologies, heightened safety and privacy concerns, and the proliferation of state and federal do not call lists (Steeh et al 2001; Curtin, Presser, Singer 2005). The integrity of RDD sampling frames has also increasingly been called into question. RDD sampling frames have always excluded the portion of the population (approximately 1.7%) in 2005, who do not have a land-line telephone in their household (Blumberg, Luke, and Cynamon 2006). Additionally, most survey organizations have adopted "list assisted" RDD sampling approaches, which exclude telephone numbers (approximately 3% to 4% of all households) in "zero blocks" -- that is, banks of 100 telephone numbers with no directory-listed households (Brick et al 1995). Noncoverage problems have been further exacerbated with the increased use of cellular telephones, with 6.7% of households reported to be cell-phone only in 2004 and this percentage is expected to increase over time (Blumberg, Luke, and Cynamon 2006). Because most RDD samples are typically drawn from area code-central office code combinations assumed to be land-line numbers (including mixed-use exchanges), most cell-phone-only households are excluded from RDD sampling frames. When we consider all sources of under-coverage in RDD frames (i.e., households with no telephones, those in zero blocks, and cell-phone-only households), the percentage of US households not covered by RDD frames may be as high as 9% to11%.

Yet, alternative probability sample designs to RDD of comparable speed, efficiency, and cost are scarce. The growth of database technology, however, has allowed for the development and maintenance of large, computerized address databases, which may provide survey researchers with an inexpensive alternative to RDD for drawing household samples. 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). On the file, 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 households in area probability surveys have proven promising (Iannacchione, Staab, and Redden 2003; Staab and Iannacchione 2003; O'Muircheartaigh, Eckman, and Weiss 2002). These initial studies showed that for a survey of the general population, the DSF offers potential coverage of 97% of the households in the United States, thereby providing a cost-effective and timely sampling frame. The standardized format of the frame also facilitates geocoding of addresses and linkage to other external data sources such as the Census Zip Code Tabulation Areas (ZCTA) data. These data can then be used to stratify the frame for sampling of target populations.

Use of the DSF does have some drawbacks. First, coverage in rural areas tends to be lower than in urban areas. Initial assessments show that the level of undercoverage is greater than 10% in a large majority of counties where less than 25% of adults live in an urbanized area (Link, Battaglia, Giambo et al. 2005).

Second, 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. Over time, therefore, the more simplified designations are expected to be replaced by full street address information. Third, the DSF contains PO boxes and multidrop addresses (i.e., addresses associated with more than one name), which may be problematic for in-person and telephone surveys where a street address is required to locate the household or identify a telephone number associated with the household. Such addresses may be less problematic for mail surveys, where the initial goal is to ensure the mailed questionnaire is delivered to the sampled household. Despite these limitations, the DSF appears to be a promising source of information for developing sampling frames of residential addresses.

To date the DSF has been tested primarily within the context of creating segment housing unit listings for area probability sampling. In this study, we extend assessment of the DSF to a comparison with RDD sampling methods for conducting surveys of the general public across a wide geographic area.

2.0 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, calculated using American Association for Public Opinion (AAPOR) response rate #4, below 40%, (2) they represent the various geographic regions of the United States, and (3) combined, they provide a good representation of the racial and ethnic mix of the U.S. population (AAPOR 2004). Data collection for the pilot was conducted March 15-May 15, 2005.

2.1 Delivery Sequence File – Mail Survey Sample

Households were sampled from the DSF sample frame, which is based on residential housing unit addresses (commonly referred to as deliverable residential addresses). The frame includes city-style addresses, and PO boxes and covers single-unit, multi-unit, and other types of housing structures. To ensure the most complete coverage possible we included units in the frame identified by the USPS as being seasonal or vacant units, as well as throwback units (i.e., addresses who do not want mail delivered to their house, but prefer to pick it up at the local post office) and drop point units (i.e., locations where mail is dropped off and the addresses associated with that drop point pick up their mail at that location, such as a general store in a rural area or a trailer park office). 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 stratified 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 6 states.

2.2 Split Sample Experiments

Embedded within the mail survey pilot were several split sample experiments designed to test the effectiveness of various contacting and within-household selection procedures. These included:

- Inclusion of surname/family name on the mailing envelope -- 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) "<State> Resident." Cases where a surname could not be matched were addressed to "<State> Resident."
- Postcard reminder -- Cases were randomized to: (i) receive a postcard 1 week after initial questionnaire mailing or (ii) not receive a postcard.
- Second questionnaire mailing -- cases were randomized to: (i) nonrespondents, who received a second mailing after 4 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 nonprobability 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.

2.3 Mail Survey Weighting
The mail survey data were weighted to adjust for probability of selection both at the residential address and within-household respondent selection levels (depending on the type of within household selection used), post-stratified by sex and age of the respondents, then ratio adjusted to equalize weighted state sample sizes. Additional information about weighting is available upon request from the authors.

2.4 Random Digit-Dialed Telephone Survey
The mail pilot surveys were conducted in parallel with the ongoing, monthly RDD data collection facilitating comparison of results across the two designs. Telephone survey data from the 6 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, post-stratified using the same gender and age categories specified for the mail survey data, and ratio-adjusted so that the sum of the final weights 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 et al. 2003) and at http://www.cdc.gov/brfss.

2.5 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 2004). 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 where the name of the respondent is known up-front, some modifications were required to deal with sampled cases which might not be identified with an eligible residence. All cases in which some type of return (either from the respondent or from USPS) was not received were considered to have unknown eligibility and the residency rate for these sampled addresses was estimated using the sample for which eligibility was determined. Cases which were returned as undeliverable from the USPS were coded according to the reason given for not being able to deliver the survey packet. Those where the packet could not be delivered due to an address problem, address no longer in service, or the unit was vacant were treated as ineligible, including those 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."

2.6 Cost Calculations
Cost is an important component in the evaluation of any survey design. The data collection costs per 1,000 completed interviews was calculated for both the telephone and mail surveys 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.

3.0 Results
3.1 Effect of Design Factors on Mail Survey Response
A total of 3,010 completed mail surveys were obtained across the 6 states, with at least 1 complete obtained from 2,550 of the sampled addresses. At the household level, the final disposition of cases was as follows: 2,550 completed questionnaires, 50 eligible non-interviews, 29 undeliverable cases with known 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 under 18 years of age).

We first examined the effect of various survey design experiments embedded in the mail survey on obtaining a completed interview from at least one respondent in the addresses sampled. Table 1 provides the results of a logistic regression model predicting the effects of the design components on the odds of obtaining a completed survey from all of the addresses to which a questionnaire was mailed. The odds of receiving a completed interview were 127% higher than all other types of addresses (i.e., seasonal, drop-point, throw-back, and vacant units) if a city style address was available and 83% higher if a post office box was used. The odds of receiving a completed questionnaire using a family name or surname on the mailing label were 83% higher compared to 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% higher). Sending a second questionnaire improved the odds of a completed survey by 58% and sending a postcard reminder 1 week after the original mailing improved the odds by 12%. 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. 2005, 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). 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 post card reminder to these two factors improved the response rate only slightly from 44.3% to 44.9%. The lowest response rates were for the groups where no surname was identified and no second mailing was sent.

3.2 Comparison of Response Rates

Comparing the response rates of the mail survey to that of the telephone survey, we found that in two of the six states (Texas and Washington) the mail survey resulted in substantially higher household-level response rates (i.e., where at least 1 mail survey was returned from the sampled address) than the telephone survey (see Table 2). Considering all cases in the mail survey, the difference in rates between the mail and telephone surveys were +4.4% for Texas and +5.8% for Washington. California, Illinois and New Jersey had rates which were statistically equivalent across the two modes. In North Carolina, the state with the highest RDD response rate, the mail survey resulted in a response rate which was 9.5 percentage points lower.

Examining, however, only those cases in the treatment group to receive a second mailing, the difference in rates is much starker, with the mail survey performing significantly better in five of the six states: Illinois (+7.0%), New Jersey (+8.0%), California (+9.8%), Washington (+10.8%), and Texas (+13.3%). In North Carolina, the second mailing markedly improved the response rates for the mail survey, making it statistically equivalent to, but not greater than the telephone survey rate.

3.3 Comparison of Demographic Characteristics

We also looked at the demographic characteristics obtained using the telephone and mail surveys and compared these to results from the 2003 Current Population Survey (CPS). The CPS was used as a "gold standard" against which the BRFSS telephone and mail results were compared. Estimates for the telephone and mail surveys were post-stratified 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 (Table 3). Most striking were the differences in education levels of the respondents. In the telephone survey, 59.7% reported having at least some college education, as did 71.8% of those responding to the mail survey. Both of these are higher than the 53.8% estimated by the CPS. The mail and telephone surveys also differed significantly from the CPS estimates with respect to metropolitan statistical area (MSA) status. Of the mail survey respondents, 89.7% lived within an MSA and 10.3% lived outside of an MSA (i.e., in a less urbanized area). This latter percentage compares to 13.2% from the RDD survey and 13.8% from the CPS, who live in outside of an MSA area. In terms of other demographic groups, the telephone survey overestimated the percentages of white, non-Hispanics and married people and underrepresented the percentages of persons with no children in the household and households with three or more adults.

Similarly, the mail survey overestimated the percentages of white, non-Hispanics, households with family incomes of $50,000 or more, and married people and underestimated the percentage of households with three or more adults. The mail survey also differed significantly from the telephone survey with regards to household education level and income as well as number of children and adults in the household.

Next, we examined the success of the mail survey in reaching cell-phone-only households and households with no telephone coverage – both of which are missed by RDD surveys. We made comparisons with estimates from interviews conducted 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 5, 6.5% of the adults responding to the DSF-based mail survey indicated that their household could only be reached by cell phone. This was similar to the 6.7% reported for the NHIS (Blumberg et al., 2006). Approximately 1% of mail survey respondents said they had no telephone access in their household compared to 1.7% of those interviewed in the NHIS.

3.4 Comparison of Costs

The operational costs of conducting the telephone survey were nearly two-and-a-half times greater than the costs associated with the mail survey: $79,578 per 1,000 completed interviews for the telephone survey versus $30,919 per 1,000 completed interviews for the mail survey. Although the cost of materials was higher for the mail survey (loaded rates: $3,938 for telephone survey, $10,211 for mail survey), mounting telephone surveys is much more labor intensive for

the same number of completed interviews compared to a mail survey (loaded rates: $75,640 for telephone survey, $20,708 for mail survey). The higher indirect rates for labor (estimated to average 150%) compared to those for materials and supplies (estimated to average 25%) further exacerbated these difference.

4.0 Conclusions

Mail surveys conducted with respondents sampled from addresses listed in the DSF show some promise as an alternative or complementary approach to RDD surveys of the general population. The mail survey approach had several advantages. First, the mail survey response rates were significantly higher than those obtained in the RDD surveys in five of the six states when a second questionnaire mailing was used. The benefit of a second questionnaire mailing is consistent with the findings of other mail surveys (Dillman 2000). Use of a reminder post-card one week after the initial mailing also appears to provide a modest boost to response rates. Additionally, there were clear differences in participation rates between those with addresses in which a surname was found and those where a surname could not be identified, with the former being more likely to respond regardless of whether the name was actually used on the mail envelop. This is similar to the differences found in RDD surveys between households in which an address can be matched to a database and those where an address cannot be matched (Link and Mokdad 2005). 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 those who are more difficult to identify. The use of surname may raise concerns, however, about confidentiality among some respondents leading them to alter their responses, particularly to sensitive questions (Link, Battaglia, Frankel et al. 2006). Second, the mail survey 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, while the latter group has always been unreachable by telephone survey. Third, the mail survey was considerably less costly to conduct. For the same number of completed interviews, the telephone survey was more than twice the cost of the mail survey.

The mail survey approach did, however, have a number of drawbacks. First, improvement in response rates were obtained only in those states where the RDD response rates were low (i.e. below 40%). In the one state, North Carolina, where the RDD response rates was above 45%, the mail survey did not out perform the telephone survey. Second, the mail survey obtained responses from a significantly lower percentage of persons who do not live in an MSA and a much higher percentage of persons with some college or more education than did either the RDD survey or the CPS. The same is true of the percentage of non-Hispanic whites who completed the surveys. This skewed distribution across these key demographic groups raises some concerns about potential bias in the estimates (see Link et al 2005 for more detailed analysis of this issue). For several other characteristics, such as marital status and number of children in the household the mail survey performed somewhat better than the RDD survey in comparison to the CPS. Third, use of the mail survey approach would likely force some fundamental changes in the way in which a study, such as BRFSS, currently operates, particularly with regards to curtailing the flexibility of the survey. The mail survey requires a longer fielding period (typically 8 weeks or more) compared to 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.

This study has some limitations. First, the DSF frame does not provide universal coverage of all households, particularly in more rural and lower income areas (Link, Battaglia, Frankel et al. 2005). Second, the number of completes obtained in each of the treatment groups (i.e., combinations of surname use, postcard reminder, and second mailing) did not allow us to look at demographic characteristics of respondents by these different groups. Third, the study was conducted in six states, which may not be representative of either the nation or other populations.

A great deal more study is needed before use of the USPS DSF can be recommended as a standard approach to sampling, and mail surveys as the preferred mode of interviewing, for an on going survey such as the BRFSS. The findings do, however, offer encouragement, particularly for states and areas with low RDD response rates, urban areas where DSF coverage is higher, and for surveys where all households are eligible. Future research efforts should continue to evaluate the expansion of DSF coverage as more rural areas adopt city-style addresses that conform to 911 emergency number rules. Use of the DSF and mail surveys also should be explored as complements to RDD in dual-frame and mixed-mode surveys designs. Future research will need to address the applicability of these methods to surveys that screen for a specific target population. Given the continued decline in RDD response rates and the increased use of cell phones (particularly as an alternative to landlines), it seems clear that an alternative design is needed to fill a

growing gap as the new workhorse for survey research.

5.0 Acknowledgements
We thank the BRFSS coordinators, including Bonnie Davis (California), Bruce Steiner (Illinois), Kenneth O'Dowd (New Jersey), Ziya Gizlice (North Carolina), Jimmy Blanton (Texas), Katrina Wynkoop Simmons (Washington), the staff at Marketing Systems Group, in particular Meg Ryan, for their timely assistance with this project.

6.0 References

American Association for Public Opinion Research. 2004. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys, 3rd Edition*. Ann Arbor, MI: AAPOR.

Battaglia, Michael P., Michael W. Link, Martin R. Frankel, Larry Osborn, and Ali H. Mokdad. 2005. "An Evaluation of Respondent Selection Methods for Household Mail Surveys." *Proceedings of the American Statistical Association, Survey Methodology Section* (CD-ROM), Alexandria, VA.

Blumberg, Stephen J., Julian V. Luke, J, and Marcie L. Cynamon. 2006. "Telephone Coverage and Health Survey Estimates: Is Concern about Wireless Substitution Warranted?" *American Journal of Public Health* (forthcoming).

Brick, J. Michael, Joseph Waksberg, Dale Kulp, and Amy Starer. 1995. "Bias in List-assisted Telephone Samples." *Public Opinion Quarterly* 59: 218-235.

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: John Wiley & Sons, Inc.

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* 76:202-210.

Link, Michael W., Michael P. Battaglia, Pamela Giambo, Martin R. Frankel, Ali H. Mokdad, and R. Sowmya Rao. 2005. "Assessment of Address Frame Replacements for RDD Sampling Frames." Paper presented at the American Association for Public Opinion Research Conference, Miami, FL.

Link, Michael W., Michael P. Battaglia, Martin R. Frankel, Larry Osborn, and Ali H. Mokdad. 2005. "Address-Based Versus Random-Digit Dial Sampling: Comparison of Data Quality from BRFSS Mail and Telephone Surveys." Proceedings of the 2005 Federal Committee on Statistical Methodology Research Conference (CD-ROM), Arlington, VA.

Link, Michael W., and Ali H. Mokdad. 2005. "Advance Letters as a Means of Improving Respondent Cooperation in RDD Studies: A Multi-state Experiment." *Public Opinion Quarterly* 69:572-587.

Mokdad, Ali H., Donna Stroup, 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, Steven Eckman, and Cheryl Weiss. 2003. "Traditional and Enhanced Field Listing for Probability Sampling." *Proceedings of the American Statistical Association, Survey Methodology Section* (CD-ROM), Alexandria, VA, pp.2563-2567.

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), Alexandria, VA, pp.4028-4033.

Steeh, Charlotte, Nicole Krigis, 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: 227-247.

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

United States Bureau of Census. 2005. "County Household Count Data (accessed April 1, 2005). Available at http://www.census.gov/popest/housing/files/HU-EST2003-ST.csv

Table 1. Logistic regression model: odds of receiving a completed survey form from all forms mailed by survey design feature

	Completed interview from total addresses mailed	
Address Type	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[1]	0.90	(0.81, 1.01)
(n)	(10,080)	

AOR = adjusted odds ratio; CI = confidence interval
Significance: * = p<.05, **= p<.01, ***= p<.001
[1] At least one completed interview received from the household.

Table 2. Comparison of DSF mail survey and RDD telephone survey response rates by state and experiment condition

	Response Rates[1]		
State	RDD telephone survey % (n)	DSF mail survey: All cases % (n)	DSF mail survey: cases in 2nd mailing group[2] (n)
California	29.4 (5,771)	31.8 (1,266)	39.2*** (597)
Illinois	35.8 (3,323)	36.2 (1,356)	42.8*** (671)
New Jersey	22.5 (14,965)	23.2 (1,250)	30.5*** (614)
North Carolina	45.8 (9,782)	36.3*** (1,200)	42.5 (602)
Texas	31.1 (6.902)	35.5** (1,122)	44.4*** (543)
Washington	34.1 (17,304)	39.9*** (1,334)	44.9*** (626)

RDD=random-digit dialed; DSF=Delivery Sequence File
(n) = estimated number of households.
Significance based on comparisons with RDD telephone survey: * = p<.05, **= p<.01, ***= p<.001
[1] Response rate calculated using American Association for Public Opinion Research Response Rate Formula #4 (AAPOR 2004).
[2] 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.

Table 3. Comparison of weighted demographic characteristics, DSF mail survey, RDD telephone survey, and Current Population Survey (CPS)

Demographic characteristics	CPS population estimates %	RDD telephone survey %[1]	DSF Mail survey %[1]	Significance levels RDD -vs- CPS	DSF -vs- CPS	RDD -vs- 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				***	***	***
White, non-Hispanic	64.9	68.5	76.1			
Other	35.1	31.5	23.9			
Education				***	***	***
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.	*	**
< $50,000	53.6	54.5	51.4			
$50,000+	46.4	45.5	48.6			
Marital status				***	**	n.s.
Married/couple	56.6	60.2	59.1			
Not married/single	43.4	39.8	40.9			
Number of children in household				***	n.s.	***
None	59.8	56.8	61.0			
One or more	40.2	43.2	39.0			
Number of adults in household				**	***	***
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)[2]				*	***	***
In MSA	86.2	86.8	89.7			
Not in MSA	13.8	13.2	10.3			
[n]	[32,963]	[18,780]	[3,010]			

CPS = Current Population Survey; RDD = random-digit dialed, DSF = Delivery Sequence File
Significance: * = p<.05, **= p<.01, ***= p<.001

[1]Data are weighted to adjust for sample design , post-stratified by sex and age, and ratio adjusted so state sample sizes are equivalent.

[2]Metropolitan Statistical Area for the telephone and mail surveys was based on a Core-Based Statistical Area (CBSA) with at least one urban area with a population of 50,000 or higher. MSA for the March 2004 CPS was based on pre-CBSA Metropolitan Areas. of average rates across survey research industry.