8-2014

Assessing Within-Household Selection Methods in Household Mail Surveys

Kristen Olson
University of Nebraska-Lincoln, kolson5@unl.edu

Mathew Stange
University of Nebraska–Lincoln, s-mstange2@unl.edu

Jolene D. Smyth
University of Nebraska-Lincoln, jsmyth2@unl.edu

Published in *Public Opinion Quarterly* (2014); doi: 10.1093/poq/nfu022

Published online August 13, 2014.

A previous version of this paper was presented at the annual meeting of the American Association for Public Opinion Research, in Orlando, FL, USA, in May 2012.

Assessing Within-Household Selection Methods in Household Mail Surveys

Kristen Olson, Mathew Stange, and Jolene Smyth

University of Nebraska–Lincoln

Corresponding author — Kristen Olson, University of Nebraska–Lincoln, Department of Sociology, 703 Oldfather Hall, Lincoln, NE 68588, email kolson5@unl.edu

Abstract

Household surveys are increasingly moving toward self-administered modes of data collection. To maintain a probability sample of the population, researchers must use probability methods to select adults within households. However, very little experimental methodological work has been conducted on within-household selection in mail surveys. In this study, we experimentally examine four methods—the next-birthday method, the last-birthday method, selection of the youngest adult in the household, and selection of the oldest adult in the household—in two mail surveys of Nebraska residents (n = 2,498, AAPOR RR1 36.3 percent, and n = 947, AAPOR RR1 31.6 percent). To evaluate how accurately respondents were selected from among all adults in the household, we also included a household roster in the questionnaire for one of the surveys. We evaluated response rates, the completed sample composition resulting from the different within-household selection methods, and the accuracy of within-household selection. The analyses indicate that key demographics differed little across the selection methods, and that all of the within-household selection methods tend to underrepresent key demographic groups such as Hispanics and persons with lower levels of education. Rates of selection accuracy were low among the four selection methods analyzed, and the rates were similar across all four methods.

Introduction

Surveys are increasingly moving from interviewer-administered modes to self-administered modes, due in large part to the availability of an address-based frame for households in the United States (Iannacchione 2011). This move to self-administered modes presents a special challenge to survey researchers: how to maintain a probability sample of adults within a household without an interviewer to administer a within-household selection procedure. Although a wide body of literature exists for within-household selection methods for telephone surveys (Oldendick et al. 1988; Lind, Link, and Oldendick 2000; Gaziano 2005; Yan 2009), empirical research on within-household selection in mail surveys is much sparser (e.g., Reich, Yates, and Woolson 1986; Battaglia et al. 2008). The present study addresses this gap by using two experimental mail surveys to examine four within-household selection methods: the next-birthday, last-birthday, oldest-adult, and youngest-adult methods.

In comparison to telephone surveys, mail surveys require greater involvement of the household informant when a respondent is selected within a household. In a telephone survey, this task involves the interviewer who makes the selection and encourages the selected household member to participate in the survey. In a mail survey, the household informant opens the mail and must carry out the task of selecting a respondent, unaided by an interviewer. To complete the within-household selection task, the informant must (1) read the within-household selection procedure in the cover letter or on the questionnaire; (2) accurately follow the instructions to select the household member; and (3) persuade the selected individual to participate. If any of these steps are not accomplished, inaccurate within-household selection will occur and/or the survey will not be returned. In telephone surveys, anywhere from less than five percent (Troldahl and Carter 1964; O'Rourke and Blair 1983) to about 20 percent (Lavrakas, Stasny, and Harpuder 2000; Lind, Link, and Oldendick 2000) of respondents are incorrectly selected. Empirical examinations of within-household selection in mail surveys are limited; the few studies we know of find generally higher error rates, ranging from less than one percent (Gallagher, Fowler, and Stringfellow 1999) to about 14 percent (Olson and Smyth 2014) to greater than 30 percent (Gallagher, Fowler, and Stringfellow 1999; Schnell, Ziniel, and Coutts 2007; Battaglia et al. 2008).

Regardless of the mode of data collection, these studies have used a variety of within-household selection methods, including the age-position and birthday methods. In the next- and last-birthday methods, the selected respondent is the person in the household with the birthday that is immediately upcoming (next) or has most recently occurred (last). These are quasi-probability methods of within-household selection, as opposed to probability methods such as Kish's household roster (Gaziano 2005). Although Battaglia et al. (2008) examined the next-birthday method

and Schnell, Ziniel, and Coutts (2007) examined the last-birthday method in mail surveys, to our knowledge no previous studies have compared these selection procedures in the same study. Other options for within-household selection are the youngest-adult or oldest-adult methods, which, as their names imply, identify the survey respondent as the person in the household who is the youngest or oldest adult, respectively, and are nonprobability methods. To our knowledge, only one mail survey has evaluated these methods (Olson and Smyth 2014), but they have not been compared to the birthday methods.

One important question for survey researchers is whether a particular within-household selection procedure yields higher or lower response rates than another procedure. Looking across multiple telephone surveys, Gaziano (2005) found that the Kish selection method had lower response rates than the last-birthday method, which in turn was lower than the youngest-male/oldest-female selection method. Response rates for other within-household selection procedures (e.g., next birthday, youngest male/oldest female, no selection) have few and mixed empirical evaluations. Two studies have examined response rates between the next-birthday and all-adult procedures in mail surveys (one study also included the any-adult procedure), finding that the next-birthday method yielded higher response rates than the all-adult method, but similar response rates as the any-adult procedure (Battaglia et al. 2008; Hicks and Cantor 2012). With little empirical research of within-household selection in mail surveys, this question remains open.

Another important question for survey researchers is whether a particular within-household selection procedure yields a better representation of the target population or higher rates of accurate selections. The representativeness of a sample that results from different within-household selection methods can be evaluated by comparing the demographic composition of the completed sample to that of the target population, using benchmark data such as the American Community Survey or the Current Population Survey. Information on household composition (e.g., was the respondent actually the youngest adult in the household?), can be used to evaluate how accurately a within-household selection technique was implemented within each household. For example, Battaglia et al. (2008) telephoned households that responded to a mail survey to evaluate whether the household completed the selection task successfully, and if not, why the errors occurred. Schnell, Ziniel, and Coutts (2007) used information from population-register data in Germany to assess the accuracy of within-household selection. We use a household roster embedded in the survey itself to determine whether the household informant selected the appropriate household respondent. We also look at characteristics of those respondents who are more likely to be selected accurately. Larger households and household members with lower education may be more likely to have the wrong member of the household respond to the survey, as might households with members having birthdays close to the reference date (Forsman 1993).

To understand how various within-household selection methods affect household mail surveys, we use data from two general-population mail surveys of Nebraska residents to ask the following questions:

1. Do response rates vary systematically across the next-birthday, last-birthday, youngest-adult, and oldest-adult within-household selection procedures in general-population mail surveys?
2. Does sample composition vary across these four selection procedures in general-population mail surveys?
3. Which of these within-household selection procedures yield samples that are closer to benchmark data about the target population?
4. Are within-household selections accurately made in general-population mail surveys, and does accuracy of selection vary across within-household selection procedures or respondent characteristics?

Data and Methods

Nebraska Annual Social Indicators Survey

The first survey is the 2011 Nebraska Annual Social Indicators Survey (NASIS), conducted by the Bureau of Sociological Research at the University of Nebraska–Lincoln. NASIS is a statewide omnibus survey asking questions on a range of topics, including attitudes about water quality, scientific knowledge, crime and the criminal justice system, media and television, vacation and travel, gender attitudes, and demographics. In 2011, the NASIS was administered by mail to a sample of 2,498 directory-listed landline households in Nebraska selected by Survey Sampling International (SSI). The survey contained 145 questions in a 12-page questionnaire booklet. The initial NASIS survey packet, including a questionnaire, cover letter, and postage-paid return envelope, was mailed on April 13, 2011. A reminder postcard was sent to all nonrespondents in all treatment groups about three weeks after the initial survey mailing. A second survey packet was sent to all remaining nonrespondents approximately 10 days after the reminder postcard, with a third survey packet mailed to remaining nonrespondents approximately five weeks later. Completed surveys were received from 906 respondents between April 19, 2011, and August 18, 2011, resulting in an AAPOR RR1 of 36.3 percent.

Nebraska Trees and Forest Survey

The second survey is the Nebraska Trees and Forest Survey (Forest Survey)—a statewide mail survey of Nebraska residents, fielded on behalf of the Nebraska Forest Service (NFS) by the Bureau of Sociological Research at the University of Nebraska–Lincoln. The Forest Survey asked Nebraska residents how they use trees and forests. It also asked demographic questions. The survey was sent to a

sample of 3,000 addresses selected from the United States Postal Service's Delivery Sequence File (USPS DFS) by Survey Sampling International (SSI). Addresses across the state of Nebraska were selected from three strata: City (n = 1,000)—addresses from zip codes in the Nebraska cities of Bellevue, Elkhorn, La Vista, Lincoln, Omaha, Papillion, and Ralston; East (n = 1,000)—addresses from counties in eastern Nebraska; and West (n = 1,000)—addresses from counties in western Nebraska. A total of 947 completed surveys were received. Three of these were omitted because the reported age of the respondent was under 19, the age of majority in Nebraska. The survey contained 19 questions in a four-page booklet. The initial survey packet containing a cover letter, questionnaire, and postage-paid return envelope was mailed on February 24, 2012. A reminder postcard was mailed to all nonrespondents approximately two weeks later, and a second survey packet (cover letter, survey, and postage-paid return envelope) was mailed to all remaining nonrespondents 10 days after the postcard. Data collection for the survey ended on May 2, 2012. The overall response rate was 31.6 percent (AAPOR RR1).

Experimental Treatments

The experimental treatments differed slightly in the two surveys. Each sampled address for the NASIS was randomly assigned to one of four different within-household selection treatment conditions: last birthday, next birthday, oldest adult, or youngest adult. Although the oldest-adult and youngest-adult treatment conditions are usually combined to obtain general-population estimates (Gaziano 2005), we will examine them separately here to identify whether they yielded different sample pools (as we would expect, at least on the age of respondent) and have different correlates of selection accuracy. Table 1 contains the respondent-selection instruction wording, which was provided in the first paragraph of the cover letter for each of the four treatment groups. The NASIS questionnaire itself was identical across all treatment groups.

Each sampled address in the Forest Survey was randomly assigned to receive one of two within-household selection treatments: next birthday or last birthday. Like the NASIS, the within-household selection instructions were presented in the cover letter (see table 1) and the survey questionnaires were the same across treatment groups.

Analyses

Our primary objectives are (1) to identify whether response rates differ across the household-selection methods; (2) to evaluate the composition of samples across the different selection methods; (3) to compare the completed samples produced by each selection method to ACS benchmark data; and (4) to examine whether or not selections are made accurately.

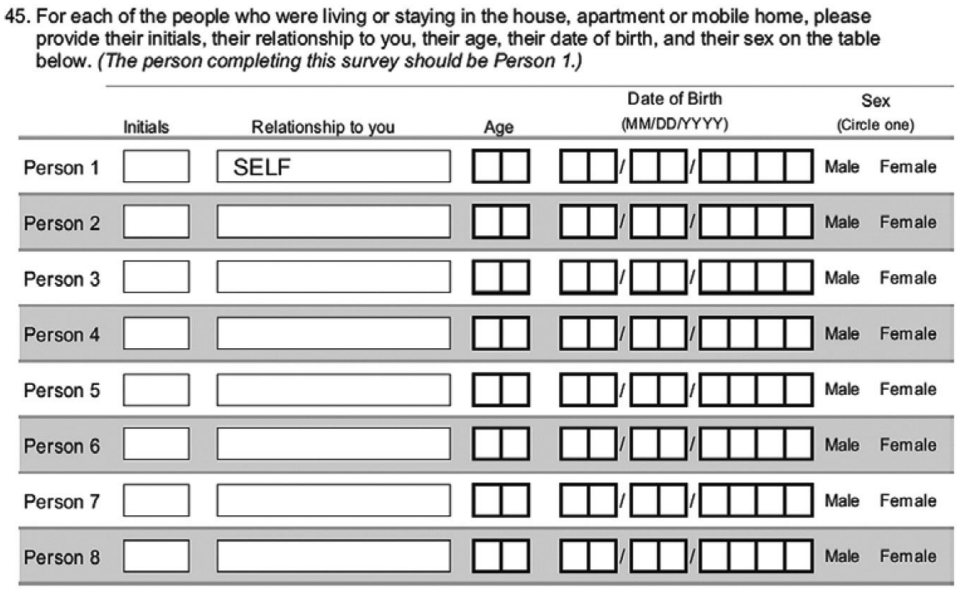
Table 1. Wording Included in Cover Letters for Each Within-Household Respondent Selection Treatment

Treatment group	Respondent selection instruction wording
	NASIS
	In order to make this study more scientific, we ask that the enclosed survey be completed by the adult (age 19 or older) in your household who…
Last birthday	*most recently celebrated a birthday.*
Next birthday	*be the next to celebrate a birthday.*
Oldest adult	*is the oldest adult in your household.*
Youngest adult	*is the youngest adult in your household.*
	Nebraska Trees and Forest Survey
	To ensure that we talk to all different types of Nebraskans, we ask that the enclosed survey be completed by the adult (age 19 or older) in your household who…
Last birthday	*most recently celebrated a birthday.*
Next birthday	*will be the next to celebrate a birthday.*

To identify whether outcomes of interest differ across the household-selection methods, we examine response rates and compositional differences between the selection methods in both surveys. We also compare the completed samples produced by each selection method in both surveys to ACS benchmarks, but this comparison is more direct in the Forest Survey because its sample was selected from a frame with very good coverage (the DSF). By comparison, differences between NASIS and ACS benchmark estimates can be due both to errors in the household-selection process and to the inherent coverage issues associated with the directory-listed sample on which the NASIS is based, although we would not expect coverage errors to vary across experimental treatments.[1] The strength of the NASIS is that it included a household roster, which we use to evaluate the accuracy of the within-household selection. No such roster is available in the Forest Survey.

For our composition analyses, we examine differences in sex, race, ethnicity, age, presence of children, family income, and education level across the within-household selection methods in both studies, and number of persons in the household, number of adults in the household, employment status, and marital status in the NASIS alone (the Forest Survey did not ask about these demographic characteristics). We expect the most substantial differences to be on age and age-related characteristics for the youngest- and oldest-adult methods, but have no expectations for differences across other characteristics. The assumption in the literature is that the next- and last-birthday methods yield roughly identical samples, and we have no reason to hypothesize otherwise.

1. For both surveys, potential coverage, nonresponse, and measurement errors in the ACS may lead to inaccurate comparisons. These errors are not likely to differ across the experimental treatments.


45. For each of the people who were living or staying in the house, apartment or mobile home, please provide their initials, their relationship to you, their age, their date of birth, and their sex on the table below. *(The person completing this survey should be Person 1.)*

	Initials	Relationship to you	Age	Date of Birth (MM/DD/YYYY)	Sex (Circle one)
Person 1		SELF		/ /	Male Female
Person 2				/ /	Male Female
Person 3				/ /	Male Female
Person 4				/ /	Male Female
Person 5				/ /	Male Female
Person 6				/ /	Male Female
Person 7				/ /	Male Female
Person 8				/ /	Male Female

Figure 1. Household Roster Questions, NASIS.

For our benchmark comparisons, we compare the sample composition of the NASIS and Forest Surveys to benchmark ACS data for Nebraska on selected demographic characteristics (e.g., race/ethnicity, education level, age) to evaluate how well the survey's target population is represented overall and in each experimental condition. We reran all of the Forest Survey analyses with probability of selection weights accounting for the unequal probabilities of selection across strata (the NASIS is an *epsem* design). None of the findings changed. To assess the accuracy of the selection process in the NASIS, we use information from the household roster included in the questionnaire to evaluate the accuracy of selections in each treatment. The household roster consisted of seven demographic items for each household member: initials, relationship to the respondent, age, sex, month of birth, day of birth, and year of birth (see figure 1).

For the next- and last-birthday treatment groups, we determine accuracy by comparing the month and day of birth information for all adults age 19 or older to the date of the initial NASIS mailing (April 13)—we assume that selection should have occurred with the initial mailing. An accurate selection occurred if the person listed as person 1 in the roster had the birthday that was closest to April 13 among all adults in the household, either following (next birthday) or preceding (last birthday) the mailing date. A limitation to this approach, however, is the potential for a change in who should be selected as the respondent in the time between receiving the initial survey and completing/returning the questionnaire. The household roster data indicate that 21.1 percent of all adults listed on the roster have a birthday during the survey field period April 13 to June 18.

For the oldest- and youngest-adult treatment groups, we examine the age of the household members listed on the roster. Accuracy was calculated by comparing the date-of-birth information for all adults age 19 or older to the date of the initial mailing. An accurate selection occurred if the first person listed on the roster was the youngest adult in the household for the youngest-adult treatment or the oldest adult in the household for the oldest-adult treatment.

We examine accuracy of selection for all households, households with two or more adults, and households with at least one adult with a birthday during the field period. Additionally, we examine the association of accuracy of selection with respondent and household composition characteristics using the same demographic characteristics that we used in our composition analyses and test whether the association between respondent and household characteristics and accuracy varies across the treatment groups. Here, we focus on households that have at least two adults because households with one adult will have an accurate selection by default.

Findings

Response Rates

We start by examining response rates across the treatment conditions in both surveys. In NASIS, the response rates for the last-birthday (38.7 percent), next-birthday (36.8 percent), and oldest-adult (37.4 percent) conditions were not significantly different from one another ($\chi^2(3) = 6.33$, $p = 0.10$), but the youngest-adult condition had a significantly lower response rate (AAPOR RR1) at 32.0 percent ($\chi^2(1) = 6.40$, $p = 0.01$). In the Forest Survey, there was no significant difference in the response rates between the last birthday (30.9 percent) and next birthday (32.3 percent) methods ($\chi^2(1) = 0.68$, $p = 0.41$).[2]

Composition

We next examine the composition of the completed samples produced by each within-household selection method. As indicated by the chi-square tests reported in table 2, there are very few differences in the characteristics of respondents across the selection methods in these two surveys. The exceptions are for age ($\chi^2(9) = 20.10$, $p = 0.02$) and marital status ($\chi^2(6) = 13.10$, $p = 0.04$) in the NASIS,

2. For the Forest Survey, we also examined factors predicting participation in the survey. We ran logistic regressions predicting response using frame data provided by SSI on estimated homeownership, length of residence, city residence, children in household, Hispanic individual(s) in household, and assignment to selection method treatment, as well as interaction terms between the frame variables and the experimental treatment. None of the variables or interaction terms were statistically significant predictors of survey participation (analyses available upon request).

Table 2. Demographic and Socioeconomic Composition of Respondents and Households

	NASIS						Forest survey				
	Total sample	**Last birthday**	**Next birthday**	**Oldest adult**	**Youngest adult**	**χ^2 (*p*-value)**	**Total sample**	**Last birthday**	**Next birthday**	**χ^2 (*p*-value)**	**ACS population estimate**
Sex											
Male	38.74*	41.53*	38.29*	39.46*	35.05*	1.95	47.48	49.45	45.57	1.41	49.5
Female	61.26*	58.47*	61.71*	60.54*	64.95*	(0.58)	52.52	50.55	54.43	(0.24)	50.5
Race											
White	94.52*	93.16	92.89	95.54	96.89*	4.53	93.66*	92.87*	94.42*	0.96	88.2
Nonwhite	5.48*	6.84	7.11	4.46	3.11*	(0.21)	6.34*	7.13*	5.58*	(0.33)	11.8
Ethnicity											
Not Hispanic	97.35*	95.45*	96.96*	99.15*	98.00*	6.76	98.52*	97.84*	99.17*	2.89	90.6
Hispanic	2.65*	4.55*	3.04*	0.85*	2.00*	(0.08)	1.48*	2.16*	0.83*	(0.09)	9.4
Age											
Mean	55.32	56.49	53.96	56.50	54.11	–	55.86	56.55	55.22	–	–
19–34	13.63*	14.86*	13.49*	11.52*	14.74*	20.10	12.09*	11.41*	12.74*	1.38	28.3
35–49	22.63*	19.82*	24.19	17.97*	29.47	(0.02)	22.22*	22.15*	22.29*	(0.71)	25.9
50–64	33.65*	30.63	35.81*	41.94*	25.26		36.27*	35.35*	37.15*		27.1
65+	30.09*	34.68*	26.51*	28.57*	30.53*		29.41*	31.10*	27.81*		18.7
Education											
≤ High school	22.09*	19.05*	19.91*	26.36*	23.28*	5.24	23.43*	22.22*	24.57*	2.31	38.0
Some college	34.55	38.10	35.07	32.27	32.28	(0.52)	35.86	38.32*	33.55	(0.32)	34.0
BA +	43.36*	42.86*	45.02*	41.36*	44.44*		40.70*	39.46*	41.88*		28.0

Table 2. *continued.*

	NASIS						Forest survey				
	Total sample	**Last birthday**	**Next birthday**	**Oldest adult**	**Youngest adult**	**χ^2 (p-value)**	**Total sample**	**Last birthday**	**Next birthday**	**χ^2 (p-value)**	**ACS population estimate**
Employment status											
Employed	65.47	63.32	69.27	65.16	64.06	2.04	–	–	–	–	–
Not employed	34.53	36.68	30.73	34.84	35.94	(0.57)	–	–	–	–	–
Family income											
Under $49,999	44.95*	46.30	45.67	45.41	41.99*	3.76	41.45*	41.94*	40.98*	12.55	51.3
$50,000–99,999	37.07*	33.33	36.06	37.68	41.99*	(0.71)	36.87*	40.45*	33.49	(0.01)	33.0
Over $100,000	17.98	20.37	18.27	16.91	16.02		21.69*	17.62	25.53*		15.7
Household size											
Mean	2.45	2.48	2.57	2.29	2.48			–	–	–	–
0 or missing	2.65	4.13	0.43	3.85	2.00	17.11	–	–	–	–	–
1	21.96	23.97	21.30	20.94	21.50	(0.31)	–	–	–	–	–
2	39.85	36.36	39.13	44.02	40.00		–	–	–	–	–
3	14.90	14.45	17.39	14.53	13.00		–	–	–	–	–
4	11.04	10.33	10.43	9.40	14.50		–	–	–	–	–
5+	9.60	10.74	11.30	7.26	9.00		–	–	–	–	–
Number of adults											
0 or missing	6.07	7.85	4.35	6.84	5.00	8.29	–	–	–	–	–
1	24.06	25.62	20.87	24.36	25.50	(0.51)	–	–	–	–	–
2	59.71	57.85	61.30	58.97	61.00		–	–	–	–	–
3+	10.15	8.68	13.48	9.83	8.50		–	–	–	–	–

Table 2. *continued.*

	NASIS						Forest survey				
	Total sample	**Last birthday**	**Next birthday**	**Oldest adult**	**Youngest adult**	χ^2 **(*p*-value)**	**Total sample**	**Last birthday**	**Next birthday**	χ^2 **(*p*-value)**	**ACS population estimate**
Any children?											
Yes	35.37*	38.91*	35.22	32.62	34.50	2.15	31.67	30.24	33.03	0.75	32.1
No	64.63*	61.09*	64.78	67.38	65.50	(0.54)	68.33	69.76	66.97	(0.39)	67.9
Marital status											
Married, cohabiting	68.76	65.81	73.66	69.37	65.98	13.10	–	–	–	–	–
Not married	23.11	26.92	17.41	25.68	22.16	(0.04)	–	–	–	–	–
Never married	8.12	7.26	8.93	4.95	11.86			–	–	–	–
n	906	242	230	234	200		947	463	484		

*Different from ACS estimate at $p < .05$ level

and family income ($\chi^2(3) = 12.55$, $p < 0.01$) in the Forest Survey. As expected, the youngest-adult condition in the NASIS brought in proportionately more younger respondents and the oldest-adult condition brought in proportionately more older respondents (test between oldest- and youngest-adult condition: $\chi^2(3) = 14.92$, $p = 0.002$). The age distribution does not significantly differ between any of the other conditions at the $p < .05$ level. The youngest-adult condition also brought in more respondents who have never been married (test between oldest- and youngest-adult condition: $\chi^2(2) = 6.73$, $p = 0.03$), an effect that is consistent with having more young respondents in the completed sample. Unexpectedly, the next-birthday method brought in more persons who were currently married or cohabiting than the last-birthday or oldest-adult conditions. As this is the only significant difference between the next-birthday method and any other method for any variable in the NASIS, we do not expect these differences to be systematic for the next-birthday method in general. Across the seven characteristics examined in the Forest Survey, only family income differed significantly across the conditions, with the next-birthday method yielding proportionately more persons with incomes over $100,000. We do not see the same pattern in the NASIS.

Because no within-household selection is needed in one-adult households, we conducted these same analyses only for households that contain two or more adults in NASIS (results available from authors on request). Differences in age across the four treatments are marginally significant when looking only at households with two or more adults ($\chi^2(9) = 16.11$, $p = 0.07$), but differences in marital status disappear. None of the other comparisons across the four treatments are statistically significant at the $p < .05$ level.[3]

Benchmark Comparisons

We compare characteristics of the completed sample from the Forest Survey to statewide data about Nebraska from the 2011 American Community Survey (ACS)[4] (see table 2). On the vast majority of the variables, the Forest Survey total sample (combining the two birthday methods), as well as each birthday method individually, differs significantly from the ACS state estimates ($p < .05$), as indicated by the asterisks in table 2. The Forest Survey overrepresents those who were white, had high family incomes, had high education levels, and were older. For example, the ACS shows that 28.0 percent of Nebraskans have completed a

3. We also examined whether substantive estimates differed by within-household selection method for variables in both NASIS and the Forest Survey. Estimates did not significantly differ by treatment group for either survey. We also found no differences in estimates in NASIS for the correctly selected respondents versus the incorrectly selected respondents (analyses available upon request).

4. We also ran analyses with older five-year ACS estimates and found similar results. For the Forest Survey, we accounted for the sampling design using the svy commands in Stata. We found similar results between the weighted and unweighted data. We report the unweighted data.

bachelor's or graduate degree, but the total sample (40.7 percent), last-birthday sample (39.5 percent), and next-birthday samples (41.9 percent) all overrepresent this education group ($p < .05$). Conversely, the total sample and both birthday methods underrepresent those who were nonwhite, had low incomes, and were younger. For example, for Hispanics, the total sample (1.5 percent), last-birthday (2.2 percent), and next-birthday (0.8 percent) methods all underrepresent the 9.4 percent of Nebraska's population that is Hispanic ($p < .05$). The total sample and both birthday methods are similar to the ACS on sex and for households containing children ($p > .05$). While there are a few differences between the samples yielded by the two selection techniques, neither yielded a sample more representative of the target population than the other. Across the seven variables examined in the Forest Survey, the average absolute difference between the last-birthday estimates and the ACS estimates is 7.38 percentage points, compared to 7.43 percentage points for the next-birthday estimates.

Comparisons between the NASIS and the ACS estimates combine coverage errors as well as nonresponse and selection errors. As can be seen in table 2, most of the demographic characteristics measured in the NASIS follow the same pattern as that seen in the Forest Survey and are significantly different from the ACS benchmark estimates. The primary difference between the Forest Survey and the NASIS is that the sex distribution in the NASIS significantly differs from the benchmark, whereas there is no significant difference between the Forest Survey and the ACS benchmark. Interestingly, the average absolute difference in estimates between the NASIS and ACS characteristics is quite similar for the same items examined in the Forest Survey. For the last-birthday method, the average absolute difference is 7.45 percentage points, compared to 7.65 percentage points for the next-birthday method, 7.47 percentage points for the oldest-adult method, and 8.26 percentage points for the youngest-adult method.

Accuracy of Selection

We now examine the accuracy of selecting respondents across the within-household selection conditions, using measures from the household roster included in NASIS. As table 3 shows, the correct household member is selected in 68.0 percent of all households, and there are no statistically significant differences across the four selection methods. This evaluation, however, includes single-adult households, which have a 100 percent accuracy rate by default. When we evaluate accuracy for households with two or more adults (i.e., households that have a chance to make an error), it falls to 57.5 percent, which is roughly equivalent to a coin flip. In these households, the next-birthday method has the highest accuracy rate (62.6 percent) and the last-birthday method has the lowest accuracy rate (52.9 percent; test between next- and last-birthday, $\chi^2(1) = 2.62$, $p = 0.11$).

Table 3. Percent Accurate Selections by Within-Household Selection Method, NASIS

	Percent selecting correct respondent			
	All households	Households with 2+ adults	Household with an adult birthday during field period	Households without an adult birthday during field period
Overall	68.04	57.53	63.10	70.83
Last birthday	65.96	52.94	59.42	69.75
Next birthday	71.04	62.59	61.84	77.57
Oldest adult	70.44	60.54	75.76	67.88
Youngest adult	64.41	53.49	55.00	69.23
n	751	551	271	480
Chi-square	2.75	4.04	6.69	3.14
p-value	0.43	0.26	0.08	0.37

Households with birthdays that occur around the field period may be more likely to have incorrect selections, specifically for the birthday-selection methods (Forsman 1993). Using the household roster data, we examined accuracy of selection by whether households contained at least one adult with a birthday during the field period (table 3). For the last-birthday method, we found that 59.4 percent of respondents were accurately selected in households with an adult birthday during the field period and 69.8 percent of respondents were accurately selected in households without a birthday occurring during the field period ($\chi^2(1) = 2.08$, $p = 0.15$). For the next-birthday method, we found that 61.8 percent of respondents were accurately selected in households with an adult birthday during the field period and 77.6 percent of respondents were accurately selected in households without a birthday occurring during the field period ($\chi^2(1) = 5.34$, $p = 0.02$). Thus, only in the next-birthday method are there significant differences by whether a birthday occurred during the field period.

One challenge with using a household roster to evaluate the accuracy of selection is missing data within the roster itself. In NASIS, 9.6 percent of respondents left the entire roster blank and 11.7 percent failed to complete any information for at least one household member. There is no difference overall across the four selection methods ($\chi^2(3) = 3.23$, $p = 0.36$) in leaving the roster completely blank.[5]

5. In terms of demographic differences between nonrespondents and respondents to the household roster, we observed statistical differences only for the percent of households with children ($\chi^2(1) = 10.81$, $p = 0.001$) and household size ($\chi^2(5) = 240.27$, $p < 0.001$). We found that respondents living with children in the household are more likely to be accurately selected in the youngest-adult procedure (63.0 percent), but less likely to be selected accurately in the oldest-adult procedure (47.7 percent) (analyses available upon request).

We next examine characteristics of respondents who were accurately selected across the four within-household selection methods among households with at least two adults (table 4). There are no significant differences in accuracy of selection across the four within-household selection methods for whites versus nonwhites, persons of different education levels, employed versus unemployed, for different sizes of households, households with or without children, by income, or by marital status (results available from authors on request).

For households with two or more adults, 59.9 percent of male respondents and 56.4 percent of female respondents are accurately selected when looking at all four selection methods combined ($\chi^2(1) = 0.68$, $p = 0.41$). The rates vary dramatically for the two age-position selection methods. In the oldest-adult condition, 86.4 percent of male respondents are correctly selected, compared to 44.2 percent of female respondents ($\chi^2(1) = 26.36$, $p < .0001$). In contrast, in the youngest-adult condition, 22.5 percent of male respondents are correctly selected, compared to 73.1 percent of female respondents ($\chi^2(1) = 31.01$, $p < .0001$). We believe that this difference occurs because in married and opposite-sex-partnered households, men are on average older than women (Simmons and O'Connell 2003), not because males are better at following the oldest-adult selection procedure and females are better at following the youngest-adult procedure. There was no difference in accuracy of selection for males and females in either of the birthday methods.

Overall, 61.1 percent of respondents age 65 and older and 64.6 percent of respondents age 19 to 34 are selected accurately, compared to just over 50 percent of respondents in the middle two age categories ($\chi^2(3) = 4.55$, $p = 0.21$). This varies by

Table 4. Percent Correct Respondent Selected by Demographic and Socioeconomic Characteristics, 2+ Adult Households Only, NASIS

	Overall	Last birthday	Next birthday	Oldest adult	Youngest Adult
Sex					
Male	59.91	56.06	70.83	86.44	22.45
Female	56.35	49.28	58.89	44.19	73.08
Chi-square	0.68	0.62	1.92	26.36****	31.01****
Age					
19–34	64.63	78.26	61.90	44.44	70.00
35–49	55.47	38.46	66.67	57.69	55.00
50–64	52.41	40.48	65.31	57.14	39.39
65+	61.11	57.58	57.69	77.14	50.00
Chi-square	4.55	10.90*	0.63	6.38	4.89
n	511	124	130	143	114

* $p < .05$; **** $p < .0001$

the selection procedure—in the last-birthday method, there is a significant association between age and the rate of accurate selections ($\chi^2(3) = 10.90$, $p = 0.01$), but there is no such association in the next-birthday method ($\chi^2(3) = 0.63$, $p = 0.89$). In the oldest-adult method, 77.1 percent of respondents age 65 and older are selected accurately, compared to 55.1 percent of respondents age 19 to 64 ($\chi^2(1) = 5.35$, $p = 0.02$). In the youngest-adult method, 70.0 percent of respondents age 19 to 34 are selected accurately, compared to 48.6 percent of respondents age 35 and older. While the size of this difference is quite large, it fails to reach traditional significance levels of $p < .05$ ($\chi^2(1) = 3.09$, $p = 0.08$), likely because there are only 20 respondents in the 19–34 age group.

One important question is whether estimates from the NASIS sample would better reflect population characteristics had the correct respondent participated. Using the household roster data, we examine the age and sex composition that would have been observed had the correct member of the household completed the survey (table 5). For the next- and last-birthday methods, the estimates for the percentage of males and females would be statistically indistinguishable from the ACS estimates had the correct respondent participated. Additionally, the age distributions would be closer to those of the ACS benchmark. The oldest- and youngest-adult methods, on the other hand, would yield estimates of the sex and age distribution that are further from the benchmark for each method individually (the methods are recommended to be used in combination). In general, for the NASIS, the oldest adult in the household should have been a male respondent and the youngest adult should have been a female respondent. Overall, this analysis suggests that we would have

Table 5. Demographic Composition for Correct Respondent from Household Roster, NASIS

	Total sample	Last birthday	Next birthday	Oldest adult	Youngest adult	□ 2 (*p*-value)	ACS population estimate
Sex							
Male	45.99*	45.65	46.96	62.03*	26.42*	43.98	49.5
Female	54.01*	54.35	53.04	37.97*	73.58*	(0.000)	50.5
Age							
19–34	17.00*	17.30*	20.99*	8.97*	22.50	39.49	28.3
35–49	19.75*	21.62	22.10	13.68*	23.00	(0.000)	25.9
50–64	30.25	30.81	31.49	38.03*	19.50*		27.1
65+	33.00*	30.27*	25.41*	39.32*	35.00*		18.7
n	800	185	181	234	200		

*Different from ACS estimate at $p < 0.05$ level

overrepresented older individuals and underrepresented younger adults even if the correct adult member of the household had responded to the NASIS, but that the gender distribution would have been fixed had the correct respondent participated.

In sum, just over half of the selections made in any of the within-household selection procedures are correctly made. There are few correlates of accuracy of selection—this is encouraging, in that errors are not being made differentially across important demographic or socioeconomic groups.

Conclusion

There are several main findings in this paper that are important for understanding within-household selection methods in mail surveys. First, few differences occurred in the response rates across the different within-household selection methods. Only the youngest-adult method had a significantly lower response rate. Given how it is particularly difficult to get young people to respond to surveys, it is possible that this lower response rate occurred because households followed the selection procedure correctly but the designated participant was unwilling to respond (i.e., a general nonresponse problem rather than strictly a within-household selection problem). Second, few differences in the demographic characteristics of respondents were found across different selection procedures. Third, consistent with previous research, all samples across the within-household selection methods we examined misrepresent some key demographic groups in ways such as overrepresentation of whites and those with the highest levels of education (e.g., Battaglia et al. 2008; Hicks and Cantor 2012). Finally, in our analyses of accuracy of selection, all four within-household selection methods examined were equally accurate in the selections made, although the accuracy rate is low. The rates, however, are comparable to those in other examinations of within-household selection in mail surveys (e.g., Schnell, Ziniel, and Coutts 2007; Battaglia et al. 2008).

Our findings are both encouraging and discouraging. On the encouraging side, few differences were found across the various conditions with respect to who ends up in the completed sample. The differences tend to be related to characteristics that should vary, such as age in the two age-related selection procedures (oldest- and youngest-adult methods). On the discouraging side, the samples produced by the selection methods are unrepresentative on certain key demographics. Non-Hispanic whites, older individuals, individuals with more education, and households with higher family incomes are overrepresented. Also discouraging is that only slightly more than half of the selections made in households with two or more people are made accurately. This is a troubling sign for within-household selection in mail surveys, as it appears that household informants are failing in at least one of the three tasks that they need to complete in order to make an accurate selection.

The question remains as to what goes wrong with the selection methods. It may be that households do not see the within-household selection instructions embedded in the cover letter. In this case, different emphasis or different placement of these instructions may lead to more accurate selections. Another possibility is that households read the instructions but do not understand how to carry them out. This is certainly possible in selection methods such as the birthday methods, but seems less likely in methods such as the youngest adult or oldest adult. A third possibility is that the instructions are read and accurately followed, but the selected person does not want to participate in the survey. In this case, the household may return a survey that was completed by a different household member. A fourth possibility is that the household informant reads and understands the within-household selection instructions but does not understand why it is important to follow the selection procedure. For example, some people may not understand why it is important to hear from all selected sample members or may not understand how birthdays relate to the goal of hearing from all respondents. Additionally, the wording in the cover letter for the birthday methods indicated that the household member who would be next to "celebrate" their birthday should be the respondent. This wording may imply that we want the adult who formally observed ("celebrated") their birthday to respond to the survey.[6] Any or all of these are likely to occur. Future research is needed to investigate what is going wrong and ways to help households carry out the within-household selection accurately.

In the end, our study is not without limitations. First, this research was conducted in a state with few racial or ethnic minorities, thus limiting our inferences about representation of these important and growing demographic groups. Second, as an omnibus survey, the NASIS may have had topics that were not appealing or interesting to many respondents. Likewise, the topic of forests and trees may have been uninteresting to many Nebraskans, perhaps contributing to the overall high rate of inaccurate selections. For example, with low-saliency topics, the wrong person may respond more often because the selected person is uninterested and chooses not to respond. High-saliency topics may result in higher within-household selection accuracy because they are more desirable generally. The role of questionnaire saliency or topic interest in within-household selection accuracy is an area in need of additional future research.

Our study assumes that the household roster used to assess accuracy was itself completed accurately, among those who filled it out. We find no differences in completion rates of the roster across the four conditions, and do not expect that the household selection methods would yield differential measurement error in the roster, but cannot evaluate that assumption from these data alone. We also cannot

6. Some households may have interpreted the word "celebrate" to indicate that the survey was a guise for selling something for "birthday celebrations." We thank an anonymous reviewer for this observation.

evaluate accuracy of selection among households who fail to complete the household roster. Next, the measures of accuracy in the NASIS are for a sample from listed landline households. There are marked differences between households with landlines and those with mobile phones only (Blumberg and Luke 2013). This coverage issue certainly contributes to the discrepancy between the composition of the NASIS respondents and the ACS benchmark. We do not know whether a sample with more cell-phone-only households would have different rates of selection accuracy or correlates of accuracy. Finally, the ACS estimates are not without potential survey errors. If households containing certain subgroups (e.g., young adults) are systematically less likely to participate in the ACS, then the benchmark estimates and resulting comparisons will be flawed. However, we do not believe that errors in the ACS systematically affect one of the experimental treatments more than another.

Despite these limitations, this study was the first to our knowledge to compare the next-birthday, last-birthday, oldest-adult, and youngest-adult within-household selection methods in general-population mail surveys. The experimental assignment of treatments and the inclusion of a household roster in NASIS to identify whether respondents were accurately selected are strengths of the study, as is the replicated test of the two birthday treatments across both surveys. Future research is needed to attempt to replicate these findings, test the findings with other sample frames (e.g., more diverse target populations), and identify design features that may assist in improving the accuracy of within-household selection in mail surveys. Other within-household selection methods, such as those based on household size (Le et al. 2013), also require testing in mail surveys.

Appendix. Question Wording for NASIS and Forest Survey

Sex (NASIS and Forest Survey)

Are you:
- Male
- Female

Race (NASIS)

What race or races do you consider yourself to be? *(Check all that apply.)*
- White (Caucasian)
- Black or African American
- Asian
- American Indian or Alaska Native
- Native Hawaiian or Other Pacific Islander
- Other race(s) (please specify):

Ethnicity (NASIS)

Do you consider yourself to be Hispanic or Latino/a?

Yes

No

Race/Ethnicity (Forest Survey)

What race or races do you consider yourself to be? *(Check all that apply.)*

White (Caucasian)
Black or African American
Asian
American Indian or Alaska Native
Native Hawaiian or Other Pacific Islander
Hispanic or Latino/a
Other race(s) (please specify):

Age (NASIS and Forest Survey)

In what year were you born?

1	9		

Education (NASIS and Forest Survey)

What is your highest level of education?

No diploma
High school diploma/GED
Some college, but no degree
Technical/associate/junior college (2 yr, LPN)
Bachelor's degree (4 yr, BA, BS, RN)
Graduate degree (master's, PhD, law, medicine)

Employment Status (NASIS)

Do you typically work full-time, part-time, go to school, keep house, or something else? *(Check all that apply.)*

Working a full-time job (35 hours or more)
Working a part-time job(s)
Unemployed, laid off, looking for work
Retired
In school
Keeping house
Disabled
Other, specify:

Family Income (NASIS)

Please indicate the category that describes your total family income in the past 12 months.

Under $5,000
$5,000 to $9,999
$10,000 to $14,999
$15,000 to $19,999
$20,000 to $24,999
$25,000 to $29,999
$30,000 to $39,999
$40,000 to $49,999
$50,000 to $59,999
$60,000 to $74,999
$75,000 to $99,999
$100,000 or more

Family Income (Forest Survey)

Please indicate the category that describes your total family income in the past 12 months.

Under $10,000
$10,000–19,999
$20,000–29,999
$30,000–39,999
$40,000–49,999
$50,000–59,999
$60,000–69,999
$70,000–79,999
$80,000–89,999
$90,000–99,999
$100,000 or more

Number of Adults (NASIS)

Including yourself, how many adults age 19 and older live in your household?

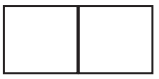
Any Children (NASIS)

How many children ages:
(Please write "0" if none.)

a. 5 and younger live in your household?
b. 6 to 12 live in your household?
c. 13 to 18 live in your household?

Children in Household (Forest Survey)

How many children, under age 19, are currently living in your household all or part of the time? View this table:



Marital Status (NASIS)

What is your current marital or relationship status?
Married
Married, living apart
Not married but living with a partner (cohabiting)
Never married
Divorced
Widowed
Separated

References

Battaglia, Michael P., Michael W. Link, Martin R. Frankel, Larry Osborn, and Ali H. Mokdad. 2008. "An Evaluation of Respondent Selection Methods for Household Mail Surveys." *Public Opinion Quarterly* 72:459–69.

Blumberg, Stephen J., and Julian V. Luke. 2013. "Wireless Substitution: Early Release of Estimates from the National Health Interview Survey, July–December 2012." National Center for Health Statistics early release report. Available at http://www.cdc.gov/nchs/data/nhis/earlyrelease/ wireless201306.pdf.

Forsman, Gosta. 1993. "Sampling Individuals Within Households in Telephone Surveys." Paper presented at the Annual Meeting of the American Association for Public Opinion Research, St. Charles, IL, USA.

Gallagher, Patricia M., Floyd Jackson Fowler Jr., and Vickie L. Stringfellow. 1999. "Respondent Selection by Mail Obtaining Probability Samples of Health Plan Enrollees." *Medical Care* 37:MS50–58.

Gaziano, Cecilie. 2005. "Comparative Analysis of Within-Household Respondent Selection Techniques." *Public Opinion Quarterly* 69:124–57.

Hicks, Wendy, and David Cantor. 2012. "Evaluating Methods to Select a Respondent for a General Population Mail Survey." Paper presented at the Annual Meeting of the American Association for Public Opinion Research, Orlando, FL, USA.

Iannacchione, Vincent G. 2011. "The Changing Role of Address-Based Sampling in Survey Research." *Public Opinion Quarterly* 75:556–75.

Lavrakas, Paul J., Elizabeth A. Stasny, and Brian Harpuder. 2000. "A Further Investigation of the Last-Birthday Respondent Selection Method and Within-Unit Coverage Error." Proceedings of the Survey Research Methods Section of the American Statistical Association, 890–95.

Le, Kien T., J. Michael Brick, Abdoulaye Diop, and Darwish Al-Emadi. 2013. "Within-Household Sampling Conditioning on Household Size." *International Journal of Public Opinion Research* 25:108–18.

Lind, Katherine, Michael Link, and Robert Oldendick. 2000. "A Comparison of the Accuracy of the Last Birthday versus the Next Birthday Methods for Random Selection of Household Respondents." Proceedings of the Survey Research Methods Section of the American Statistical Association, 887–89.

Oldendick, Robert W., George F. Bishop, Susan B. Sorenson, and Alfred J. Tuchfarber. 1988. "A Comparison of the Kish and Last Birthday Methods of Respondent Selection in Telephone Surveys." *Journal of Official Statistics* 4:307–18.

Olson, Kristen, and Jolene D. Smyth. 2014. "Accuracy of Within-Household Selection in Web and Mail Surveys of the General Population." *Field Methods* 26:56–69.

O'Rourke, Diane, and Johnny Blair. 1983. "Improving Random Respondent Selection in Telephone Surveys." *Journal of Marketing Research* 20:428–32.

Reich, James, William Yates, and Robert Woolson. 1986. "Kish Method for Mail Survey Respondent Selection." *American Journal of Public Health* 76:206.

Schnell, Rainer, Sonja Ziniel, and Elizabeth Coutts. 2007. "Inaccuracy of Birthday Respondent Selection Methods in Mail and Telephone Surveys." Paper presented at the Annual Meeting of the European Survey Research Association Conference, Prague. Available at http://www.unidue. de/soziologie/schnell_publikationen_presentations.php.

Simmons, Tavia, and Martin O'Connell. 2003. "Married-Couple and Unmarried-Partner Households: 2000." Census 2000 Special Reports CENSR-5. Online at http://www.census. gov/prod/2003pubs/censr-5.pdf

Troldahl, Verling C., and Roy E. Carter Jr. 1964. "Random Selection of Respondents Within Households in Phone Surveys." *Journal of Marketing Research* 1:71–76.

Yan, Ting. 2009. "A Meta-Analysis of Within-Household Respondent Selection Methods." Paper presented at the Annual Meeting of the American Association for Public Opinion Research, Hollywood, FL, USA.

Ziniel, Sonja. 2003. *The External Validation of the Last-Birthday Method as Respondent Selection Method in Mail Surveys.* Master's thesis, Department of Politics and Management, University of Konstanz.

◊ ◊ ◊ ◊

Acknowledgments — The authors thank Stacia Jorgensen for initial research assistance, the Bureau of Sociological Research at UNL for data collection, and three anonymous reviewers for comments that improved the manuscript.

The Authors

Kristen Olson and **Jolene Smyth** are associate professors in the Department of Sociology and the Survey Research and Methodology Program at the University of Nebraska–Lincoln, Lincoln, NE, USA. **Mathew Stange** is a graduate student in the Survey Research and Methodology Program at the University of Nebraska–Lincoln, Lincoln, NE, USA.