

ARTICLES

# Within-Household Selection and Dual-Frame Telephone Surveys: A Comparative Experiment of Eleven Different Selection Methods

Jennifer Marlar<sup>1</sup>, Manas Chattopadhyay<sup>1</sup>, Jeff Jones<sup>1</sup>, Stephanie Marken<sup>1</sup>, Frauke Kreuter<sup>2</sup>

<sup>1</sup> Gallup, <sup>2</sup> Joint Program in Survey Methodology, University of Maryland, University of Mannheim & IAB

Keywords: within-household selection, research methods, telephone survey, survey sampling

<https://doi.org/10.29115/SP-2018-0031>

---

## Survey Practice

Vol. 11, Issue 2, 2018

---

Numerous within-household selection methods have been tested in general population surveys since the advent of telephone interviewing. However, very few selection studies, if any, have been conducted with a dual frame (landline and cell phone) sample. Landline and cell phone frames are known to represent demographically different groups of respondents, and selection methods that may result in more representative demographics in a landline frame may actually skew the results when combined with the cell phone frame. This study tested 11 different within-household selection methods with approximately 11,000 landline respondents. A parallel cell phone sample was also collected with 1,000 respondents, and the frames were combined for analysis. The selection methods tested included one probability-based method, four quasi-probability methods and six nonprobability methods. The methods were evaluated on four criteria: response rates, accuracy, demographic representation and substantive results. The demographic representativeness of each method was examined for the landline frame only and for the dual (landline and cell phone) frame combination. The probability method had the lowest response rate, while the nonprobability at-home methods had the highest. Accuracy rates were lowest for the quasi-probability birthday methods. There were few demographic differences between selection methods, and no substantive differences, when combined with the cell phone sample.

## Introduction

For decades, within-household selection has been a standard component of probability-based telephone surveys. However, within-household selection methods that were designed to result in more representative samples from landline frames may no longer be as effective when combined with the demographically different cell phone frame. The present study evaluates 11 selection methods and their resulting response rates, accuracy of the selection, demographic representativeness, and substantive results. The landline samples were then combined with a cell phone sample to assess demographic and substantive differences between treatment groups.

## ***SELECTION METHODS***

Within household selection techniques can be grouped into three main categories, with the probability component being the key distinction: probability methods, quasi-probability methods and nonprobability methods (Gaziano 2005). Methods vary in terms of complexity and invasiveness, adherence to true probabilistic methodology, and their utility for obtaining interviews from traditionally underrepresented demographics.

Probability-based methods enumerate the adult members of the household, have a random component for selecting the respondent in each household, and every adult member of the household has a known probability of selection (Kish 1949, 1965). While probability-based methods may result in unbiased selection, the enumeration of all household members can seem intrusive and burdensome to respondents (Paisley and Parker 1965; Salmon and Nichols 1983) and may result in nonresponse bias (Binson, Canchola, and Catania 2000; O'Rourke and Blair 1983).

Quasi-probability selection traditionally employs a next or last birthday method. While birthday methods have a random component for selection, if any variables of interest are related to the date of birth, the selection may bias those estimates (Gaziano 2005).

Nonprobability methods include techniques that are designed to simplify the selection process, such as selecting from only the people at home and/or to improve the age and gender distribution of the sample. These methods include "in household" methods that select a person within the household and methods that select among only the people "at home." All nonprobability methods lack a true random component, and certain members of the household may have no opportunity to be selected.

## **Data and Methods**

### ***SELECTION METHODS***

This study compared one probability method, four quasi-probability, and six nonprobability methods.

The probability method that we used in this experiment was a variant of the Rizzo-Brick-Park (R-B-P) method (Rizzo, Brick, and Park 2004). Single-member households did not require a selection component, respondents in two-member households were selected randomly by a computer, and the Kish grid was used to select respondents in households with three or more members (Kish 1949, 1965).

The four quasi-probability methods we used were: most recent birthday, last birthday, next birthday, and a multipart most recent birthday.

We tested three “in-household” and three “at-home” nonprobability methods: youngest adult in-household, youngest male/youngest female (YMYF) in-household, YMYF at-home, multipart YMYF in-household, multipart YMYF at-home, and youngest male/oldest female (YMOF) at-home.

Gender was assigned at random in the YMYF methods, and if an adult of the selected gender did not reside in the household, the interviewer asked for the youngest adult of the other gender. For the in-household version of YMYF, if the selected person was not available, the interviewer scheduled a call back. In the at-home version, if the selected gender was not available, the interviewer asked to speak with the youngest adult of the other gender.

The YMOF method always asked for the youngest adult male first. If an adult male did not live in the household or was not at home, the interviewer asked to speak to the oldest female at home. If the male was at home but was not available to complete the interview, a call-back was scheduled.

Within these groups, we also tested single-question and multipart question formats. Figure 1 shows the exact wording we used for each of these methods. We hypothesized that respondents may find it difficult to process all the information that is typically included in within-household selection items. Take for example the following request: *"May I please speak to the person, 18 years of age or older, living in this household, who had the most recent or last birthday?"* In that single question, respondents must consider three components: the required age, residency in the household and birthday. Breaking the query into separate items may help the respondent accurately comprehend and report the information required.

## ***DATA COLLECTION***

Gallup conducted 10,999 random-digit-dial (RDD) landline interviews from June 24–July 19, 2013. Overall, the study had an 11.3% response rate (AAPOR – RR3). Interviews were conducted in English and Spanish with adults age 18 and older living in the United States. Sampled telephone numbers were randomly assigned to the treatment group prior to data collection. A three call design was used, and the selection method remained the same for the household for all contact attempts.

The survey was approximately five minutes in length and included nine political and economic questions and 14 demographic items. The final question was a household roster. Respondents were asked to report the number of adults in the household and the age, gender and birth date (month and day) of each adult.

Probability	Rizzo-Brick-Park	1) Can you please tell me how many people age 18 or older currently live in your household, including yourself? If 1 adult HH: Interview person on phone If 2 adult HH: Computer randomly selects respondent If 3+ adult HH: Record a list of HH members—computer randomly selects respondent. 2) To make sure our sample is completely random, I need to get a little more information about the people currently living in your household. Can you please tell me the first name or initial of the youngest person age 18 or older living in your household, including yourself? (Probe:) What is the first name or initial of the next youngest person age 18 or older living in your household?
Quasi-probability	Most recent birthday	May I please speak to the person, 18 years of age or older, living in this household, who had the <u>most recent</u> or <u>last</u> birthday? Is that you or someone else?
	Multisequence most recent birthday	1) Are you currently 18 years of age or older? 2) Including yourself, how many adults, 18 years of age or older, are currently living in your household? 3) May I please speak with the adult who had the <u>most recent</u> birthday?
	Last birthday	Please think about everyone living in your household age 18 or older. May I please speak to the person who had the <u>last</u> birthday?
	Next birthday	Please think about everyone living in your household age 18 or older. May I please speak to the person who will have the <u>next</u> birthday?
Nonprobability, in household	Youngest adult, in household	May I please speak to the youngest person age 18 or older living in your household?
	Youngest male/youngest female, in household	Please think about everyone living in your household age 18 or older. May I please speak to the youngest male/female age 18 or older living in your household?
	Multisequence youngest male/youngest female, in household	1) Are you currently 18 years of age or older? 2) Including yourself, how many adults, 18 years of age or older, are currently living in your household? 3) May I please speak to the youngest male/female living in your household?
Nonprobability, at home	Youngest male/youngest female, at home	May I please speak to the youngest male/female 18 years of age or older, who is now at home? (If not home, ask for other gender)
	Multisequence youngest male/youngest female, at home	1) Are you currently 18 years of age or older? 2) Including yourself, how many adults, 18 years of age or older, are currently living in your household? 3) May I please speak to the youngest male/female who is now at home?
	Youngest male/oldest female, at home	May I please speak to the youngest male, 18 years of age or older, who is now at home? (If not home, ask for oldest female)

Figure 1. Selection Method Question Wording

## INTERVIEWERS

Selection methods were randomly assigned to the sample case, not the interviewer. A chi-square test of independence by interviewer and completed selection method was not significant ( $X^2=525.683$ ,  $df=520$ ,  $p=.422$ ) confirming that the assignment mechanism worked and that the association of selection treatment group and an interviewer's completed interviews did not depart from randomness. In other words, there were not instances where some interviewers more successfully completed interviews with some treatment groups than with others.

## CELL PHONE SAMPLE

To create a dual frame sample for each of the 11 treatment groups, we randomly selected a sample of 1,000 completed cell phone interviews that were conducted on Gallup's daily tracking survey June 24–July 19, 2013. This survey has an average response rate of 10.5% (AAPOR RR3) and uses a similar methodology, survey content, and call design as the within household experiment, making it an ideal source of cellular cases. Cell phones were treated as personal devices, and the individual who answered the phone was asked to complete the survey.

## Results

Response rates (AAPOR-RR3), refusal rates (AAPOR – REF2), cooperation rates (AAPOR – COOP1) and contact rates (CON2) are presented in Table 1 (American Association for Public Opinion Research 2016).

Table 1. Response rates by selection method.

	Condition	Response rate (RR3)	Refusal rate (REF2)	Cooperation rate (COOP1)	Contact rate (CON2)
Probability	(1) Rizzo-Brick-Park	9.9% <sup>a</sup>	31.9% <sup>f</sup>	22.1% <sup>a</sup>	44.70%
Quasi-probability	(2) Most recent birthday	12.1% <sup>b</sup>	30.70%	26.7% <sup>i</sup>	45.40%
	(3) Multisequence most recent birthday	11.8% <sup>b</sup>	30.90%	25.9% <sup>b</sup>	45.40%
	(4) Last birthday	11.5% <sup>b</sup>	32.7% <sup>g</sup>	24.2% <sup>e</sup>	47.4% <sup>l</sup>
	(5) Next birthday	11.4% <sup>c</sup>	32.0% <sup>f</sup>	24.6% <sup>j</sup>	46.2% <sup>m</sup>
Nonprobability, in household	(6) Youngest adult, in household	11.1% <sup>d</sup>	30.30%	25.1% <sup>k</sup>	44.30%
	(7) Youngest male/youngest female, in household	10.2% <sup>e</sup>	32.3% <sup>f</sup>	22.4% <sup>e</sup>	45.50%
	(8) Multisequence youngest male/youngest female, in household	10.1% <sup>e</sup>	30.9% <sup>h</sup>	22.9% <sup>e</sup>	44.00%
Nonprobability, at home	(9) Youngest male/youngest female, at home	12.30%	29.30%	27.40%	44.80%
	(10) Multisequence youngest male/youngest female, at home	12.20%	29.60%	27.20%	44.80%
	(11) Youngest male/oldest female, at home	12.60%	29.30%	27.80%	45.10%

\* All differences listed below are significant using the Benjamini-Hochberg procedure with a false discovery rate equal to 0.05 or 5%.

<sup>a</sup> Significantly different from 2, 3, 4, 5, 6, 9, 10, and 11

<sup>b</sup> Significantly different from 7 and 8

<sup>c</sup> Significantly different from 7, 8, and 11

<sup>d</sup> Significantly different from 8, 9, and 11

<sup>e</sup> Significantly different from 9, 10, and 11

<sup>f</sup> Significantly different from 6, 9, 10, and 11

<sup>g</sup> Significantly different from 2, 3, 6, 8, 9, 10, and 11

<sup>h</sup> Significantly different from 9 and 11

<sup>i</sup> Significantly different from 4, 7, and 8

<sup>j</sup> Significantly different from 7, 9, 10, and 11

<sup>k</sup> Significantly different from 7, 8, 9, and 11

<sup>l</sup> Significantly different from 1, 2, 6, 7, 8, 9, 10, and 11

<sup>m</sup> Significantly different from 8

All refusals occurred before selection was completed. Refusals include hard and soft refusals that were unresolved at the end of the field period. Although some respondents may desire a call-back, or the selected respondent is not available at the time of the call, some individuals may request a call-back as a polite way of declining the interview (an implied refusal), for example saying, “Now isn’t a good time,” instead of explicitly refusing. The sample management system did not distinguish between types of call-backs, and we elected to be conservative with refusal rates and include call-backs in refusal rates.

Most recent birthday and the three nonprobability at-home methods had the highest response rates—all exceeding 12%. These methods also had among the lowest refusal rates and highest cooperation rates. The R-B-P method and the YMYF in-household methods had lower response rates than the other methods, which can be explained in part by lower cooperation rates.

This finding for R-B-P, which used a Kish grid for households with more than two adults, is consistent with past methods comparisons which found that respondents are more likely to break off when the Kish method is used (American Association for Public Opinion Research 2016; Gaziano 2005; O'Rourke and Blair 1983).

The response rates for multiquestion selection methods were nearly identical to those that used the same selection method with only a single question screener. It does not appear that breaking the selection process down into a series of shorter questions positively or negatively affected response rates.

## ***ACCURACY***

The birth month and day, age, and gender of household members collected on the household roster were compared against the selected household member and the date of interview to determine if the correct household member completed the interview. The selection check requires an assumption that the household roster is accurate, which may not always be the case.

At-home methods were excluded from the analysis because the roster did not ask which household members were home during the interview. The R-B-P method was also excluded from analysis. Since the computer randomly selected the household member to be interviewed (in cases of 2+ person households) an error in the R-B-P method was not possible. It is important to point out that this does not mean the R-B-P results in perfect accuracy. This method still relies on an informant to give the interviewer accurate information. An informant may either be unable (Martin 2007) or unwilling (Tourangeau et al. 1997) to provide an accurate household roster, and previous work has found that errors do occur (Groves and Kahn 1979).

Accuracy results are displayed in Table 2. Households with only one member are, by default, accurate. Therefore, the table also includes an accuracy figure for households with two or more members.

The most recent birthday method was the least accurate, resulting in errors in nearly 30% of households with two or more members, which is consistent with other research on selection accuracy (Battaglia et al. 2008; Olson and Smyth 2014; Olson, Stange, and Smyth 2014). At the other end of the spectrum, the YMYF in household methods, when asked as a single question or a multipart question, resulted in near perfect accuracy rates.

Breaking screening questions into multipart questions seemed to improve accuracy, although the findings are limited. The most recent birthday item had an accuracy rate of 71.9% when asked as a single item and 79.1% accuracy when broken out into a multipart question. The YMYF in household method had an already extremely high accuracy rate, so the gains with the multipart version are minimal.

Table 2. Accuracy of selection methods.

	Method	% Accurate– households with complete roster, all households (N)	% Accurate– households with complete roster, 2+ adult members (N)
Quasi-probability	(2) Most recent birthday	81.1 (935) <sup>a</sup>	71.9 (629) <sup>a</sup>
	(3) Multisequence most recent birthday	87.9 (952)	79.1 (550)
	(4) Last birthday	87.5 (937)	80.9 (611)
	(5) Next birthday	89.5 (932)	84.0 (611) <sup>c</sup>
Nonprobability, in household	(6) Youngest adult, in household	86.3 (944)	79.1 (618)
	(7) Youngest male/youngest female, in household	98.8 (930) <sup>b</sup>	98.1 (576) <sup>b</sup>
	(8) Multisequence youngest male/youngest female, in household	99.7 (927) <sup>b</sup>	99.4 (534) <sup>b</sup>

\* All differences listed below are significant using the Benjamini-Hochberg procedure with a false discovery rate equal to 0.05 or 5%.

<sup>a</sup> Significantly different from all other groups

<sup>b</sup> Significantly different from 2, 3, 4, 5, and 6

<sup>c</sup> Significantly different from 2, 3, 6, 7, and 8

## DEMOGRAPHICS

In general, selection method demographics for the landline frame did not vary significantly by treatment, unless the method was designed to attempt to oversample a certain demographic characteristic. In these cases, significant differences in gender and age were observed. YMOF at-home, which always asks for the available male first, yielded a sample that was 56% male ( $X^2(10) = 73.552, p < .001$ ). In contrast, the youngest adult in the household method yielded a sample that was 33% male.

The nonprobability methods designed to target younger adults all successfully recruited more young people into the survey than the probability and quasi-probability methods ( $X^2(30) = 102.769, p < .001$ ). This finding is consistent with recent within-household selection research (Olson, Stange, and Smyth 2014). However, all methods ultimately significantly underrepresented young people, which is consistent with other work comparing household selection techniques (Battaglia et al. 2008) and is not unexpected given the landline frame.

There were no significant differences between treatment groups by race and ethnicity [white: ( $X^2(10) = 11.199, p = .342$ ; black: ( $X^2(10) = 10.737, p = .378$ ; Hispanic: ( $X^2(10) = 31.880, p = .373$ ) and educational attainment ( $X^2(50) = 61.55, p = .127$ ], but there were differences by marital status ( $X^2(10) = 86.459, p = .001$ ) and household size ( $X^2(20) = 119.427, p < .001$ ).

Table 3. Demographic characteristics of sample by selection method, 100% landline sample, unweighted.

		Probability method	Quasi-probability birthday methods				Nonprobability-In household methods			
	Value in population (2013 CPS figures)	R-B-P (1)	Most recent (2)	MultiQ most recent (3)	Last (4)	Next (5)	Youngest adult (6)	Youngest male/ Youngest female (7)	MultiQ Youngest male/ Youngest female (8)	You mal You fem (9)
Male	48.3	39.1	43.7	42.9	42.5	45.1 <sup>a</sup>	33.2 <sup>b</sup>	44.0 <sup>a</sup>	41.0	40.0
White (alone or in combo)	80.8	89.6	87.9	87.7	87.7	90.0	87.3	86.3	86.9	87.0
Hispanic	15.0	3.9	3.5	3.2	3.7	4.1	3.4	4.4	3.1	4.7
Black (alone or in combo)	12.9	5.5	7.8	7.1	7.5	6.8	8.1	7.7	8.9	7.8
18–24 yrs	12.7	1.7	2.1	1.8	2.1	1.8	6.7 <sup>c</sup>	5.2 <sup>d</sup>	4.4 <sup>d</sup>	4.7 <sup>e</sup>
25–34 yrs	17.6	3.2	3.0	3.9	4.3	3.2	4.9	5.2	3.8	3.5
35–44 yrs	16.8	9.9 <sup>e</sup>	7.4	7.2	8.8	8.4	7.9	8.0	5.7	7.9
45–54 yrs	18.3	12.4	13.4	11.8	14.4	14.4	13.0	12.1	13.0	14.0
55–59 yrs	8.8	9.9	11.6	8.5	10.8	9.5	11.1	9.2	10.0	11.0
60+ yrs	25.7	62.9	62.5	66.8 <sup>f</sup>	59.7	62.7	56.3 <sup>g</sup>	60.1	62.1	58.0
College grad	28.9	40.2	43.9	42.2	42.1	42.5	41.2	41.3	37.0	41.0
Married	53.2	50.2	57.7 <sup>h</sup>	52.1	55.2	55.9	53.6	50.2 <sup>i</sup>	51.0	57.0
1 Adult in HH	33.6	40.6 <sup>j</sup>	30.8	40.3 <sup>j</sup>	32.9	32.5	33.8	37.0 <sup>k</sup>	41.6 <sup>j</sup>	31.0
2 Adults in HH	50.3	49.7	54.2	47.4 <sup>l</sup>	54.3	55.1	53.9	49.8	48.2	50.0
3+ Adults in HH	16.1	9.6 <sup>m</sup>	15.0	12.3	12.7	12.4	12.3	13.2	10.1 <sup>n</sup>	17.0
N	-	1,002	1,001	998	1,000	1,001	999	1,000	1,000	999

\* All differences listed below are significant using the Benjamini-Hochberg procedure with a false discovery rate equal to 0.05 or 5%.

<sup>a</sup> Significantly different from group 1

<sup>b</sup> Significantly different from all other groups

<sup>c</sup> Significantly different from groups 1, 2, 3, 4, 5, 8, 10, and 11

<sup>d</sup> Significantly different from groups 1, 2, 3, 4, and 5

<sup>e</sup> Significantly different from group 8

<sup>f</sup> Significantly different from groups 4, 6, 7, 9, 10, and 11

<sup>g</sup> Significantly different from groups 1, 2, 5, and 8

<sup>h</sup> Significantly different from groups 1, 7, and 8

<sup>i</sup> Significantly different from group 9

<sup>j</sup> Significantly different from groups 2, 4, 5, 6, 9, 10, and 11

<sup>k</sup> Significantly different from groups 2, 9, and 11

<sup>l</sup> Significantly different from group 5

<sup>m</sup> Significantly different from groups 2, 7, 9, 10, and 11

<sup>n</sup> Significantly different from group 2

<sup>o</sup> Significantly different from groups 3, 4, 5, 6, 7, and 8

<sup>p</sup> Significantly different from groups 3 and 8



### ***Integrating a Cell Phone sample***

To construct a sample that mimics a dual-frame study, a random sample of 1,000 cell phone interviews was drawn and combined with the landline samples. The same cell phone cases were combined with each of the within-household selection treatment samples to yield a sample of 50% cell phone and 50% landline completes. Since the same cell phone respondents were combined with each treatment group, any variation between treatment groups is due to the variability in the landline sample.

Two sets of weights were created for each of the 11 landline and cell phone samples, following procedures outlined by Kennedy (Kennedy 2008). The first weight included adjustments for probability of selection, nonresponse, and the overlapping dual frame design. The second weight is the final post-stratified weight, which matched targets from the U.S. Census Bureau by age, gender, ethnicity, race, population density, and region, and National Health Interview Survey (NHIS) estimates for phone status.

### **RESULTS WITH DUAL FRAME WEIGHTING**

Table 4 shows the demographic distributions for the landline/cell phone samples by each of the selection methods, prior to post-stratification. There are very few significant differences between methods when combined with cell phone sample. Gender was significantly different across groups ( $X^2(10) = 18.099, p = .05$ ). YMOF resulted in significantly more males than any other treatment group. In fact, combining this method with a cell phone sample only exaggerated the overrepresentation of males. When combined with the cellphone frame, youngest adult in the household still underrepresented males. The differences in household size also remained significant across treatment groups ( $X^2(20) = 45.199, p = .001$ ) when combined with the cell phone sample.

Although youngest adult in the household continued to perform best in terms of age, this was not statistically different from the other treatment groups ( $X^2(50) = 31.00, p = .98$ ). Treatment group comparisons of Hispanic ( $X^2(30) = 10.826, p = .999$ ); white ( $X^2(30) = 3.889, p = 1.00$ ); black ( $X^2(30) = 3.379, p = 1.00$ ); education level ( $X^2(50) = 23.441, p = 1.00$ ); and marital status ( $X^2(50) = 31.443, p = .981$ ) were also not significantly different when combined with the cell phone sample.

### **RESULTS WITH POSTSTRATIFICATION WEIGHTING**

Although it is useful to look at characteristics of the underlying unweighted sample, the demographic differences between treatment groups may be of little consequence once poststratification weighting is applied. In fact, once the sample is poststratified, significant age, gender and household size differences no longer exist.

Table 4. Demographic characteristics of sample by selection method, 50% landline + 50% cell phone, dual frame weighting.

		Probability method	Quasi-probability birthday methods				Nonprobability–In household methods			
	Value in population (2013 CPS figures)	R-B-P (1)	Most recent (2)	MultiQ Most recent (3)	Last (4)	Next (5)	Youngest adult (6)	Youngest male/ Youngest female (7)	MultiQ Youngest male/ Youngest female (8)	You mal You fem (9)
Male	48.3	49.1	50.6	51.4	50.8	51.8	45.3 <sup>a</sup>	51.3	50.2	50.0
White (alone or in combo)	80.8	81.6	80.9	81.0	80.8	82.0	80.6	80.0	79.9	80.0
Hispanic	15.0	10.4	9.8	10.1	10.1	10.2	10.0	10.6	10.2	10.0
Black (alone or in combo)	12.9	8.6	9.8	9.3	9.6	9.1	9.5	9.7	10.4	9.5
18–24 yrs	12.7	10.7	10.7	10.7	10.7	10.6	13.8	13.1	12.6	12.0
25–34 yrs	17.6	13.1	12.5	13.4	13.3	12.8	13.9	13.8	13.8	12.0
35–44 yrs	16.8	14.4	12.6	13.1	13.3	13.1	13.0	13.0	12.1	13.0
45–54 yrs	18.3	15.0	15.5	14.4	15.8	16.4	14.8	14.7	15.3	16.0
55–59 yrs	8.8	8.8	10.5	8.4	9.7	8.4	9.2	8.6	9.5	9.5
60+ yrs	25.7	37.9	38.2	40.0	37.2	38.8	35.3	36.7	36.6	35.0
College grad	28.9	39.8	41.0	41.2	40.4	40.3	39.9	40.2	38.3	40.0
Married	53.2	55.0	58.1	55.6	56.5	57.2	55.5	53.7	54.4	57.0
1 Adult in HH	33.6	18.4 <sup>c</sup>	14.9 <sup>d</sup>	18.3 <sup>e</sup>	15.7	15.5	16.1	16.9	18.9 <sup>f</sup>	14.0
2 Adults in HH	50.3	54.4	54.8	52.7	55.4	56.1	55.1	53.5	53.3	52.0
3+ Adults in HH	16.1	27.7	30.3	29.0	28.8	28.4	28.8	29.5	27.8	32.0
N	-	2,002	2,001	1,098	2,000	2,001	1,999	2,000	2,000	1,999

\* All differences listed below are significant using the Benjamini-Hochberg procedure with a false discovery rate equal to 0.05 or 5%.

<sup>a</sup> Significantly different from groups 2, 3, 4, 5, 7, 8, 9, 10, and 11

<sup>b</sup> Significantly different from all other groups

<sup>c</sup> Significantly different from groups 2, 4, 5, 9, 10, and 11

<sup>d</sup> Significantly different from groups 3 and 8

<sup>e</sup> Significantly different from groups 4, 5, 9, 10, and 11

<sup>f</sup> Significantly different from groups 4, 5, 6, 9, 10, and 11

## SUBSTANTIVE RESULTS

This analysis also explored differences in two key attitudinal variables by selection method—approval of the president and party identification—by selection method.

Table 5. Demographic characteristics of sample by selection method, 50% landline + 50% cell phone, poststratified.

		Probability method	Quasi-probability birthday methods				Nonprobability-In household methods			
	Value in population (2013 CPS figures)	R-B-P (1)	Most recent (2)	MultiQ Most recent (3)	Last (4)	Next (5)	Youngest adult (6)	Youngest male/ Youngest female (7)	MultiQ Youngest male/ Youngest female (8)	Youngest male/ Youngest female (9)
Male	48.3	48.9	49.4	49.1	49.4	49.3	48.6	49.7	49.0	49.4
White (alone or in combo)	80.8	78.0	78.4	78.4	78.5	78.1	77.9	77.7	78.3	77.7
Hispanic	15.0	13.9	14.0	13.6	14.1	13.9	13.5	14.1	13.7	14.2
Black (alone or in combo)	12.9	12.0	12.0	12.2	12.3	12.5	12.7	12.7	12.6	12.8
18–24 yrs	12.7	13.4	13.7	13.4	13.4	13.4	14.7	14.3	14.2	14.8
25–34 yrs	17.6	13.4	13.7	14.7	14.6	13.7	13.4	14.5	14.5	13.3
35–44 yrs	16.8	17.1	16.3	16.2	16.8	16.9	16.2	16.3	15.6	16.4
45–54 yrs	18.3	18.6	18.9	18.4	18.8	19.6	18.2	18.2	18.9	18.3
55–59 yrs	8.8	8.7	8.4	8.0	8.9	7.4	9.1	8.1	8.4	8.7
60+ yrs	25.7	28.7	29.0	29.3	27.5	29.1	28.3	28.6	28.4	28.3
College grad	28.9	30.0	30.2	30.0	29.9	30.1	29.8	29.8	29.9	29.8
Married	53.2	54.0	55.4	53.6	54.4	55.2	54.0	53.0	53.6	54.4
1 Adult in HH	33.6	16.8	13.6	16.5	15.1	14.5	15.1	15.4	16.8	13.6
2 Adults in HH	50.3	52.5	52.7	51.4	53.7	53.6	54.0	52.1	53.3	51.4
3+ Adults in HH	16.1	30.7	33.7	32.1	31.3	31.9	30.9	32.5	30.0	34.2
N	-	2,002	2,001	1,098	2,000	2,001	1,999	2,000	2,000	1,999

\* There were no significant differences using the Benjamini-Hochberg procedure with a false discovery rate equal to 0.05 or 5%.

There were no statistically significant differences between groups on job approval or party identification in the unweighted landline-only samples. This finding is consistent with Gaziano's meta-analysis (Gaziano 2005), which also concluded that the various selection methods resulted in very few differences on key substantive items.

Although many of the presidential approval and party identification estimates changed when the landline sample was combined with the cell phone sample, there were no significant differences between treatment groups, for either the sample weighted for selection probability or with the final poststratification weight.

Table 6. Attitudinal characteristics of sample by selection method, 100% landline, unweighted.

	Probability method	Quasi-probability birthday methods				Nonprobability–In household methods			
	R-B-P	Most recent	MultiQ Most recent	Last	Next	Youngest adult	Youngest male/ Youngest female	MultiQ Youngest male/ Youngest female	Youngest male/ Youngest female
Approve Obama	42.8%	43.4%	42.5%	39.6%	40.5%	42.0%	43.2%	42.3%	41.9%
Disapprove Obama	51.3%	51.1%	51.1%	54.9%	54.1%	52.2%	50.0%	52.3%	51.4%
Republican	26.1%	26.5%	25.6%	29.3%	28.5%	29.4%	25.7%	28.0%	26.7%
Democrat	33.1%	30.6%	32.7%	28.0%	30.3%	30.0%	30.1%	29.9%	31.1%
Independent	37.1%	39.3%	37.5%	38.5%	37.0%	36.0%	40.2%	36.9%	37.0%
N	1,002	1,001	998	1,000	1,001	999	1,000	1,000	999

\* There were no significant differences using the Benjamini-Hochberg procedure with a false discovery rate equal to 0.05 or 5%.

Table 7. Attitudinal characteristics of sample by selection method, 50% landline + 50% cellphone, dual frame weighting for probability of selection.

	Probability method	Quasi-probability birthday methods				Nonprobability–In household methods			
	R-B-P	Most recent	MultiQ Most recent	Last	Next	Youngest adult	Youngest male/ Youngest female	MultiQ Youngest male/ Youngest female	Youngest male/ Youngest female
Approve Obama	48.2	48.3	48.3	46.2	46.3	48.0	48.5	48.0	48.1
Disapprove Obama	51.8	51.7	51.7	53.8	53.7	52.0	51.5	52.0	51.9
Republican	27.6	27.5	27.0	29.3	28.9	29.1	27.1	28.6	27.7
Democrat	32.7	31.2	32.8	30.1	31.0	31.4	31.7	31.5	31.8
Independent	38.5	40.1	38.8	39.6	38.9	38.3	39.9	38.7	39.3
N	2,002	2,001	1,098	2,000	2,001	1,999	2,000	2,000	1,999

\* There were no significant differences using the Benjamini-Hochberg procedure with a false discovery rate equal to 0.05 or 5%.

## Conclusions

There is no clear evidence, based on the findings of this study, that one within-household selection method is preferable for landline frames when conducting a dual-frame telephone survey. When taking into account response rates, accuracy of selection, demographics, and substantive differences, there is not one method that clearly outperforms all others. A decision of which method to use may come down to other considerations not directly explored by this study, such as the target population, the need to adopt a probability-based method, and budgetary and timeline constraints.

Many nonprobabilistic methods attempt to compensate for nonresponse bias by purposefully selecting respondents who tend to be underrepresented, but after combining the landline sample with the cell phone sample, we found few significant demographic or substantive differences between the

Table 8. Attitudinal characteristics of sample by selection method, 50% landline + 50% cellphone, poststratified.

	Probability method	Quasi-probability birthday methods				Nonprobability–In household methods			Nonprobability	
	R-B-P	Most recent	MultiQ Most recent	Last	Next	Youngest adult	Youngest male/Youngest female	MultiQ Youngest male/Youngest female	Youngest male/Youngest female	
Approve Obama	49.1	49.2	48.8	46.6	46.6	48.3	49.1	48.0	48.9	4
Disapprove Obama	50.9	50.8	51.2	53.4	53.4	51.7	50.9	52.0	51.1	5
Republican	27.9	28.2	26.6	30.0	29.0	29.4	27.6	28.7	27.1	2
Democrat	32.4	31.2	32.3	30.1	30.5	31.1	31.6	31.1	31.8	3
Independent	38.5	39.5	39.4	38.5	39.2	38.2	39.6	39.0	39.9	3
N	2,002	2,001	1,098	2,000	2,001	1,999	2,000	2,000	1,999	1

\* There were no significant differences using the Benjamini-Hochberg procedure with a false discovery rate equal to 0.05 or 5%.

selection method groups. Given that the nonprobability methods had no clear advantage in terms of representativeness of the demographics, researchers may find it preferable to use a probability or quasi-probability based method, which adhere to the underlying probability assumptions of most sample surveys and statistical analyses.

Considering each method individually, we did find possible opportunities for improvement. Our modified R-B-P method deviated from the procedures outlined by Rizzo et al. (Rizzo, Brick, and Park 2004), and the refusal rates and average household size for this group indicate the Kish grid may have been too intrusive for households with more than three adults. Results may have differed if we had used a quasi-probability method for households with three or more adults, instead of the Kish grid.

If a quasi-probability method is used, it may be advisable to use next or last birthday, rather than most recent birthday, which had the lowest accuracy rate. The “most recent” language may not have been as easily understood by respondents and is a finding that may have practical application to within household selection for other modes.

Of the three nonprobability at-home methods, the YMOF method significantly overrepresented males. Cell phone samples tend to result in a higher proportion of males than females (Kennedy, McGeeney, and Keeter 2016), and this overrepresentation persisted when YMOF was combined with the cell phone sample. Based on the findings of this study, it seems prudent to avoid YMOF when conducting a dual-frame study.

This study is not without limitations. We assigned selection methods to sampled cases, and interviewers administered all treatment groups. The advantage of this design is that interviewer variance is distributed across methods, rather than within a method. The downside of this approach is that

interviewers switched methods with each contact, even though we found no evidence in quality reviews that interviewers had difficulty adapting to this change.

Our method for obtaining cellular cases was another limitation. Rather than collecting cellular cases for this specific study, we randomly selected cases from a parallel study that used identical sampling and call design methodology but with a slightly different survey instrument. Our results may have been different if the cellular respondents had completed the same survey instrument as the landline respondents.

Although response rates for this study are in line with comparable surveys conducted by Gallup and other commercial firms (Keeter et al. 2017), response rates were considerably lower than what might be expected from studies with more rigorous call designs. It is unknown how a more extensive call design would change the results of this experiment.

It is conceivable that in the near future the majority, if not all, telephone surveys will be completed from the cellular frame, making within household selection irrelevant for telephone. However, the results of this study highlight designs that result in improved accuracy and cooperation. These comparisons have possible applications to the design of face-to-face and mail modes, which will continue to rely on within household selection.

## REFERENCES

- American Association for Public Opinion Research. 2016. "Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys (6th Ed)." [http://www.aapor.org/AAPOR\\_Main/media/publications/Standard-Definitions20169theditionfinal.pdf](http://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf).
- Battaglia, M., M. Link, M. Frankel, L. Osborn, and A. Mokdad. 2008. "An Evaluation of Respondent Selection Methods for Household Mail Surveys." *Public Opinion Quarterly* 72 (3): 460–69.
- Binson, D., J. Canchola, and J. Catania. 2000. "Random Selection in a National Telephone Survey. A Comparison of the Kish, next-Birthday, and Last-Birthday Methods." *Journal of Official Statistics* 16 (1): 53–59.
- Gaziano, C. 2005. "Comparative Analysis of Within-Household Respondent Selection Techniques." *Public Opinion Quarterly* 69 (1): 124–57.
- Groves, R., and R. Kahn. 1979. *Surveys by Telephone: A National Comparison with Personal Interviews*. New York: Academic Press.
- Keeter, S., N. Hatley, C. Kennedy, and L. Arnold. 2017. "What Low Response Rates Mean for Telephone Surveys." Pew Research Center. <http://www.pewresearch.org/2017/05/15/what-low-response-rates-mean-for-telephone-surveys/>.
- Kennedy, C. 2008. "Evaluating the Effects of Screening Telephone Service in Dual Frame RDD Surveys." *Public Opinion Quarterly* 71 (5): 750–71.
- Kennedy, C., K. McGeeney, and S. Keeter. 2016. "The Twilight of Landline Interviewing." Pew Research Center. <http://www.pewresearch.org/2016/08/01/the-twilight-of-landline-interviewing/>.
- Kish, L. 1949. "A Procedure for Objective Respondent Selection within the Household." *Journal of the American Statistical Association* 44 (247): 380–87.
- . 1965. *Survey Sampling*. New York: Wiley.
- Martin, E. 2007. "Strength of Attachment: Survey Coverage of People with Tenuous Ties to Residence." *Demography* 44 (2): 427–40.
- Olson, K., and J. Smyth. 2014. "Accuracy of Within-Household Selection in Web and Mail Surveys of the General Population." *Field Methods* 26 (1): 56–69.
- Olson, K., M. Stange, and J. Smyth. 2014. "Assessing Within-Household Selection Methods in Household Mail Surveys." *Public Opinion Quarterly* 78 (3): 656–78.
- O'Rourke, D., and J. Blair. 1983. "Improving Random Respondent Selection in Telephone Surveys." *Journal of Marketing Research* 20 (4): 428–32.
- Paisley, W., and E. Parker. 1965. "A Computer-Generated Sampling Table for Selecting Respondents within Households." *Public Opinion Quarterly* 29 (3): 431–336.
- Rizzo, L., J.M. Brick, and I. Park. 2004. "A Minimally Intrusive Method for Sampling Persons in Random Digit Dial Surveys." *Public Opinion Quarterly* 68 (2): 267–74.
- Salmon, C., and J. Nichols. 1983. "The Next-Birthday Method of Respondent Selection." *Public Opinion Quarterly* 47 (2): 270–76.
- Tourangeau, R., G. Shapiro, A. Kearney, and L. Ernst. 1997. "Who Lives Here? Survey Undercoverage and Household Roster Questions." *Journal of Official Statistics* 13 (1): 1–18.