The ABS Frame: Quality and Considerations

Shelley Brock Roth[1], Daifeng Han[1], Jill M. Montaquila[1]
[1]Westat, 1600 Research Boulevard, Rockville, MD 20850

Abstract:
Address based sampling (ABS) is increasingly being used in the current survey environment of declining response rates and coverage concerns associated with landline random digit dial surveys. The goals of this research are to determine the implications of using ABS frame variables for data collection, to evaluate data quality of demographic variables provided on ABS frames, and to examine the possibility of using ABS frame variables to guide survey design. This research will focus on ABS with mail as the mode of collection, although some of the findings will be pertinent to ABS surveys using other modes of data collection. The implications of using ABS on data collection will be evaluated by examining the association of frame data with eligibility rates, response rates, and characteristics of survey respondents. Quality of frame data will be evaluated by examining rates of availability of these items and by comparing these data to information reported by respondents. The ability to use ABS frame information for stratification and to guide operations will also be considered in light of the findings of these evaluations.

Key Words: Address based sampling, mail survey, data quality

1. Introduction

With landline coverage rates and response rates to random digit dial (RDD) surveys declining, survey practitioners are increasingly considering address based sampling (ABS) using residential address frames derived from the U.S. Postal Service (USPS) Computerized Delivery Sequence (CDS) file. The ABS frames maintained by reputable survey sampling frame vendors provide nearly complete coverage of residential addresses in the U.S. (Iannacchione 2011). Furthermore, the ability to mail to the addresses, to locate the addresses for in-person interviewing (with the exception of P.O. box addresses and the small proportion of rural route, highway contract, and simplified addresses), and to append landline telephone numbers to the addresses offer a variety of choices in data collection methods.

While the acronym "ABS" is generally used to refer to a survey design that involves using the USPS address lists as the basis for the sampling frame, various data collection methods have been used for ABS, with varying degrees of success. In this paper, we consider a two-phase design that uses mail for the first and second phase of data collection for all households. This approach was tested in the 2009 Pilot Study of the National Household Education Surveys Program (NHES), with results indicating higher response rates than the most recent NHES landline RDD survey (Brick, Williams, and Montaquila 2011). To investigate the effectiveness of alternative approaches within a two-phase ABS context, a very large-scale methodological field test was conducted for NHES in 2011.

1.1 Research Goals

The goals of this research were to determine the implications of using ABS frame information to make decisions about mail data collection procedures, to evaluate the quality of the demographic information provided by a reputable data vendor on the ABS frame, and to examine the possibility of using these appended demographics to guide survey design. Such information could be used for stratification purposes at the sample design stage, to provide additional mode of contact (through matched landline telephone numbers) that could help to improve response during the data collection process, or to examine potential nonresponse bias after data collection.

The ABS frame information available for this analysis included standard USPS characteristics of the addresses as well as information appended by the vendor, as shown in Exhibit 1. The availability of information on an ABS frame can vary depending on the vendor providing the list and the procedures that the vendor uses to append information. General postal (standard USPS) information is provided in the same or similar way regardless of the frame vendor, provided the vendor has a computerized delivery sequence (CDS) file license. However, appended items such as telephone number, surname, and additional demographics can vary among vendors in the way the information is collected and presented. These types of items are appended through address matching so they are address level (not person level) characteristics. The sample used for the research presented here was provided by Marketing Systems Group (MSG). We specifically included and examined a set of appended demographic characteristics that were available both on the ABS frame and from NHES.

Exhibit 1: Address and Appended Demographic Characteristics from the ABS Frame

Variable type	*Description*
Standard USPS	Carrier route (PO box, city contract, highway delivery/rural route) Seasonal address Vacant address Type of dwelling unit (multi- or single-family) Drop point address (single address serving several residences)
Appended	PO box only way to get mail Telephone number Surname Presence of children in the household Educational attainment Ethnicity Gender Annual household income Home tenure

In this paper, we explore the association of the ABS frame information with address eligibility rates, response rates, and characteristics of survey respondents. Quality of the ABS frame data is evaluated by examining missing rates of several of these items and by comparing these data to information reported by respondents. The ability to use ABS frame information for stratification is evaluated by examining the effect of oversampling targeted subgroups on nominal and effective yield for the subgroup, and on overall effective yield.

DiSogra et al. (2010) examined correlations between ancillary data from a probability-based nationally representative ABS sample and self-reported survey data from that same sample, finding that these data have some value for examining nonresponse in surveys

and possibly for mail strategies that target certain demographic characteristics. Amaya et al. (forthcoming) researched various address type flags for potential use in enhancing frame construction and sampling. This research expands on these studies by focusing on how specific address types and demographic characteristics that can be appended to addresses are related to eligibility and response rates, as well as examining further how they may be used to improve data collection efforts.

1.2 Overview of the NHES:2011 Field Test

Sponsored by the National Center for Education Statistics (NCES), NHES is a set of periodic education surveys. NHES surveys have targeted various subgroups, including infants, preschoolers, school-age children, and adults, and have covered various topics related to the care and education of children and adult education. The surveys included in the NHES:2011 Field Test targeted two subgroups: school-aged children (those in kindergarten through twelfth grade) were eligible for the Parent and Family Involvement in Education questionnaire (PFI) and children ages six years or younger who had not yet started kindergarten were eligible for the Early Childhood Program Participation questionnaire (ECPP). The Field Test contained a nationally representative sample (n = 41,260), along with a supplemental sample of addresses that were more likely to contain Spanish speakers (n = 18,740). All of the analyses presented here are restricted to the nationally representative sample.

For the NHES:2011 Field Test, the first phase of the survey was to determine whether a sampled address had any eligible children. Items collected in the NHES screener included the age, gender, school enrollment, and grade for each person of age 20 or younger living at the sampled address. Any first-phase responding household with at least one child eligible for either the PFI or ECPP survey was sampled for a second phase or topical survey. Only one child was sampled per household, and the appropriate topical questionnaire was sent to the parents/guardians of the sampled child.

The data collection lasted five months with an approximately one-month overlap between the screener and topical phase. Up to three questionnaire mailings were administered in each phase. A series of experiments were included aiming to increase response rate and decrease potential nonresponse bias (Montaquila, Brick, and Kim 2012).

2. Implications of Standard ABS Frame Information on Data Collection

2.1 Address Eligibility and Topical Respondent Child Characteristics

Address eligibility rates were computed by three ABS address characteristics in order to determine the potential implications on mail data collection procedures. For NHES, an address was considered ineligible if no completed screener was received *and* either a screener mailing was returned as non-deliverable or was identified as a business address. The overall address eligibility rate was 88 percent and was computed as the total number of eligible addresses divided by the total number of sampled addresses.

It may be tempting to use variables such as vacant and seasonal indicators to restrict the frame, perhaps by excluding such addresses. Here, we examine the potential consequences of such restrictions. Table 1 shows the NHES address eligibility rates for addresses flagged as seasonal, vacant, or drop point. Most notable is the low eligibility rate of 18 percent for addresses that were flagged as vacant. However, eligibility rates for addresses flagged as seasonal or drop point are fairly high. Additionally, the proportion

of these types of addresses is small, such that including them for coverage reasons may be worth the additional cost of fielding them, depending on survey budget and goals.

Table 1: Association of ABS Address Characteristics with Weighted Address Eligibility Rate

ABS Frame Characteristic		*Total number of addresses*	*Percent of addresses*	*Address eligibility rate*
Seasonal address	Yes	310	1	69 (2.5)
	No	40,950	99	88 (0.2)
Vacant address	Yes	2,620	6	18 (0.7)
	No	38,640	94	93 (0.1)
Drop point address	Yes	740	2	89 (1.3)
	No	40,520	98	88 (0.2)

NOTE: Standard errors are given in parentheses.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2011 Field Test of the National Household Education Survey (NHES).

For addresses flagged as vacant or drop point, a set of characteristics of sampled children as reported by the topical respondents was examined and tested for overall statistical significance using chi-square tests (seasonal addresses were excluded from further analysis due to small sample size). These child characteristics were chosen because they reflect the demographic and socio-economic status (SES) of the respondents, and may be correlated with the key measures of the survey. The goal was to determine whether the topical respondents differed significantly for these types of addresses, which are often believed to increase costs and/or complicate survey operations.

Table 2: Characteristics of Children, by Vacant and Drop Point Address Indicators

	Vacant		*Drop point*	
Characteristic	*Yes (% of children)*	*No (% of children)*	*Yes (% of children)*	*No (% of children)*
Born in the U.S.	97 (1.6)	96 (0.3)	96 (2.2)	96 (0.3)
Hispanic origin	14 (4.6)	18 (0.6)	26 (7.9)	18 (0.6)
Not English only spoken at home	14 (4.6)	20 (0.7)	46** (8.9)	19 (0.7)
Household size of 6 or more members	22 (11.1)	18 (0.9)	36* (9.5)	18 (0.9)
Lived at address for less than 2 years	83** (4.7)	25 (0.7)	28 (9.3)	26 (0.8)
Parents highest education is less than high school	4 (2.3)	7 (0.4)	9 (7.1)	7 (0.4)
Cell phone only household	39* (7.5)	25 (0.7)	20 (7.6)	25 (0.7)
Income of under $40,000 per year	54* (9.7)	33 (0.9)	56** (7.7)	33 (0.9)
Rent/other	69** (6.9)	26 (0.7)	38 (8.2)	26 (0.8)
Single parent household	35 (8.1)	22 (0.7)	39* (9.4)	22 (0.7)

*Chi-square p value <0.05.
**Chi-square p value <0.01.
NOTE: Standard errors are given in parentheses. Significance test results are indicated in the "Yes" columns.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2011 Field Test of the National Household Education Survey (NHES)

Table 2 above shows that children living at addresses flagged as vacant or drop point are likely to differ from other children on many of these characteristics. Although excluding these types of addresses may simplify survey procedures and/or lower costs, there is also potential for bias, given these differences.

2.2 Association of Frame Characteristics and Response Rate

Another potential use of variables available for ABS samples (both USPS address characteristics and other variables that can be appended to the sampled addresses) is in weighting adjustments for nonresponse and/or for nonresponse bias analyses. In this section, we examine the extent to which these variables were associated with nonresponse to the screener request in the NHES:2011 Field Test. The screener response rate is a household-level rate calculated among all of the households (with or without children). The overall weighted screener mail response rate for the NHES:2011 Field Test was 69 percent.

Table 3 shows screener response rates by standard USPS address characteristics as well as three of the appended characteristics. All of the characteristics in this table were found to be significantly associated with screener response (chi-square p value <0.05). In a multivariate analysis using NHES:2011 Field Test screener respondents, carrier route type, dwelling unit type, vacant and seasonal address indicators, and telephone number match were all significantly associated with screener response. In addition to their potential use in post data collection weighting adjustments or analysis, these variables can also be used during data collection in a responsive design context (Groves 2006, Wagner 2008); by monitoring response rates by variables available for the entire sample, actions can be taken during data collection in an effort to limit the variation in response rates, with the goal of limiting nonresponse bias.

Table 3: Association of ABS Address Characteristics with Weighted Screener Response Rates

Characteristic		*Total number of addresses*	*Percent of addresses*	*Screener response rate*
Carrier route type	PO Box	4,590	11	61 (1.0)
	City delivery	24,350	59	68 (0.3)
	Highway contract/rural route	12,320	30	73 (0.4)
Dwelling unit type	M: multi-family	9,000	22	59 (0.6)
	S: single family	27,680	67	73 (0.2)
	Missing	4,590	11	61 (1.0)
Seasonal address	Yes	310	1	80 (2.8)
	No	40,950	99	69 (0.2)
Vacant address	Yes	2,620	6	77 (1.3)
	No	38,640	94	69 (0.2)
Drop point address	Yes	740	2	59 (2.0)
	No	40,520	98	69 (0.2)
PO box only way to get mail	Yes	1,590	4	59 (1.6)
	No	39,670	96	69 (0.2)
Telephone match	Yes	17,320	42	76 (0.3)
	No	23,950	58	63 (0.3)
Surname available	Yes	33,490	81	71 (0.2)
	No	7,780	19	53 (0.6)

NOTE: Standard errors are given in parentheses.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2011 Field Test of the National Household Education Survey (NHES).

Although response rates are widely viewed as essential quality measures, nonresponse bias is present only if the nonrespondents differ from the respondents. To examine this,

we analyzed characteristics of the topical respondents by characteristics available either on the ABS frame or appended to the sampled addresses.

Many significant differences were found for the topical respondent characteristics across all of the address characteristics (results not shown). In general, lower SES indicators were found to be more likely to be associated with P.O. Box addresses or multi-family dwelling types, as also noted earlier for vacant and drop point addresses. Higher SES among respondents is associated with availability of telephone number matches and surnames.

3. Quality of Additional Appended Variables

3.1 Availability of Telephone Number and its Association with Respondent Characteristics

As shown in Table 3, 42 percent of addresses in the NHES sample had a telephone number matched to the address, and the screener response rate for these matched cases was 76 percent. For the 58 percent of addresses without a telephone number match, the screener response rate was 63 percent. Whether a telephone number can be matched to an address is a highly significant predictor of response propensity, even when the mode of administration is mail, as in the NHES:2011 Field Test. Table 4 below presents various child- and household-level characteristics for addresses with and without telephone number matches. Many statistically significant differences were observed; not obtaining a telephone number match to the address is generally associated with lower SES.

One data collection model used by survey practitioners is to make efforts to try to obtain a telephone number for the non-matched addresses and then attempt to conduct the interview by phone for all addresses. However, because many of the characteristics of addresses with and without a telephone number match differ substantially, large differences in response rates could result in nonresponse bias. As a result, it is important to use procedures that aim to limit the differences in response rates between the matched and non-matched cases to the extent possible.

Table 4: Estimated Percent of Children (and their households) having Various Characteristics, by Telephone Number Match Status

	Telephone number match	
Characteristic	*Yes (n=2,800) (% of children)*	*No (n=2,790) (% of children)*
Born in the U.S.	96 (0.5)	96 (0.4)
Hispanic origin**	14 (0.8)	22 (0.9)
Not English only spoken at home**	17 (0.8)	23 (1.0)
Household size of 6 or more members	17 (1.1)	19 (1.4)
Lived at address for less than 2 years**	16 (0.8)	36 (1.2)
Parents highest education is less than high school*	5 (0.6)	9 (0.7)
Cell phone only household**	10 (0.9)	40 (1.1)
Household income of under $40,000 per year**	25 (1.0)	42 (1.3)
Rent/other**	15 (0.8)	38 (1.1)
Single parent household**	17 (0.9)	27 (1.1)

**Chi-square p value <0.01.

NOTE: Standard errors are given in parentheses.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2011 Field Test of the National Household Education Survey (NHES).

It has been shown that reverse-matching rates and quality of telephone number matches can vary by different types of addresses (Amaya et. al. 2010). Since the NHES:2011 Field Test mail screener included a question requesting a household telephone number, the data collected can be compared to the ABS matched telephone number to assess how asking for a telephone number on a mail screener might enhance the ability to call sampled households to conduct an extended interview. The set of cases used for this comparison was restricted to screener responding households with children, i.e., the households eligible for the topical survey. The ABS matched telephone number was compared to telephone number provided by screener respondents, resulting in five possible outcomes as shown in Table 5. For respondents for whom the screener telephone number is different from the ABS matched telephone number, there is no indication of which, if either (or both), telephone number is correct. Most notable among these results is that for 35 percent of the cases under evaluation, telephone numbers were provided by screener respondents for whom there was no matched ABS telephone number available.

Table 5: Telephone Availability by Source for Screener Responding Households with Children

Source of telephone number	*Number of households*	*Percent of households*
No telephone number available	1,350	17 (0.4)
ABS phone number, no Screener phone number	1,160	15 (0.3)
Screener phone number same as ABS phone number	1,430	14 (0.5)
Screener phone number different from ABS phone number	1,110	18 (0.3)
Screener phone number, no ABS phone number	2,760	35 (0.6)
Total	7,810	100

NOTE: Standard errors are given in parentheses.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2011 Field Test of the National Household Education Survey (NHES).

For topical respondents, telephone number availability was re-categorized as shown in Table 6 (this is a collapsed version of the categories shown in Table 5). The first category contains cases where information was available from the ABS but no new information was obtained from respondents. The second category reflects cases for which all or some new telephone number information was obtained from screener respondents, and the third category contains cases for which there is no telephone number information. The association of various characteristics with this collapsed variable was examined using chi-square tests.

For characteristics for which the association was found to be significant, pairwise t-tests for all three combinations of availability of telephone number were conducted. Eight of the ten characteristics were significantly associated with availability of telephone number. Of the eight sets of pairwise comparisons, for households in the first category (ABS information), significance was observed for all eight characteristics when comparing to households in the second category (all or some new information obtained from respondents). In other words, asking for a phone number in the screener brought in significantly more Hispanic, non-English only, newer, lower education, lower income, cell phone only, non-owned, or single parent households, compared to households for which a phone number is already available via ABS information. Similar patterns of significance were observed for seven of the characteristics for households where no telephone number was available compared to households in the first category (ABS information available). Finally, the second category (all or some new information from

respondents) was significantly different from the set of cases where no telephone number was available in terms of speaking a language other than English at home.

These results indicate that in an ABS mail survey, many characteristics of survey respondents vary by the availability of telephone numbers (obtained through reverse-matching or from respondents). For surveys using telephone to administer the second-phase survey after fielding a screener by mail, asking respondents for a telephone number on the screener (in addition to appending telephone numbers to sampled addresses through reverse-matching) will likely result in better coverage and response to the second-phase survey than using only the appended telephone numbers.

Table 6: Estimated Percent of Children (and Their Households) Having Various Characteristics, by Telephone Number Availability and Source

Characteristic	*1- Screener phone number same as ABS phone number, or ABS phone number only (% of children)*	*2- Screener phone number different from ABS phone number, or screener phone number only (% of children)*	*3- No phone number available (% of children)*
Born in the U.S.	96 (0.6)	95 (0.4)	97 (0.5)
Hispanic origin**[ab]	12 (0.9)	21 (0.9)	21 (1.8)
Not English only spoken at home**[ac]	15 (1.0)	24 (1.1)	19 (1.5)
Household size of 6 or more members	16 (1.1)	19 (1.3)	18 (2.3)
Lived at address for less than 2 years**[ab]	14 (0.9)	32 (1.2)	35 (2.1)
Parents highest education is less than high school**[ab]	4 (0.6)	9 (0.8)	7 (1.1)
Cell phone only household**[ab]	6 (0.7)	35 (1.1)	40 (2.0)
Household income of under $40,000 per year**[ab]	21 (1.1)	40 (1.3)	41 (2.0)
Rent/other**[ab]	12 (0.8)	34 (1.0)	34 (2.1)
Single parent household**[ab]	15 (0.9)	26 (1.1)	27 (2.0)

**Chi-square p value <0.01.
[a]Category 1 vs. category 2 t-test result is significant.
[b]Category 1 vs. category 3 t-test result is significant.
[c]Category 2 vs. category 3 t-test result is significant.
NOTE: Standard errors are given in parentheses.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2011 Field Test of the National Household Education Survey (NHES).

3.2 Availability of Demographic Information

Here we examine the demographic variables appended to the sampled addresses for completeness. Completeness of this kind of information may be associated with some latent characteristics of the household and these characteristics may be associated with response rates. The variables examined include educational attainment, ethnicity, gender, household income, and home tenure. It is important to note that all of these variables are provided at the address level. Table 7 shows missingness rates for these variables, for the entire NHES sample and for just the NHES respondents. For the overall sample, both educational attainment and ethnicity had the highest missingness rates of 43 percent, and gender and household income had the lowest missingness rates of 20 percent. Similar

patterns were observed among NHES respondents, but with missingness rates consistently lower than among the sample, indicating that availability of this information is associated with screener response propensity. Similar results were observed for these demographics by DiSogra (2010).

Table 7: Missingness Rates of Demographic Items Appended to Addresses

	Entire NHES sample		*NHES respondents only*	
Characteristic	*Sample size*	*Percent missing*	*Sample size*	*Percent missing*
TOTAL	*41,260*		*5,590*	
Education	17,950	43 (0.2)	1,680	30 (0.7)
Ethnicity	17,950	43 (0.2)	1,680	30 (0.7)
Home tenure	10,760	26 (0.2)	780	14 (0.6)
Gender	8,420	20 (0.2)	500	9 (0.5)
Household income	8,320	20 (0.2)	490	9 (0.5)

NOTE: Standard errors are given in parentheses.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2011 Field Test of the National Household Education Survey (NHES).

3.3 Concordance of Demographic Information Appended to ABS Addresses with Information Provided by Survey Respondents

In this section, we examine the quality and accuracy of the variables that may be appended to the ABS addresses. We begin by examining presence of children. Both true and false positive rates were computed. The true positive rate, defined as the percent of screener responding households with children where the appended variable also indicated children was 41 percent. The false positive rate, defined as the percent of screener responding households without children where the appended variable indicated children was 31 percent. These rates suggest that presence of children as indicated by the appended variable is highly inaccurate and should be used with caution in survey design or data collection procedures. One minor caveat of this analysis is that in NHES, presence of children (to determine eligibility) is defined as having at least one child age 20 or younger who is not yet enrolled in school or is enrolled in grades K-12, whereas the appended variable indicates addresses thought to have at least one child age 18 or younger (regardless of enrollment status).

Besides presence of children, for four items collected in the topical questionnaire, it was possible to measure concordance of the respondent-provided value with the value for the appended item. These characteristics include parents' educational attainment (5 levels), Hispanic ethnicity, household income (4 levels), and home tenure. Concordance was measured both including and excluding the cases with missing data for the appended variable, resulting in different denominators for the calculations. When cases missing the appended variable are included, the denominator consists of all NHES:2011 national sample topical respondents and the match rate is referred to as "true match rate" below. This rate measures the proportion of cases for which the item could be appended to the address and the appended value matched the respondent report. The second rate is the rate at which the provided appended value matched the respondent report; even if this rate is high, the appended item could have very limited utility if its missing rate is high.
Table 8 gives both rates for each of the items included in this analysis. When cases with missing data for the appended variable are excluded, Hispanic ethnicity matches about 92 percent of the time; however, the true match rate is only 64 percent. Among cases with

non-missing appended data, home tenure matches about 87 percent of the time, but the true match rate is only 75 percent. Household income and educational attainment have lower concordance. These results indicate that in general, true match rates between NHES respondents and ABS frame information are fairly low and therefore may be unreliable in guiding survey design.

Table 8: Concordance of Information From Appended Variables and Topical Respondents

	Concordance rate-all topical respondents[a]		*Concordance rate-topical respondents with non-missing appended item*[b]	
Characteristic	*Number of cases*	*Percent*	*Number of cases*	*Percent*
Educational attainment	5,590	26 (0.7)	3,910	37 (0.9)
Hispanic ethnicity	5,590	64 (0.8)	3,910	92 (0.5)
Household income	5,590	49 (0.8)	5,100	54 (0.8)
Home tenure	5,590	75 (0.7)	4,810	87 (0.7)

[a]Includes cases missing data for appended variable in denominator.
[b]Excludes cases missing data for appended variable from denominator.
NOTE: Standard errors are given in parentheses.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2011 Field Test of the National Household Education Survey (NHES).

4. Analysis of Appended Demographic Variables for Stratification Purposes

Six of the appended demographic variables were analyzed in order to determine their potential use for stratification purposes. These included: household annual income less than $40,000 or not; presence of children in the household or not; parents' educational attainment less than high school or not; home is rented or otherwise not owned; Hispanic ethnicity; and Hispanic surname availability. For each characteristic, strata for potential oversampling were defined using the appended variable, and NHES respondent data were used to compute misclassification rates. Cases with missing data for the appended variable were included in these analyses. For each variable, several scenarios involving different oversampling rates for the stratum containing the subgroup (based on the appended variable) were examined.

Three measures were computed to quantify the effects of oversampling: the overall design effect due to stratification, the nominal increase in yield for the subgroup, and the effective yield for the subgroup (which accounts for the design effect due to differential sampling as well as misclassification). Figures 1 through 6 plot each of the three measures as the relative sampling rate for the targeted subgroup increases from one to ten. (The non-targeted subgroup is sampled at a relative rate of 1 in all scenarios.) For each characteristic, the nominal yield for the subgroups can be increased by a factor of two or more by heavily oversampling the high-density stratum. However, for household income less than $40,000, poor results for effective yield are observed when sampling at twice the rate or higher for low income households. Similar patterns are seen for presence of children, parents' educational attainment, and home tenure. Slightly modest increases in effective yield for higher sampling rates may be possible when stratifying by Hispanic ethnicity or Hispanic surname.

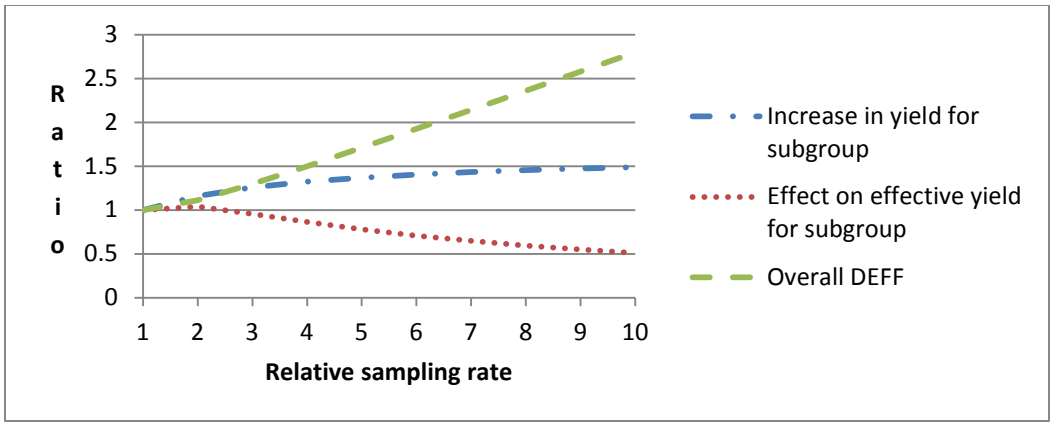

Figure 1: Effect of Oversampling on Yield: Annual Household Income less than $40,000

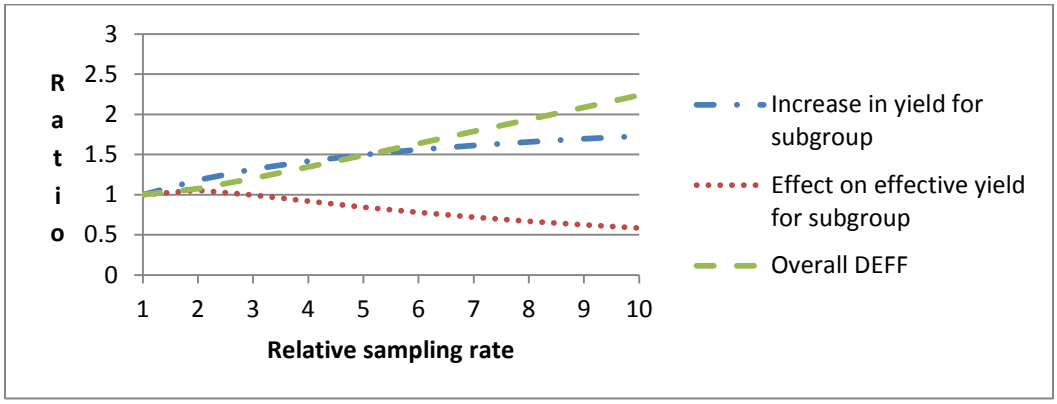

Figure 2: Effect of Oversampling on Yield: Presence of Children

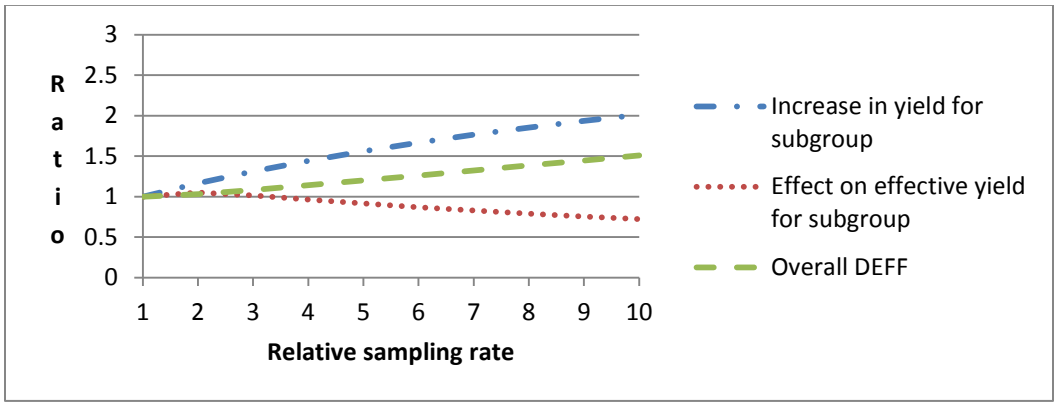

Figure 3: Effect of Oversampling on Yield: Educational Attainment less than High School

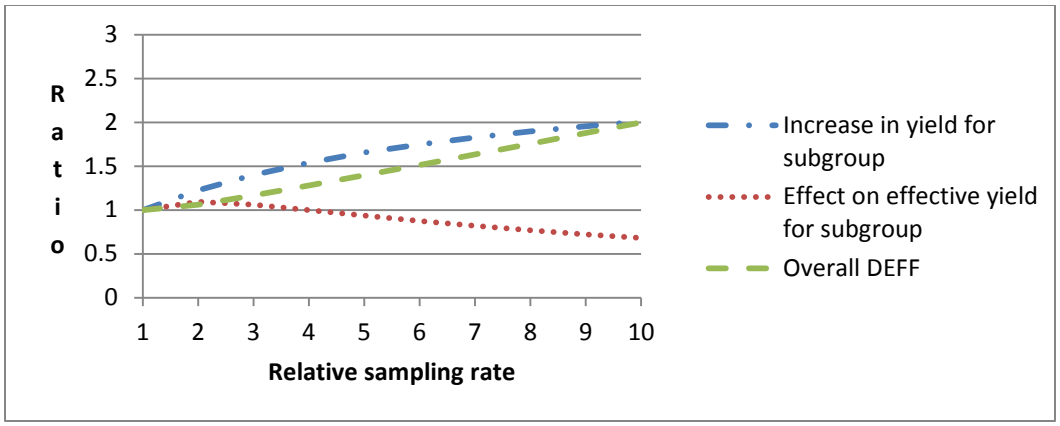

Figure 4: Effect of Oversampling on Yield: Home Tenure: Rented

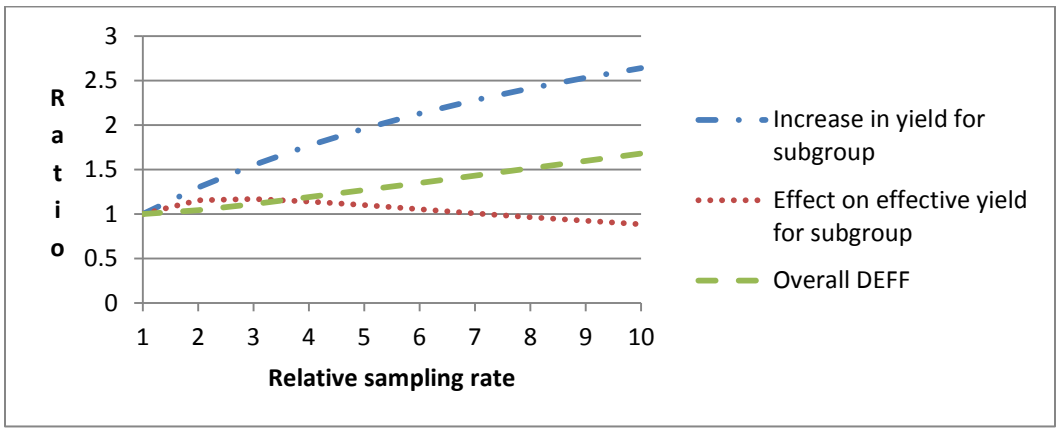

Figure 5: Effect of Oversampling on Yield: Hispanic Origin

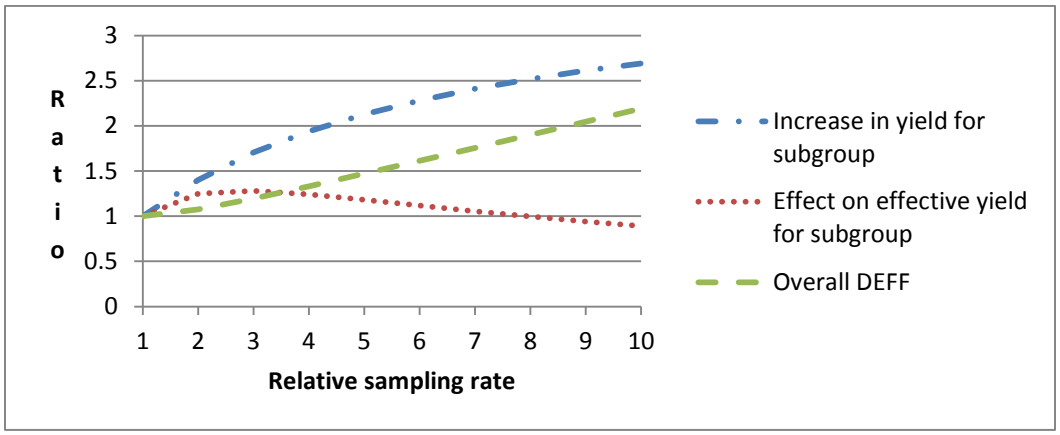

Figure 6: Effect of Oversampling on Yield: Hispanic Surname

The Hispanic ethnicity variable actually has three possible relevant outcomes: Hispanic ethnicity, ethnicity other than Hispanic, and missing. To determine whether the potential for stratification by Hispanic ethnicity could be improved by creating three separate strata defined as Hispanic, non-Hispanic, and missing, the same measures were computed using this three-stratum design; in this design, those identified as Hispanics are oversampled most heavily, those with missing ethnicity are sampled at a relative rate between 1 and the rate for Hispanics.

Table 9 shows the results when the relative sampling rate for Hispanics is held at three times the rate for non-Hispanics, the relative sampling rate for non-Hispanics is held at 1, and the relative sampling rate for missing cases is 1.1, 1.3, or 1.5. The results indicate that while further increases in nominal yield could be obtained for Hispanics when the missing data stratum is also oversampled, the overall design effect is much larger than in the two-stratum design. Therefore, if estimates for Hispanics are of particular interest and overall estimates are not a priority, this stratification scheme may be of use, but otherwise using the two-stratum design (Hispanics or not) may be a better choice. However, further research is needed to compare this approach to using aggregate data (e.g., tract-level percent Hispanic) for stratification. Also, the cost associated with appending demographics should be considered in light of the modest gains in effective sample sizes for Hispanics of these stratification approaches.

Table 9: Effect of Oversampling on Yield: Hispanic Origin Including Missing Stratum

Relative sampling rate for Hispanic category	*Relative sampling rate for "Missing" category*	*Increase in yield for subgroup*	*Effect on effective yield for subgroup*	*Overall DEFF*
3	1.1	1.54	1.19	1.18
3	1.3	1.52	1.23	1.31
3	1.5	1.51	1.24	1.45

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2011 Field Test of the National Household Education Survey (NHES).

5. Conclusions and Discussion

In this study, we evaluated information provided on or appended to an ABS sampling frame, with the goal of determining the viability of using this information to inform sample designs of future studies. First, we assessed eligibility of some types of undesirable addresses in terms of NHES screener respondents. We discovered that while eligibility is low for vacant addresses, it is higher for seasonal and drop point addresses. Given that these comprise a small proportion of addresses, gains in coverage of including them may justify the costs of fielding them, depending on the goals of the survey.

The correlations between topical characteristics of children and ABS address characteristics indicate that there are many differences among NHES respondents for many of these characteristics, so consideration of them may be useful for monitoring data collection efforts or identifying potential nonresponse bias. For example, some types of addresses (such as vacant, drop point, P.O. Boxes or multi-family dwelling units) are associated with lower SES. Household surveys tend to under-represent those with lower SES, so it is important to ensure adequate representation of these cases during data collection. This argues not only for including these types of addresses, but also for using the appended demographics to monitor response rates during data collection in an effort to limit nonresponse bias.

Telephone number availability is an important consideration in many surveys. Forty-two percent of NHES sample cases had telephone number matches on the ABS frame. When asked in the screener, 68 percent of screener respondents provided a telephone number, 15 percent had an ABS matched only telephone number, and 17 percent had no available

telephone number. Similar to other findings, households with lower SES were less likely to have ABS telephone number matches. When the topical respondent characteristics of children were examined for significance as related to telephone number availability and source, all but two of the characteristics showed significant differences. These results indicate that to reduce bias, it is important to have procedures that effectively elicit response from households for which a telephone number match is not available.

Another goal was to assess the availability and quality of variables that can be appended to ABS records. Among the variables we examined (educational attainment, ethnicity, gender, household income, and home tenure) we found that many have relatively high rates of missing values.

We measured concordance between the appended variables and the respondent reports for several key demographic variables. Of households with children in NHES, only 41 percent of them had presence of children indicated from the appended ABS variable. Among the characteristics educational attainment, Hispanic ethnicity, household income, and home tenure, only home tenure and Hispanic ethnicity had concordance rates that were better than what could be obtained with a flip of a fair coin. We examined the potential of using certain appended demographic variables for stratification purposes. By examining overall and subgroup design effects using NHES respondent data to determine rates of misclassification based on the appended variables, we concluded that stratifying by household income, presence of children, parents' educational attainment, or home tenure (and oversampling in the high-density stratum) does not improve the effective yield of respective subgroups. Stratifying by Hispanic ethnicity or surname and oversampling in the high-density stratum may result in modest increases in subgroup effective yield at the expense of overall effective yield. Further research is needed to compare this approach with stratification based on aggregate area-level (e.g., census tract-level) characteristics.

One limitation of this study is that validation information is only available from NHES survey respondents, which limits the set of variables that could be examined. Additionally, the respondent-provided information is treated as the "truth"; any measurement error in these responses is not accounted for in this analysis. Finally, the quality of the information provided on ABS frames may vary given that it can be provided by different vendors.

In summary, our investigation sheds some light on the quality and potential uses of the information that is routinely available on ABS frames, as well as the variables that may be appended to ABS records. The address information that originates from the USPS address list file may help to improve several aspects of household surveys, including data collection efforts, monitoring and reducing variations in response rates, and maximizing coverage and response. With the exception of Hispanic indicators (the ethnicity and surname variables), the appended demographic variables considered do not show much promise as having value for stratification purposes. The Hispanic indicators may warrant further examination, as do other appended variables not considered in our analysis.

References

Amaya, Ashley, Felicia Leclere, Lee Fiorio, and Ned English. Forthcoming. Improving the Utility of the DSF Address-Based Frame through Ancillary Information. Field Methods.

Amaya, Ashley, Ben Skalland, and Karen Wooten. 2010. “What’s in a Match?” Survey Practice, December: www.surveypractice.org.

Boyle, John, Anna Fleeman, Courtney Kennedy, Faith Lewis and Andy Weiss. 2011. “Sampling Cell Phone Only Households: A Comparison of Demographic and Behavioral Characteristics from ABS and Cell Phone Samples.” Survey Practice, February: www.surveypractice.org.

Brick, J. Michael, Douglas Williams, and Jill M. Montaquila. 2011. “Address-Based Sampling for Subpopulation Surveys.” Public Opinion Quarterly, Vol. 75, No. 3, 409-428.

DiSogra, Charles, J. Michael Dennis, and Mansour Fahimi. 2010. “On the Quality of Ancillary Data Available for Address-Based Sampling.” Proceedings of the Survey Research Methods Section of the Joint Statistical Meetings.

Gentry, Robin, Michelle Cantave, Nicole Wasikowski, and Yelena Pens. 2010. ‘‘To Mail or to Call: How to Reach the Hard-to-Reach.’’ Paper presented at the 65th Annual Meeting of the American Association for Public Opinion Research, Chicago.

Groves, Robert M., and Steven Heeringa. 2006. "Responsive design for household surveys: tools for actively controlling survey errors and costs." Journal of the Royal Statistical Society Series A: Statistics in Society, 169 (Part 3), 439-457.

Iannacchione, Vincent G. 2011. “Research Synthesis: The Changing Role of Address-Based Sampling in Survey Research.” Public Opinion Quarterly, Vol. 75, No. 3, 556-575.

Link, Michael W.,Gail Daily, Charles D. Shuttles, L. TracieYancey, Anh Thu Burks, and H. Christine Bourquin. 2009. ‘‘Building a New Foundation: Transitioning to Address-Based Sampling after 30 years of RDD.’’ Paper presented at the 64th Annual Meeting of the American Association for Public Opinion Research, Hollywood, FL.

Montaquila, J.M., J. M. Brick, and K. Kim. 2012. Transitioning from RDD to ABS in the National Household Education Surveys Program. Paper presented at the Joint Statistical Meetings, San Diego, CA.

Wagner J. 2008. Adaptive survey design to reduce nonresponse bias. Ph.D. Dissertation, University of Michigan.