A Meta-analysis of Within-Household Respondent Selection Methods on Demographic Representativeness

Ting Yan
Roger Tourangeau
Rose McAloon

Westat
1600 Research Blvd
Rockville MD 20850

Introduction

Surveys of general population usually start with a probability sample of households identified by telephone numbers or mailing addresses. However, a probability sample of households does not automatically translate into a probability sample of people with characteristics of interest. Random selection of a respondent within a sampled household is essential for maintaining the probability nature of the resulting sample and for making inference from the household sample to the general population. The latest review (Gaziano, 2005) lists at least 15 selection methods that have been used to generate a sample of persons from a probability sample of households. The fifteen selection methods are grouped into three categories based on the probability nature of the resulting sample of persons – probability methods, quasi-probability methods, and non-probability methods.

Probability methods require a listing of all people living in the sampled household in order to compute the probability of selection for each individual. The Kish method (Kish, 1949), Age-order or Age-only method (Denk and Hall, 2000; Forsman, 1993), and full enumeration methods (Denk et al., 2000; Srinivasan, Christiansen, and Tortora, 1996) fall into this category.[1] Quasi-probability methods bypass household listing in order to reduce the perceived intrusiveness and sensitivity associated with household listing and to decrease the administration time. Typical examples of quasi-probability selection methods are birthday methods (Salmon & Nichols, 1983), including the next birthday method and the last birthday or most recent birthday method. Non-probability methods were created to streamline selection process and only tried to approximate population age and gender distributions. Troldahl and Carter method (Troldahl and Carter, 1964) is a typical example of non-probability methods; so are various modifications or variants on the Troldahl and Carter method such as Paisley and Parker Standard Modification (1965), Bryant's Correction for Too Many Females (1975), Groves and Kahn's Modification (1979), Czaja-Blair-Sebestik modification (1982), Hagan-Collier Alternative (1983), Youngest Male Oldest Female (YMOF) (Srinivasan et al., 1996).

The major advantage of probability methods lies in their ability to produce consistent and unbiased estimates. However, it is commonly believed that the listing of all household members add to the length of the interview and increase the likelihood of encountering household refusals due to the perceived sensitivity and burden of the listing questions. Birthday methods allow all household members to have an equal chance of selection under the assumption that births are random. However, births are not truly random and tend to heap in certain months. Therefore, the increase in cooperation and reduction of cost associated with birthday methods is achieved through sacrificing true randomness. Non-probability methods trade randomness for increased cooperation and reduced cost.

Two recent selection methods are conditional on household size. The Rizzo-Brick-Park method (Rizzio, Brick, and Park, 2004) starts by asking the household contact the total number of adults living in the household. If there is only one adult in the household, that adult is chosen for the interview. If there are two or more adults in a household, one adult is randomly selected with a probability equal to the inverse of the total number of adults. If the household informant is selected, the selection process ends. If the household informant is not selected and there are two adults

[1] Interested readers are referred to Gaziano (2005) for a detailed description of each method.

in the household, the informant is told that the other adult is selected for the interview and the selection process ends. If the household informant is not selected and there are more than two adults in the household, another selection method (such as the Kish method or the last birthday method) can be used to select an adult after excluding the household contact. This method takes the advantage of the fact that more than eight in ten households in the United States have two or one adults. As a result, this method has the potential to significantly reduce the effort required to make a selection.

The Le-Brick-Diop-Alemadi method is proposed for areas with larger households (Le, Brick, Diop, and Alemadi, 2013). This method also starts by asking a household contact the number of adults living in the household. For households with one or two adults, no additional question is asked and the household contact is automatically selected in one-person households and randomly selected half of the time for two-person households. For households with three or four adults, again no additional question is needed and the household contact is selected either 33% or 25% of the time. When the household contact is not selected, the older or the younger of the two adults in a three-person household is selected and the oldest, the youngest, or the second oldest of the other three adults in a four-person household is randomly selected with an equal chance. For households with five or more adults, one more question is asked about the number of males in the household and selection is made based on the answer. Obviously, both the Rizzo-Brick-Park method and the Le-Brick-Diop-Alemadi method are suitable for interviewer-administered modes or computerized self-administered modes rather than mail surveys.

A good within-household respondent selection method should be able to randomly select a respondent within the household without appearing intrusive or burdensome to potential respondents. Gaziano's review (2005) compared and evaluated a few selection methods on three dimensions – demographic representativeness, cooperation and refusal rates, and cost. Qualitative conclusions about the potential impact of various selection methods are drawn in the review. For instance, the Kish method maybe is not as intrusive as feared. The last birthday method is advantageous in cooperation rates and cost but not in demographic representation. The next birthday method is found to be more confusing than the last birthday method. Some nonprobability methods (e.g., the YMOF method) are slightly cheaper and better at demographic representation.

What is missing from the review (and the survey literature at large) are quantitative measures of the impact of different within-household selection methods on these three dimensions; as a result, it is hard for survey practitioners to properly use the literature to inform decisions on which respondent method one should use and what would happen if a different method were used. As concluded by Gaziano, "little systematic, accessible evidence exists to guide choice of respondent selection method" (2005, 124).

Meta-analysis of empirical studies on within-household selection methods produces quantitative effect sizes that can be used to inform decisions on respondent household selection. One meta-analysis (Yan, 2009) attempted to produce effect sizes (odds ratios) to quantify the effect of different within-household selection methods on the odds of obtaining completed interviews (as a measure of cooperation rates) and obtaining female respondents (as a measure of demographic representativeness). As shown in Table 1, the meta-analysis found that the Kish method has lower odds to produce completed interviews than the last birthday method as well as any other method, indicating that the Kish method is perceived to be more intrusive and burdensome than the other selection methods. As a matter of fact, the probability methods in general have lower odds to obtain completed interviews than other selection methods, again speaking to the perceived sensitivity associated with the listing process of the probability methods. The odds of producing completed interviews under birthday methods (including the last birthday method) are not statistically different from the odds under other selection methods (e.g., the YMOF method). When it comes to the representation of females in the resulting sample, the odds fir probability methods (including the Kish method) to find female respondents are about 0.9 times of the odds for other methods. The odds for birthday methods to recruit females are higher than the odds for other methods. The odds of having females in the sample when no selection is done are 1.3times higher than when any selection method is used to select a respondent. This meta-analysis clearly shows that probability methods (including the Kish method) reduce the over-representation of female respondents at a cost of lower cooperation and birthday methods over-represent females without clear advantage in cooperation.

Table 1: Meta-Analytic Results from Yan (2009)

	Mean Effect Sizes (Odds Ratio)	
	Completes vs. Refusals	Females vs. Males
Probability vs. Other	0.79*	0.92*
Kish vs. Last Birthday	0.78*	0.89*
Kish vs. Other	0.79*	0.91*
Birthday vs. Other	1.13	1.12*
Last Birthday vs. Youngest Male Oldest Female	0.92	1.33*
No Selection vs. Any selection		1.30*

Note: Odds ratios with an asterisk are statistically significant at p=0.05 level.

It is not clear, however, from Yan (2009)'s meta-analysis which selection method yields demographic distributions closer to the population distribution. This paper continues and extends Yan (2009)'s meta-analysis to address this issue in particular. The primary goal of this paper is to present a quantitative summary of empirical studies on the accuracy of different selection methods in demographic representativeness.

Meta-analysis Method

Selection of Studies

We went back to the articles included in Yan (2009) and supplemented it with a new search for empirical studies of within-household selection methods in various databases available (e.g., JSTOR, Ebsco, LexisNexis, PubMed) and online search engines (e.g., Google Scholars), using as key words within-household, respondent, selection, method, survey. We also searched *the Proceedings of the Survey Research Methods Section of the American Statistical Association*. These proceedings publish papers presented at two major conferences of survey methodologists (the Joint Statistical Meeting and the annual conferences of the American Association for Public Opinion Research) where survey methods studies are often presented.

We included in this meta-analysis studies that provided quantitative information on demographic distribution, focusing on studies that reported both sample estimates and population benchmarks. Studies that provided statistics not appropriate for meta-analysis (e.g., proportions of respondents incorrectly selected) are excluded from this meta-analysis. For studies that only reported sample estimates (e.g., proportion of female respondents) and didn't provide population benchmarks, we went to the Census tables published by the US Census Bureau to find the population proportions corresponding to the year in which the study was conducted. We dropped from this meta-analysis studies for which we couldn't find the corresponding population proportions either because the demographic characteristics do not match the Census definition or population benchmarks for local areas are not available in the Census tables.

A total of 27 research papers reporting empirical results met our inclusion criteria and are listed in Appendix 1. Displayed in Table 2 are the number of research papers and the number of sample estimates included in these research papers by the year in which the studies were conducted. The sample estimates are more likely to come from studies conducted in the 1990s or 2010s.

Table 2: Studies Included in Meta-Analysis

Year of Study	Number of Papers	Number of Sample Estimates Reported
1980s	7	90
1990s	9	246
2000s	6	150
2010s	5	244
Total	27	730

Analytic Procedures

As measures of demographic representativeness, we compared sample estimates of proportions of respondents with certain demographic characteristics (such as gender, race, and marital status) reported in a study to the corresponding population proportions and created two effect size measures. The first is simply the difference between a sample proportion and its population counterpart, which essentially is the bias in the sample proportion. For instance, if one study reported a proportion of females in the sample as 48% and the population proportion of females is 50%, then the bias for this sample estimate of the proportion of females is -2%. In other words, this study underrepresented females by two percentage points. We used bias to calculate overall effect sizes when the direction of bias for estimates is known from the literature.

The second effect size measure is absolute bias. In the same example mentioned above, the absolute bias for this estimate of proportion of females is 2%. Absolute bias is used when deviation from population estimates is of more importance than direction.

For each study, we also coded the specific within-household respondent selection method used. Mean overall effect sizes are calculated for each *type* of selection methods and for each popular selection method.

We carried out all meta-analyses reported in this paper using SAS's PROC SURVEYMEANS. The SAS procedure calculates the overall effect size for a particular within-household selection method ($\overline{ES}$ below) as the weighted average of effect sizes across all sample estimates obtained under that particular selection method (ES_i below), taking into account the clustering of individual sample estimates within a study.

$$\overline{ES} = \sum_i (w_i * ES_i) \,/ \sum_i w_i$$

The weight w_i is the inverse of the variance v_i for estimate *i:*

$$w_i = \frac{1}{v_i} = n_i/(p_i * (1 - p_i))$$

where $p_i$is a sample proportion reported in a study for estimate i and n_i is the sample size for estimate i.

PROC SURVEYMEANS provides a "design-based" estimate of the standard error of the overall effect size (see, e.g., Wolter, 1985). It uses the variation in the (weighted) mean effect sizes across studies to calculate a standard error for the overall estimate, without making any assumptions about the variability of the individual estimates. Results from PROC SURVEYMEANS are largely consistent with results from the random-effects model as specified in Lipsey and Wilson (2001) (see, e.g., Tourangeau and Yan, 2007).

Results

We first examined the distribution of the 730 estimates by type of within-household selection methods. Close to half of the estimates (45%) are from studies employing birthday methods. Almost one-third of the sample estimates (28%) used some kind of non-probability selection methods. Probability methods produced 6% of the estimates whereas the Rizzo-Brick-Park method yielded 4 estimates. Close to one-fifth of the estimates are from studies where no selection method is used. In addition, close to 60% of the sample estimates (59%) are from interviewer-administered studies and 40% are from mail surveys. Only 13 estimates (2%) are from web surveys.

Impact of Within-household Selection Methods on Bias and Absolute Bias in Proportion of Females

We then examined the impact of within-household selection methods on the representation of females in the sample. The survey literature has documented that female respondents have a higher probability to answer the phone or the door bell and a higher probability to agree to a survey request than males (e.g., Groves and Couper, 1998). As a result, the selection method that reduces the representation of females is considered better. We focused on bias in estimates of proportion of females in our analysis but also presented absolute bias in Table 3.

Table 3. Impact of Within-Household Respondent Selection Methods on Proportion of Females

	Number of Estimates	Effect Size (Bias)		Effect Size (Absolute Bias)	
		Mean	SE	Mean	SE
Probability Methods	14	3.3%	0.7%	3.5%	0.6%
Kish	7	4.1%	0.6%	4.1%	0.6%
Birthday Methods	35	4.8%	1.0%	5.7%	0.7%
Last Birthday	19	4.9%	1.8%	6.5%	1.0%
Next Birthday	14	4.8%	0.7%	5.0%	0.8%
Non-Probability Methods	21	-0.1%	0.7%	2.9%	0.4%
Troldahl and Carter and variations	8	-1.7%	0.8%	2.1%	0.6%
Youngest Male Oldest Female	5	-0.9%	0.1%	2.8%	0.8%
Rizzo-Brick-Park	2	3.3%	2.2%	3.3%	2.2%
No Selection	11	8.0%	2.5%	8.5%	2.5%

Several trends can be noted from Table 3. First, sample estimates of the prevalence of females are statistically different from the population benchmark when no within-household respondent selection is used; females are overrepresented by an average of eight percentage points. Second, probability methods and birthday methods also produce estimates of female proportion statistically significant from population benchmarks; both methods resulted in an overrepresentation of females. Third, the average bias in sample estimates of proportion of females is close to zero for non-probability methods and is statistically different from average bias in sample estimates produced by other methods. The average absolute bias in estimates of female proportion for non-probability methods is close to 3%.

With regards to the effect of specific selection method on sample estimates of females, the Kish method and the two birthday methods have an average bias of 3 to 5%, leading to an overrepresentation of females in the resulting samples. The average bias in sample estimates of female proportion under the Troldahl and Carter methods is not statistically significant from zero whereas the average bias in sample estimates under the YMOF method is -0.9%, statistically different from 0. Unlike all other methods, the YMOF leads to a slight underrepresentation of females. In addition, the standard errors for the overall effect sizes are larger for birthday methods (especially the last birthday method) and when no selection is done.

Impact of Within-Household Respondent Selection Methods on Age

Studies reported age distributions in various ways. To examine the impact of selection methods on the age distribution, we conducted two sets of analyses. The first set looks into sample estimates of the proportion of old people. Old people are defined as those aged 60 or above for some studies and aged 65 or above for other studies. Survey literature demonstrates that older people have a higher likelihood to respond to survey requests than younger people (e.g., Groves and Couper 1998). We focus on the bias in the resultant estimates of old people and show both bias and absolute bias in Table 4.

Table 4. Impact of Within-Household Respondent Selection Methods on Proportion of Old People

	Number of Estimates	Effect Size (Bias)		Effect Size (Absolute Bias)	
		Mean	SE	Mean	SE
Probability Methods	4	4.6%	1.8%	4.6%	1.8%
Kish	4	4.6%	1.8%	4.6%	1.8%
Birthday Methods	16	5.5%	1.9%	6.4%	1.5%
Last Birthday	7	4.4%	1.3%	4.7%	1.2%
Next Birthday	9	5.9%	2.8%	7.3%	2.2%
Non-Probability Methods	14	3.2%	1.6%	4.2%	1.4%
Troldahl and Carter and variations	6	3.6%	1.7%	4.6%	1.4%
Youngest Male Oldest Female	2	-1.0%	1.0%	1.6%	0.3%
No Selection	10	2.1%	2.1%	5.2%	1.5%

It is clear from Table 4 that all within-household selection methods overrepresented old people. Furthermore, average bias in estimates of the prevalence of old people is not worse when no household selection is done than when a household selection method is implemented. Non-probability methods don't seem to have the same advantage in producing accurate estimates of the proportion of old people as they do with estimating the proportion of females.

The second set of analyses make use of estimates of proportions of all age categories reported in the studies instead of focusing only on the proportion of people aged 60 or 65 and above. We calculated mean absolute bias in all estimates pertaining to age proportions by different selection methods. The goal is to compare deviations from population proportions across all age categories when different selection methods are used. As shown in Table 4, estimates of age proportions are statistically different from population benchmarks for all selection methods. Birthday methods (especially the next birthday method) produce the largest deviation.

Table 5. Impact of Within-Household Respondent Selection Methods on Age Proportions

	Number of Estimates	Effect Size (Absolute Bias)	
		Mean	SE
Probability Methods	17	4.6%	0.7%
Kish	17	4.6%	0.7%
Birthday Methods	94	7.4%	1.5%
Last Birthday	37	4.0%	1.3%
Next Birthday	57	9.2%	1.5%
Non-Probability Methods	75	4.7%	0.5%
Troldahl and Carter and variations	28	5.3%	0.4%
Youngest Male Oldest Female	19	3.9%	1.3%
No Selection	45	5.8%	1.4%

Impact of Within-Household Respondent Selection Methods on Other Demographic Representation

We also examined absolute bias in sample estimates of demographic proportions pertaining to education, employment, race and ethnicity, marital status, and income. Due to the dramatic differences in demographic categories reported by the studies, we decided to investigate mean absolute bias in all proportion estimates measuring, for instance, education, rather than looking at estimates of the proportion of a particular education category (e.g., proportion of people with a college degree). Displayed in Figure 1 are mean absolute bias in sample estimates by types of selection methods and demographic characteristics of interest (Appendix II provides the number of estimates included in calculation of each overall effect size, the mean overall effect sizes, and the standard errors of the mean overall effect sizes).

Average absolute bias in sample estimates is significantly different zero for all cases except two. Average bias in racial estimates under probability methods is not statistically significant from zero, but it is an average calculated over only two estimates. Average bias in employment estimates under non-probability methods, calculated over 4 individual estimates, is also not statistically significant from zero. A common trend apparent from Figure 1 is that sample estimates are further off from population benchmarks when no within-household selection method is used across all five demographic domains. The three types of selection methods do not differ significantly from each other in terms of deviation from population benchmarks across the five demographic domains. Furthermore, there are large variations in absolute biases in sample estimates of income proportions.

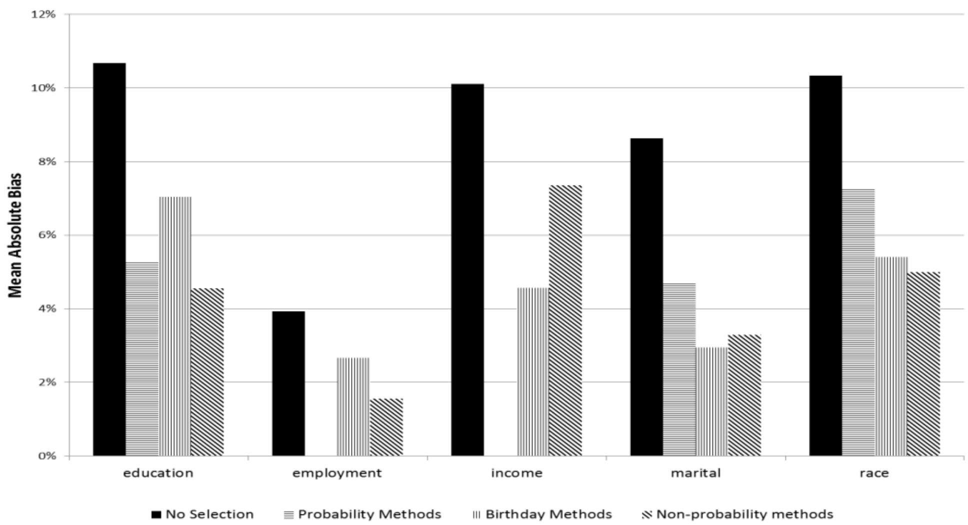

Figure 1. Impact of Within-Household Respondent Selection Methods on Other Demographic Representations

Discussion and Conclusions

This paper reports a meta-analysis on within-household selection methods. Meta-analysis systematically combines information from existing research and produces a quantitative summary of what is known. This paper takes advantage of the analytical power of meta-analysis to summarize findings across empirical studies on within-household selection methods and to produce quantitative effect sizes that can be used in future guidance.

Several conclusions are drawn. First, mean absolute bias in sample estimates of demographic representation is large when no within-household selection is carried out; sample estimates deviate from population benchmarks by 5 to 10 percentage points. This finding suggests that the implementation of no within-household selection results in inaccurate representation of demographic characteristics. Our meta-analysis clearly shows the danger of skipping within-household selection in a survey of general population.

Second, even though probability methods in general and the Kish method in particular are shown to effectively reduce the overrepresentation of females when compared to other methods (Yan, 2009), they still yield an overrepresentation of females. Consistent with Yan (2009), birthday methods produce positive bias in sample estimates of female prevalence. By contrast, non-probability methods produce gender distribution closest to the population distribution, especially the Youngest Male Oldest Female method.

Third, all selection methods over-represent old people. The next birthday method produces larger absolute bias in sample estimates of age proportions than other selection methods, producing the east accurate age representation.

Fourth, sample estimates of education, employment, race, marital status, and income deviate from population benchmarks for almost all types of selection methods. Selection methods do not differ significantly from each other in the amount of deviation.

The conclusions provide guidance for survey practitioners and researchers. Based on our findings, we definitely recommend future surveys to implement within-household selection whenever possible to ensure better representation of key demographic characteristics. In addition, we recommend any non-probability selection method because it produces accurate gender distribution without hurting representativeness of other demographic characteristics. Despite their wide-spread use in the survey field, birthday methods are neither advantageous in producing more completed interviews (Yan, 2009) nor advantageous in producing accurate demographic distributions. As a result, we don't recommend them over non-probability methods.

Besides quantitatively summarizing what is known, meta-analysis is also able to reveal what is needed from future research. Three limitations are noted. First, the calculation of overall effect sizes draws on a small number of sample estimates in some cases, limiting the power to make inferences. Second, the paper is missing comparisons of several specific techniques (e.g., the two methods conditional on household size) due to the lack of empirical studies and the lack of sufficient information reported. Third, some papers reported weighted estimates and other papers reported unweighted estimates. Still some papers do not clearly indicate whether their estimates are weighted or not. We have some evidence indicating that weighted estimates are in general closer to population benchmarks than unweighted estimates and our next step is to examine how weighting moderates the overall effect sizes by selection methods. Furthermore, we also plan to look into the moderating effect of the mode of data collection on the overall effect sizes by selection methods.

References

Battaglia, Michael, Michael W. Link, Martin F. Frankel, Larry Osborn, and Ali H. Mokdad. 2008. "An Evaluation of Respondent Selection Methods for Household Mail Surveys." *Public Opinion Quarterly* 72: 459-469

Beebe, Timothy J., Michael E. Davern, Donna D. McAlpine, and Jeanette K. Ziegenfuss. 2007. "Comparison of Two Within-Household Selection Methods in a Telephone Survey of Substance Abuse and Dependence." *Annals of epidemiology* 17:458-463.

Carr. Kathleen, and Joan Hertvik. 1993. "Within-Household Selection: Is Anybody Listening?" In *Proceedings of the Survey Research Methods Section of the American Statistical Association,* 1119-1123.

Czaja. Ronald. Johnny Blair, and Jutta P. Sebestik. 1982. "Respondent Selection in a Telephone Survey: A Comparison of Three Techniques." *Journal of Marketing Research* 19:381-85.

Denk. Charles E., Thomas M. Guterbock, and Daniel B. Gold. 1996. "Modelling Selection of Respondents Within Household in Telephone Surveys." Paper presented at the annual meeting of the American Association for Public Opinion Research, Sail Lake City, UT.

Denk, Charles E., and John W. Hall, 2000. "Respondent Selection in RDD Surveys: A Randomized Trial of Selection Performance." Paper presented at the annual meeting of the American Association for Public Opinion Research, Poriland, OR.

Forsman, Gosta. 1993. "Sampling Individuals Within Households in Telephone Surveys." In *Proceedings of the Survey Research Methods Section of the American Statistical Association,* 1113-1118.

Gaziano, Cecile. 2005. "Comparative Analysis of Within-Household Respondent Selection Techniques." *Public Opinion Quarterly* 69:124-157.

Grandjean, Burke D., Martha G. Leighty, Patricia A. Taylor, and Ying Xu. 2004. "Is Target Selection by Last Birthday 'Random Enough'? A Split-Ballot Test." *Proceedings of the Survey Research Methods Section, American Statistical Association*. Pp.4789-4793.

Groves, Robert M. and Mick P. Couper. 1998. *Nonresponse* in *Household Interview Surveys.* New York: Wiley.

Goyder, John, Maryanna Basic, and M. E. Thompson. 2001. "Within Household Sample Selection: A Reassessment." Paper presented at the annual meeting of the American Association for Public Opinion Research, Montreal, Canada. (Revised post-meeting version.)

Hagen, Dan E. and Charlotte M. Collier. 1983. "Must Respondent Selection Procedures for Telephone Surveys Be Invasive?" *Public Opinion Quarterly* 47: 547-556.

Hicks, Wendy and David Cantor. 2002. "Evaluating Methods to Select a Respondent for a General Population Mail Survey." Paper presented at the Annual Conference of the American Association for Public Opinion Quarterly.

Hill, Craig A., Karen Donelan, and Martin R. Frankel. 1999. "Within-Household Respondent Selection in an RDD Telephone Survey: A Comparison of Two Methods." Paper presented at the annual meeting of the American Association for Public Opinion Research. St. Petersburg, FL.

Keeter, Scott. and Kevin Fisher. 1997. "A Comparison of 'Last Birthday' and 'Youngest Male/Oldest Female' Respondent Selection Procedures." Paper presented at the annual meeting of the American Association for Public Opinion Research. Norfolk. VA.

____.1997-98. "Comparison of Two Respondent Selection Methods." *National Network of State Polls (NNSP) Newsletter* 31 (Winter): 1-3. Available online at http://survey.rgs.uky.edu/nnsp/newsletters/nnsp31.pdf (accessed September 2009).

Kennedy, John M. 1993. "A Comparison of Telephone Survey Respondent Selection Procedures." Paper presented al the annual meeting of the American Association for Public Opinion Research, St. Charles, IL. Available online at http://www.indiana.edu/~csr/aapor93.html (accessed September 2009).

Lavrakas, Paul J., and Sandra L. Bauman (1993). "The Last-Birthday Selection Method and Within-Unit Coverage Problems." In *Proceedings of the Survey Research Methods Section, American Statistical Association*. Pp.1107-1112.

Le, K. T., Brick, J. M., Diop, A., and Alemadi, D. 2013. "Within-household Sampling Conditioning on Household Size." *International Journal of Public Opinion Research*, 25, 108-118.

Link, Michael W., and Ali Mokdad. 2006. "Can Web and Mail Survey Modes Improve Participation in an RDD-based National Health Surveillance?" *Journal of Official Statistics*, 22: 293-312.

Oldendick, Robert W., George F, Bishop, Susan B. Sorenson, and Alfred J. Tuchfarbor. 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, Matthew Stange, and Jolene D. Smyth. 2014. "Assessing Within-Household Selection Methods in Household Mail Surveys." *Public Opinion Quarterly*, 78:656-678.

Olson, Kristen, Matthew Stange, Jolene D. Smyth, and Paul Lavrakas (2015). Monograph on Within-household Selection Methods.

O'Rourke, Diane, and Johnny Blair. 1983. "Improving Random Respondent Selection in Telephone Surveys," *Journal of Marketing Research* 20:428-232.

O'Rourke, Diane and Edward Lakner. 1989. "Gender Bias: Analysis of Factors Causing Male Underrepresentation in Surveys." *International Journal of Public Opinion Research*, 1:164-176.

Prairie Research Associates, Inc. 2001. "Comparing Two Respondent Selection Methods in Telephone Surveys." Manitoba, Canada. Available online at http://www.pra.ca/resources/birthday_e.pdf (accessed September 2009).

Rizzo, L., J. Michael Brick, and Inho Park. 2004. "A Minimally Intrusive Method for Sampling Persons in Random Digit Dial Surveys." *Public Opinion Quarterly*, 68: 267-274

Sabin, Martha C. and Susan H. Godley. 1987. "Mental Health Citizen Surveys: A Comparison of Two Within Household Telephone Sampling Techniques." *Evaluation and Program Planning* 10: 137-141.

Srinivasan, Rajesh, Elaine H. Christiansen, and Robert Tortora. 1996. "Comparative Study of Three Respondent Selection Techniques for Telephone Survey Research." Paper presented at the annual meeting of the American Association for Public Opinion Research. Sail Lake City, UT.

Stange, Matthew, Jolene D. Smyth, and Kristen Olson. 2015. "Using a Calendar and Explanatory Instructions to Aid Within-household Selection in Mail Surveys." *Field Methods*, 28: 64-78.

Tarnai, John. Eugene A, Rosa, and Lesli Peterson Scott. 1987. "An Empirical Comparison of the Kish and the Most Recent Birthday Method for Selecting a Random Household Respondent in Telephone Surveys." Paper presented at the annual meeting of the American Association for Public Opinion Research, Hershey, PA.

Tourangeau, Roger, and Ting Yan. 2007. "Sensitive Questions in Surveys." *Psychological Bulletin,* 133: 859-833.

Wolter, Kirk. 1985. *Introduction to variance estimation*. New York:Springer-Verlag

Yan, Ting. 2009. "A Meta-analysis of Within-Household Respondent Selection Methods." In *Proceedings of the Survey Research Methods Section, American Statistical Association*. Pp.6134-6147.

Zukin, Cliff, Bob Carter, and Mark Schulman. 1987. "Trade-offs in Respondent Selection Methods: Theoretical Purity versus Real World Concerns." Paper presented at the annual meeting of the American Association for Public Opinion Research, Hershey, PA.

Appendix I: Studies included in meta-analysis

	Study	Selection Methods Used
1	Battaglia, Link, Frankel, Osborn, and Mokdad (2008)	Next Birthday, no selection
2	Beebe, Davern, McAlpine, and Ziegenfuss (2007)	Rizzio-Brick-Park, Next birthday
3	Carr and Hertvik (1993)	Birthday, Kish (simulated)
4	Czaja, Blair, and Sebestik (1982)	Kish, Bryant's Correction, Czaja-Blair-Sebestik modification
5	Denk and Hall (2000)	Last birthday, Kish, age-order, full enumeration
6	Denk, Guterbock, and Gold (1996)	Kish, last birthday, no selection
7	Forsman (1993)	Age-order, Groves and Kahn's Modification, Birthday
8	Grandjean, Leighty, Taylor, and Xu (2004)	Birthday, Targeted selection
9	Goyder, Basic, and Thompson (2001)	Next birthday, no selection
10	Hagen and Collier (1983)	Troldahl-Carter, Hagen-Collier
11	Hicks and Cantor (2012)	Next birthday, no selection
12	Hill, Donelan, and Frankel (1999)	Last birthday, Youngest Male Oldest Female
13	Keeter and Fisher (1997)	Last birthday, Youngest Male Oldest Female
14	Kennedy (1993)	Kish, Bryant's Correction, Hagen-Collier, Last Birthday, no selection
15	Lavrakas and Bauman (1993)	Last birthday
16	Le, Brick, Diop, and Alemadi (2013)	Le-Brick-Diop-Alemadi
17	Link and Mokdad (2006)	Troldahl-Carter, no selection
18	Oldendick, Bishop, Sorenson, and Tuchfarbr (1998)	Kish, Last birthday
29	Olson, Stange, and Smyth (2014)	Last birthday, Next birthday, Youngest adult, Oldest adult
20	Olson, Smyth, Stange, and Lavrakas (2015)	Last birthday
21	O'Rouke and Blair (1983)	Kish, Last birthday
22	O'Rouke and Lakner (1989)	Last birthday
23	Praire Research Associates (2001)	Next Birthday, no selection
24	Rizzo, Brick, and Park (2004)	Rizzo-Brick-Park
25	Sabin and Godley (1987)	
26	Srinivsan, Christiansen, and Tortora (1996)	Last birthday, full enumeration, Youngest Male Oldest Female
27	Stange, Smyth, and Olson (2015)	Last birthday
28	Zukin, Carter, and Schuman (1987)	Last birthday, no selection

Appendix II: Impact of Within-Household Respondent Selection Methods on Other Demographic Representation

	Number of Estimates	Effect Size (Absolute Bias) Mean	SE
Education Proportions			
Probability Methods	4	5.3%	0.7%
Birthday Methods	46	7.0%	1.3%
Non-Probability Methods	27	4.6%	1.3%
No Selection	23	10.7%	1.6%
Marital Status Proportions			
Probability Methods	6	4.7%	1.2%
Birthday Methods	13	3.0%	0.3%
Non-Probability Methods	11	3.3%	0.7%
No Selection	5	8.6%	1.7
Race Proportions			
Probability Methods	2	7.3%	1.3%
Birthday Methods	50	5.4%	0.6%
Non-Probability Methods	36	5.0%	1.0%
No Selection	14	10.3%	1.5%
Employment Proportions			
Birthday Methods	8	2.7%	1.0%
Non-Probability Methods	4	1.6%	0.8%
No Selection	4	3.9%	0.4%
Income Proportions			
Birthday Methods	41	4.6%	0.3%
Non-Probability Methods	22	7.4%	3.5%
No Selection	13	10.1%	4.0%