

Address-based Sampling (ABS) as an “Alternative” to RDD: A Test in California

Matt Jans¹, David Grant¹, Annie Lee¹, Royce Park¹,
Sherman Edwards², John Rauch², Ismael Flores-Cervantes²

¹UCLA Center for Health Policy Research, California Health Interview Survey,
10960 Wilshire Blvd., Ste. 1550, Los Angeles, CA 90034

²Westat, 1600 Research Boulevard, Rockville, MD 20850-3129

Abstract

Address-based sampling (ABS) from the USPS Delivery Sequence File (DSF) presents a potential method to overcome historical coverage decreases in landline random digit dial (RDD) frames, and reduce costs relative to dual-frame cell/landline RDD samples. DSF coverage tends to be better in urban areas than rural areas, and the contact and response options available with ABS make it an interesting method for reaching otherwise “hard-to-survey” populations, such as ethnic and linguistic minorities, and households with only a cell phone. ABS was pilot tested as a special project of the California Health Interview Survey (CHIS), and the procedural results are presented here. Two California communities comprised the target population (with a sample of n=8,277 addresses). Communities were chosen based on population size/density, ethnicity, and percentage speaking Spanish. The mailing protocol included three “full-packet” mailings of a one-page screener form designed to replace screening traditionally done via phone. A reminder postcard was mailed between the first and second mailings. Households providing a phone number were called to complete the standard CHIS telephone interview. Households not providing a phone number were called if one was matched to their address through a record match. Initial results suggest that ABS may work as a replacement to RDD screening but with several caveats.

Key Words: address-based sampling, ABS, hard-to-reach populations, RDD

1. Introduction

Landline random-digit dial (LL-RDD) surveys have decreased precipitously in coverage of the general household population over the past decade as people and households rely more on cell phones (Blumberg & Luke, 2007, 2010; Blumberg et al., 2011, 2012). The California Health Interview Survey (CHIS) has moved to a dual-frame LL/cell-RDD sample design to reduce coverage error related to this trend. Yet, adding a second sampling frame increases the complexity of sampling, data collection, and weighting. Further, precise geographic targeting is difficult with cell-RDD samples because phone numbers and the phones themselves are not tied to a specific geographic location as reliably as landline numbers are. Thus, sampling households by sampling addresses has benefits over RDD sampling for surveys of small geographic areas. Further, address-based sampling (ABS) may lead to more efficient data collection if households can be contacted and screened by mail instead of telephone because no interviewers are involved

and geographic screening will be more efficient (i.e., fewer sample units that are not in the target area).

When considering an ABS design, phone surveys like CHIS need to decide whether to use an ABS approach for contact and screening only or for data collection as well. Several factors make transitioning large phone surveys like CHIS to a mail mode. CHIS is conducted in five languages, has complex skip patterns, and interviews multiple randomly-selected people in each household, each of which makes computerized interviewer-administered options a better choice. The research question for this pilot test was “Can we (and how do we) use a mail mode to obtain phone numbers at which to call sampled addresses?”

The ABS pilot test reported here has the following goals:

- 1) Test the logistics and feasibility of using an ABS sample to screen households and obtain phone numbers by mail and computer match, but conduct interviews by phone.
- 2) Evaluate how well the cell phone only (CPO) population can be reached by this sequential multi-mode hybrid.
- 3) Compare results across different types of communities, specifically those with large Spanish-speaking populations

2. Methods

2.1 Sample Design¹

Two communities were chosen for the pilot from the 14 communities that are part of the Building Healthy Communities (BHC) program funded by The California Endowment (TCE). They were Boyle Heights in Los Angeles and parts of Merced County. Of the 14 communities available, some were excluded due to the number of CHIS contact attempts and interviews already conducted in the community. For the remaining communities, demographic and language use estimates were used to choose communities that provided variability in urbanicity, ethnicity, and language. A current version of the United States Postal Service Delivery Sequence File (DSF) was obtained from Marketing Systems Group (MSG) and served as the sampling frame. A total of 8,277 addresses were sampled across both communities. The primary geographic inclusion criterion was that the address must fall within the Census block groups that define each BHC area. Only 98 addresses that had been contacted by CHIS in the past were excluded. P.O. boxes were included, and made up 10.4% of the final sample. There were no addresses flagged as only receiving their mail at a P.O. box.

Table 1 shows how the two selected communities differ in demographic, geographic, language, and housing characteristics. Sample sizes in Table 1 are the number of sampled addresses before any eligibility screening was done based on the responses to the mail screener form. Screening in Boyle Heights followed standard CHIS eligibility criteria,

¹The data collection was carried out by Westat and design of the study was a collaboration between Westat and CHIS staff at the UCLA Center for Health Policy Research.

requiring only that there was an adult age 18 or older in the household. In Merced, BHC-specific eligibility criteria excluded adults age 41 or older without children (i.e., households were included if they had children or if they included any adult younger than 41 whether or not children were present). The BHC criteria are screening criteria that were used in previous data collection with BHC sites.

Table 1: Description of Sampled Communities

<i>Community Characteristic</i>	<i>Boyle Heights (n = 4,466)</i>	<i>Southwest Merced and East Merced County (n = 3,811)</i>
Location	Los Angeles	Central Valley
Description	One of the oldest LA neighborhoods in Central/East LA	Agricultural and small college town with the youngest UC campus
Adult age 18-39 years	35%	33%
Hispanic	92-98%	60-63%
Speak Spanish at home	84%	44%
Families below poverty line	33%	30%
Renter-occupied housing units	78%	54%
Eligibility criteria assessed from mail screener form and used to determine eligibility for calling	At least one adult in household age 18+	Excludes adults age 41+ without kids

Source: Estimates come from: a) www.HealthyCity.org (reporting Census 2010 data), BHC 2009 health profiles (<http://healthpolicy.ucla.edu/chis/bhc/Pages/default.aspx>), and 2008 Claritas data based on U.S. Census 2000 urbanized areas.

MSG attempted phone number matches for all sampled addresses, but all addresses were sent the mail screener form even if a phone number was found so that the return on each method of number retrieval could be measured.

2.2 Contact materials and protocols

Households were mailed a screener packet that included a cover letter signed by CHIS Director Dr. David Grant, a list of frequently-asked questions (FAQ), and a one-page screener form that included a phone number request. All materials mailed were in English and Spanish, printed on both sides to reduce the weight and perceived burden of the packet. Details about the screener packet contents are below in the order the respondent opening the packet would have seen them. The cover letter and forms were printed on 8.5"x11" paper and mailed in a 9"x12" windowed envelope. The envelope was printed in a multi-shaded blue, black, and white with a return address and logo that are used on other CHIS mailing material and the CHIS website. The return envelope was addressed to Westat.

Screener Packet Details

- 1) **Cover letter:** English version of the cover letter on CHIS letterhead and signed by CHIS Director Dr. David Grant. The Spanish cover letter was printed on the back and there was a Spanish phrase on the English side to inform Spanish speakers of the translation. A \$2 bill was glue tacked to the cover letter.
- 2) **Screener form:** Each mailing included an English and Spanish form. Forms were printed so that the screening questions were on the front and the FAQ was on the back. The top of the form included instructions for filling out the form and a phone number that could be used to call with questions or to complete the interview. The first 5 questions asked about health and health behavior topics that these questions would pique respondents' interest to continue further (Montaquila, J. M., Brick, J. M., et al., 2013). The next 5 question asked demographic questions required for screening. Finally, the last two questions asked for the language in which the household wanted to be contacted and phone numbers at which the household could be contacted (see Figure 1).
- 3) A pre-paid return envelope.

The California Health Survey is the primary source of data on health issues in California. It helps set healthcare policy in California.

We would like to call you to ask some more questions about your health.

9. In what language would you prefer to be interviewed when we call?

English

Spanish

10. Please provide at least one telephone number at which you can best be reached.

Home: - -

Cell: - -

Figure 1: Questions asking for phone numbers on ABS mail screener form

Figure 2. shows the overall packet mailing and reminder timeline. The study used a common mailing and reminder approach that included up to three full packet mailings and a reminder postcard. Merced received a supplemental sample toward the end of the field period to boost the completed sample size.

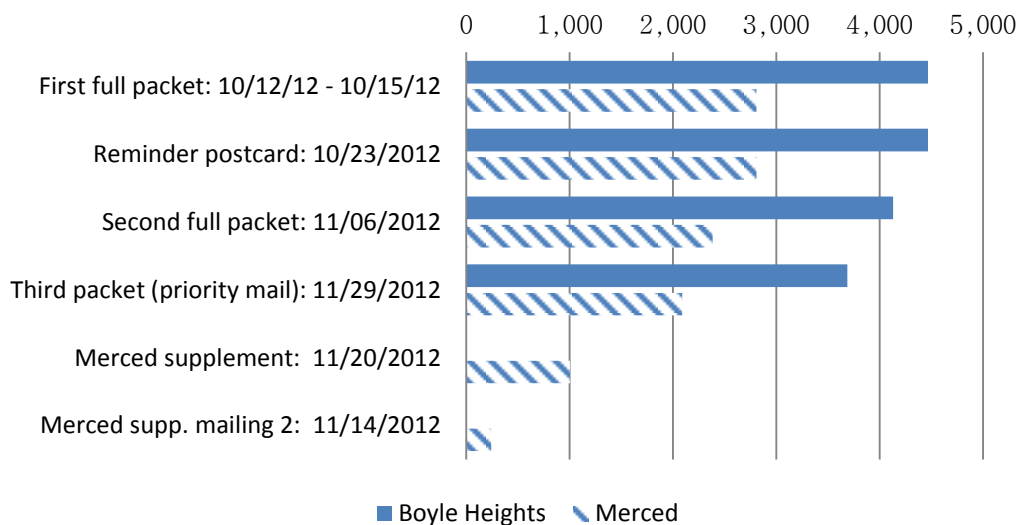


Figure 2. Screener form mailing timeline

2.3 Interview phase

Phone numbers that were returned on the form or obtained by matching were called for screening/verification and extended interview if the household was eligible (see Table 1 for eligibility criteria). Only slight changes from the usual CHIS protocols were made to the introduction and screener interview that were used once the respondent was on the phone. The extended interview was identical to what was used in CHIS 2011-2012 data collection. The CHIS adult interview asks about health insurance coverage, healthy and unhealthy behaviors, doctor and dental visits, general medical history and diagnoses, and other health-related topics. It takes about 35 minutes to complete on average. The interview is longer for those who have children, who are uninsured, or who have complex insurance situations, and for those with multiple health issues. In addition to the adult interview, CHIS conducts one interview with a sampled teen (12-17 years), and an interview with a parent or guardian about a sampled child (0-11 years). The child and teen interviews are much shorter than the adult interview. Interviews were conducted in English and Spanish.

3. Results

3.1 Mail-based screening results

To assess the effectiveness of the mail screening, Table 2 shows dispositions for all the 8,277 mailed addresses. The largest problem in mail screening was the high number of addresses never returning a form. It is interesting that even using the official delivery list of the USPS results in about 9% of mailings being returned as undeliverable, with fairly wide variation between about 7% in Boyle Heights and about 12% in Merced.

Table 2. Dispositions of addresses from mail screening phase

<i>Mailing Disposition</i>	<i>Southwest Merced and East Merced</i>		<i>Total</i>
	<i>Boyle Heights</i>	<i>County</i>	
Not returned	72% 3,194	65% 2,477	69% 5,671
Returned with a phone number*	18% 823	20% 748	19% 1,571
Returned Not Deliverable by USPS	7% 304	12% 465	9% 769
Returned partially-complete but no phone number	3% 124	3% 104	3% 228
Returned Blank	0.3% 13	0.3% 10	0.3% 23
Refused	0.2% 8	0.2% 7	0.2% 15
Total Mailed	100%** 4,466	100% 3,811	100% 8,277

* Only 160 addresses returned multiple forms. All numbers received from each address were kept.

** Percentages do not add up to 100% due to rounding

Of the 1,571 forms received with a phone number, 65% (n = 1,028) were completed in English, and 35% (n = 543) were completed in Spanish.

Figure 3 shows the percentage of addresses out of those eligible to be called that returned only a cell phone number, only a landline phone number, or both across all forms received from the address. A large proportion (almost 40%) gave us only a cell number, suggesting that the mail screener is reaching people who are at least “cell-mostly” if not cell-only.

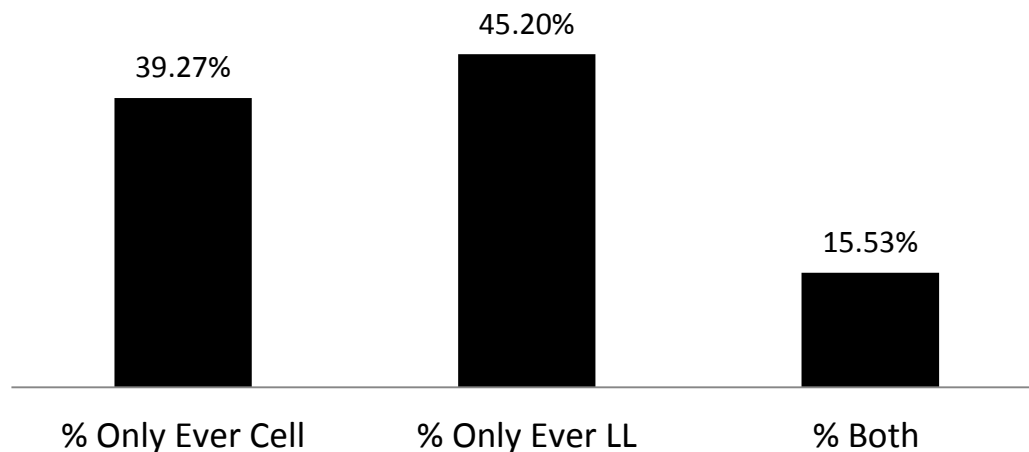


Figure 3. Percent of eligible mail returns providing each type of phone number (only cell, only landline, or both) across all mailings returned (i.e., accounts for addresses returning multiple forms). Percentages are out of n=1146 forms that were received, provided a phone number, and were considered eligible for the study. An address could return more than one form.

3.2 Phone number match

The phone number match rate for all sampled cases was 38%. Yet, in a mail-first design, the ultimate utility of the match process comes from numbers provided via match that were not obtained from the mail form. Of the addresses that did not provide a number on

the form (n=6,706), 37% (n=2,457) were matched to at least one number via directory search. This suggests that the mail form and computer match may be getting numbers for different people.

3.3 Interview results

A total of 3,880 households were called. 802 completed a phone screener. The remainder of the phone numbers were obtained by directory match. These totals exclude cases that were screened out after the mailing phase using the BHC eligibility criteria. By the end of data collection, 512 interviews were completed across the two communities. Table 3 shows the breakdown of interview type and total number of unique households by community.

Table 3. Percentage of interviews in each community that fall into the adult, child, and teen categories.

<i>Interview Type</i>	<i>Boyle Heights</i>	<i>Southwest Merced and East Merced County</i>
Adult Interview (n=337 ²)	72% 220	57% 117
Child Interview (n=141)	23% 70	34% ³ 71
Teen Interview (n=34)	5% 15	9% 19
Total Interviews in each community	100% 305	100% 207
Unique households with at least one interview (n=375) ⁴	241	134

Finally, one important question is “How well did this method represent the CPO population?” Of the 337 completed adult interviews, 28.4% (n=96) said that their household only had a cell phone. In Boyle Heights and Merced, 21% and 43% of households interviewed were CPO. There were an estimated 27.9% (s.e. 0.7) CPO households in California between January and December 2011 (Blumberg et al, 2012). The estimate for Los Angeles, which includes Boyle Heights was 28.3% (s.e., 1.3). Fresno County, which borders Merced County had an estimated 30.5% (s.e., 2.6) CPO households. Exact comparisons are difficult to make because of differences in geography between our sample and these estimates, and the additional screening done in Merced. However, it appears that the piloted ABS method captures most of the CPO population in these two communities.

4. Conclusions

4.1 Summary

This small ABS test provides some insight into how ABS may complement a large and complex state-wide RDD phone survey. In summary, the key findings at this point and initial interpretations are summarized below.

² Of adult interviews, n=336 were with sampled adults directly, and n=1 was by proxy.

³ BHC screening in Merced likely led to higher rates of child and teen interviews in Merced.

⁴ CHIS uses a “child first” procedure for completing the child interview if the most knowledgeable adult is available to be interviewed before the sample adult is available. Thus, at the close of the survey, an adult interview was not necessarily completed among all households. A “teen-first” interview could be completed if there was an sample child in the household.

- 1) The initial completion rate for the short screener form, including only those households that returned a form with a phone number, was 19% if non-deliverables are included in the denominator and 21% if they are excluded. That is much lower than our experience with screener cooperation rates in either LL-RDD or cell-RDD.
- 2) The large portion of households that never returned the form (69% of forms mailed, 66% after excluding those returned as not deliverable by the USPS) is a major concern. The first step of the two-step contact and screening process should be revised.
- 3) The extended interview completion rate in the phone phase (42%, given screener completion)⁵ is not particularly low, but this is not a final response rate. Further, combining the completion rates of both stages gives an overall rate of 10%. Such rates are not unheard of in some RDD surveys.
- 4) Comparing ABS performance results between communities at the screener stage shows that a larger percentage of addresses never returned a packet in Boyle Heights, which is urban, has more Hispanics and Spanish speakers, and has more renter-occupied housing units than Merced. Yet, Boyle Heights also had half the number of undeliverable screener packets as Merced. There was a small difference in the number of addresses returning a phone number (18% v. 20% for Boyle Heights and Merced respectively). When undeliverable addresses are removed (i.e., those that never had a chance for a respondent to open and complete), those rates become 20% in Boyle Heights and 22% in Merced. Using these adjusted return rates, it appears that the form may have worked slightly better in Merced than in Boyle Heights.
- 5) Comparing the number of CPO households we interviewed with those reported elsewhere shows that this method generally reaches this growing population. External estimates of the CPO population are not geographically precise enough for us to make exact comparisons to our results.

It is clear from these results that increasing the number of respondents who return the screener form will be essential for ABS designs with a mail screener and phone interview mode combination. To increase the form return rate, design changes or increases/changes in incentives will likely be necessary. Other mail screening ABS efforts have obtained higher better results than this pilot (Montaquila, J. M., Brick, J. M., et al., 2013). Without knowing whether respondents are opening the envelope it is hard to know exactly what to change to increase screener response rates. If the problem is that respondents aren't opening the packet, envelope redesign is a logical path to follow. If respondents are opening the packet (and at least skimming the materials), different changes would be in order. Without further experimentation or qualitative research about how the packet was perceived by the respondent, we can only speculate about what happened in the black box between mailing the form and receiving completed forms.

⁵ Calculated as the percent of completed adult interviews (n=337) out of completed screeners (n=802)

To increase overall response rates, not just screener form return rates, we could better use the matched phone numbers by calling earlier to either do the survey or at least remind the household to send back the form. The percentage of sampled addresses that were matched to a number, and the match rate based on households that did not return a phone number on the mail screener form (38% and 37% respectively), are both high. The characteristics of these households and quality of the phone numbers need to be investigated in more detail, but it seems clear that the phone match is an important part of an ABS design. Further research on the quality of those matches, how many were working numbers, and how many unique numbers were obtained by the match that were not obtained by the screener form will aid recommendations about how to best use these numbers to gain efficiencies in the overall design.

4.2 Limitations of the Pilot Test

This pilot test was conducted as a special study, with two unique communities, so comparisons should be made cautiously. We tested ABS in two communities that have characteristics of hard-to-survey populations (e.g., high poverty rates, large immigrant populations, low rates of English proficiency, and many young people who are highly mobile). While we can compare the two, the findings should be cautiously generalized to other types of communities. We also did not present any statistical tests of differences, so comparisons need to be interpreted in that light.

Further, only one way of doing ABS was tested, so inferences about performance are limited to a multi-mode (mail screener with phone interview) design that used a \$2 incentive, with a mail-first approach for all addresses including those that were matched to a phone number, and that interview in only English and Spanish. Generalizations of these results to ABS designs with other features should be made carefully as well.

Finally, the scheduling of the field dates was not ideal. The mailings went out between October 10, 2012 and December 14, 2012, and calling happened between November 13, 2012 and January 13, 2013. This was the beginning of the holiday season of an election year. People are busier during both of these societal events, and receive a lot of mail and many phone calls from fundraisers and pollsters. Respondents likely experience request fatigue that can affect surveys as well during periods like this.

The body of ABS research is still young and future tests must examine the role of each of the components independently to help inform ABS survey designs, particularly designs for RDD surveys that wish to move to ABS. Turning to research on mail surveys more generally, and on the social and psychological processes behind decisions to participate may help better understand response mechanisms operating in ABS designs.

Acknowledgements

The authors thank The California Endowment for their continuing support of CHIS, and specifically for the support of this ABS pilot test.

References

- Blumberg, S. J., & Luke, J. V. (2007). Coverage bias in traditional telephone surveys of low-income and young adults. *Public Opinion Quarterly*, 71(5), 734–749. doi:10.1093/poq/nfm047
- Blumberg, S. J., & Luke, J. V. (2010). Wireless substitution: early release of estimates from the National Health Interview Survey, January–June 2010. *National Center for Health Statistics. December*. Retrieved from <http://www.literacynewjersey.com/nchs/data/nhis/earlyrelease/wireless200912.pdf>
- Blumberg, S. J., Luke, J. V., Ganesh, N., Davern, M. E., Boudreaux, M. H., & Soderberg, K. (2011). *Wireless substitution: State-level estimates from the National Health Interview Survey, January 2007-June 2010*. US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics. <http://www.cdc.gov/nchs/data/nhsr/nhsr039.pdf>
- Blumberg, S. J., Luke, J. V., Ganesh, N., Davern, M. E., & Boudreaux, M. H. (2012). *Wireless substitution: State-level estimates from the National Health Interview Survey, 2010- 2010*. US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics. Retrieved from <http://www.cdc.gov/nchs/data/nhsr/nhsr061.pdf>
- Montaquila, J. M., Brick, J.M., Williams, D., Kim, K., and Han, D. (2013). A Study of two-phase mail survey data collection methods. *Journal of Survey Statistics and Methodology*, 1(1), 66-87.)